

Angeles, L. (2019) On the nature of banks. *Kyklos*, 72(3), pp. 381-399.

There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

This is the peer reviewed version of the following article:

Angeles, L. (2019) On the nature of banks. *Kyklos*, 72(3), pp. 381-399. (doi:10.1111/kykl.12204)

This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving.

http://eprints.gla.ac.uk/190972/

Deposited on: 25 July 2019

On the Nature of Banks

Luis Angeles*
University of Glasgow

Abstract

Two views exist regarding the nature of the banking business. The dominant view defines banks as financial intermediaries - institutions in the business of transferring money from savers to borrowers. An alternative view advances that banks finance borrowers via money creation. I explain the differences between these two views and argue for the superiority of the latter one as a description of modern banking. I discuss the implications for economic analysis and explain how the connection between bank lending and money creation helps us understand the effects of banking on economic activity.

^{*} Adam Smith Business School (Economics), University of Glasgow. University Avenue, Glasgow G12 8QQ, United Kingdom. Email: luis.angeles@glasgow.ac.uk

1. Introduction

Economists define banks as financial intermediaries. A financial intermediary is an institution in the business of collecting money from savers and transferring it to borrowers. The money is collected by issuing liabilities to the money-saving public and loaned out against liabilities issued by borrowers; the financial intermediary thus becoming debtor to the first one and creditor to the latter ones. Banks would differ from other financial intermediaries - such as mutual funds or insurance companies - only in the specific types of liabilities which they are able to issue and receive. Such differences are often regarded as secondary, and most standard analyses of the financial system group banks and other financial institutions together into a financial intermediaries sector.

In the context of real economic models, models which abstract from the existence of money, representing the business of collecting and transferring money would be, by definition, impossible. Accordingly, banks are incorporated into such models as institutions in the business of transferring real resources from savers to borrowers. The approach is meant to capture the economic role of banks and other financial institutions at some fundamental level, and is the dominant method for the economic analysis of banks in the macroeconomic literature.

This paper challenges the validity of viewing banks as financial intermediaries. In its place, I advance that banks are better viewed as institutions in the business of financing borrowers via money creation. The distinction between these two views may be blurred by the fact that the standard view of banks as financial intermediaries also ascribes to them the capacity to create money. In the standard view, however, money creation appears as a by-product of the banking business; banks are assumed to have no control over it. The alternative view of banks put forward here is different in that banks may create money independently of the actions of the general public. In a very meaningful sense, banks are in control of the process of money creation.

While the view of banks as financial intermediaries dominates academic circles, this has not always been the case. Little of what is to follow would have surprised some of the brightest economic minds of the first half of the twentieth century - including Irving Fisher, John Maynard Keynes and Joseph A.

Schumpeter.¹ Furthermore, the view that banks finance borrowers via money creation is found often enough in policy circles - most particularly among Central Bank officials. Just focusing on the last few years, the reader may refer to McLeay et al. (2014) and Jakab and Kumhof (2015) for work coming from the Bank of England, Unger (2016) and Aldasoro and Unger (2017) for work coming from the German Bundesbank, and, perhaps most notoriously, King (2016) and Turner (2016) for the views of the former Governor of the Bank of England and the former Chairman of the Financial Services Authority of the United Kingdom. Of course, none of these works ought to be taken as the official view of the institutions in which their authors were employed; but rather as indication of a debate of ideas taking place in them. The present paper wishes to engage more academic economists into this debate.

The next three sections of the paper discuss what is wrong with our current view of banking, and how it should be amended, at three levels of analysis: banking viewed from the perspective of the individual institution, banking at the aggregate level, and banking in relation to the Central Bank. These sections may also be read as a blueprint for how to teach money and banking, as they repeatedly refer to current teaching practice in the area. This is then followed by two sections in which the role of the banking system as an economic actor is interpreted, and this interpretation is contrasted with the way in which banking is incorporated into economic models. If the first three sections are of most immediate relevance to the teaching of economics, these last two derive the implications for research, and advocate a future direction for economic analysis in the area. A final section offers some concluding remarks.

2. Banking at the level of the individual institution

The business of banking is routinely introduced to students of economics using a series of accounting operations which I report in figure 1. According to this version of events, the core business of banking consists in collecting currency deposits from the public and using most of that currency to make loans to borrowers. The currency is collected by issuing a bank liability which goes by the name of a bank deposit, and is given to the public in exchange for its currency. Next, most of this currency is transferred to a borrower against a liability issued by this last one, which is an asset for the bank and goes by the name of a bank loan (or, alternatively, a corporate bond or other type of security). These two steps are shown in panels (b) and (c) of figure 1, where the resulting changes in the bank's

-

¹ See Fisher (1922, chapter 3), Keynes (1930, pp. 20-21) and Schumpeter (1934, pp. 98-99). I also like to think that Adam Smith would have agreed - see Smith (1776, pp. 316-317).

balance sheet are represented. Albeit it appears obvious at this stage, I have noted that the bank depositor and the bank borrower are different persons, so that the first finances the second via the actions of the bank. Figure 1, then, describes the business of banking as financial intermediation.

[Figure 1]

As it turns out, the actions in figure 1 would be an acceptable description of banking as it took place in classical Greece and imperial Rome - what we may refer to as "ancient banking". Ancient banking is the business model which arises when currency is the only form of money in circulation - or, to be more precise, the only widely accepted medium of exchange. If only currency is money, banks would have no other method for financing borrowers than collecting this currency from the public and transferring it to borrowers. This was the case in classical antiquity, with currency taking the form of precious metal coinage.²

Ancient banking is to be distinguished from what I shall call "modern banking" - a business model which, in the Western world, has been in continuous existence since the late 12th century AD.³ Modern banking arises when bank deposits - that is to say, when a liability issued by a bank - begins to be accepted as a means of payment. This may be said to take place as soon as a seller of goods or services accepts as payment not a delivery of currency but a credit to a bank deposit in his name; a credit which the bank agrees to put in place against a simultaneous debit to the bank deposit of the corresponding buyer. When this operation becomes commonplace, bank deposits may be referred to as money.

When bank deposits become money, banks realize that borrowers may be financed not just by a transfer of currency, but by the creation of new bank deposits. As noted long ago by Abbott P. Usher, "Banking [by which he means *modern banking* as defined here] begins only when loans are made in

² For a detailed account of the history of banking see Bogaert et al. (1994). For a concise account of the history of both money and banking see Angeles (2019).

³ See De Roover (1974, p. 200) and Bogaert et al. (1994, pp. 83-84). Some aspects of modern banking had already been tried in ancient Rome - see Harris (2008, p. 189) and von Reden (2010, p. 118).

bank credit" (Usher 1934, p. 399). Where ancient banks would finance borrowers via monetary transfer, the modern bank finances borrowers via money creation.

The process of bank lending via money creation is represented in figure 2. Panel (a) of this figure shows an initial situation, with a bank holding a mix of assets and liabilities - the specific composition of each being, at this stage of the analysis, of no relevance. In panel (b) of this figure a borrower approaches the bank for a loan, which is approved. The bank issues a liability against itself and in favour of the borrower - a new bank deposit which appears on the liabilities side of the bank's balance sheet. In exchange, the borrower issues a liability against itself and in favour of the bank - the new bank loan on the assets side. Since the total amount of bank deposits in the aggregate banking system has increased, money has been created.

[Figure 2]

The importance of presenting the business of banking at the level of the individual institution using figure 2, rather than figure 1, is critical. Given their name and their historical origin, it is natural to believe that bank deposits can only be created following a physical deposit of currency at a bank; and therefore that banks cannot create bank deposits "out of nothing" as in figure 2. This notion is fallacious because a bank deposit is a liability, not a record for currency received. To be sure, the reception of currency will give rise to such a liability, but liabilities may be issued under a number of alternative scenarios - in particular, a bank may issue a liability in exchange for the liability of a borrower. This exchange of liabilities is what we call bank lending. Banks are not different from other firms and private individuals in their capacity to issue liabilities to whoever is willing to give something in return for them. Banks are only different in that their liabilities, unlike those of any other private entity, circulate as money.

To the mind of some teachers of economics, figure 2 is ultimately no different from figure 1 since, once the borrower spends the deposit granted to him, the bank's balance sheet will look the same under both figures. Indeed, if we were to follow panel (b) of figure 2 by a third panel where the deposit is spent, we would see that bank lending leads to no expansion of deposits on the liabilities side of this particular bank, and a decrease in reserves which matches the increase in bank loans on the assets

side. In other words, we would observe the same balance sheet changes as in panel (c) of figure 1, but the action would have been split in two stages. This, however, does not imply the two figures are equivalent.

Figure 2 is a correct description of how banks make loans, figure 1 isn't: no modern bank disburses loans using currency. Figure 2 makes clear that bank deposits may be created under circumstances other than the deposit of currency; figure 1 perpetuates the opposite idea. Figure 2, finally, advances that the business of banking is not dependent on the collection of a form of money which, in most advances economies, is of residual importance. To give a specific example, currency amounts to only 3% of the money supply in the United Kingdom (McLeay et al. 2014) - it would be exceedingly troublesome if banks were dependent on the willingness of the public to deposit this form of money with them in order to make further loans.

Given the above, it is worth asking ourselves why figure 1 is so often preferred to figure 2 in textbooks and university courses. In my opinion, this has much to do with the fact that figure 1 makes immediately clear that banks need to find a source of funds for the loans they make. Figure 1, however, renders this idea in a way which is at odds with reality - banks do not need to collect currency before making loans, let alone collect currency for a value equal to the loans they wish to make. As I will describe below, the way in which banks fund their loans in reality is best described by the banking model from figure 2, but only once we extend the analysis to the aggregate level.

Figure 1 may also offer a more reassuring vision of the business of banking, as banks are portrayed as merely the agents of depositors. Depositors decide on the level of financing available to borrowers, and banks are only the middlemen in the process. As we shall see, figure 2 leads into a different understanding of the economic role of banks, giving them far more agency in the process of financing borrowers than what figure 1 allows for.

Figure 1 is typically presented to students of economics as an adequate description of the business of banking as it is practised today. In reality, the type of banks it describes were becoming obsolete in Western Europe by the late Middle Ages. Figure 2 contains the core of a superior description of banking, albeit the analysis needs to be carried further before a well-rounded picture can emerge.

3. Banking at the aggregate level

Figure 2 did not incorporate the spending of the new bank deposit created for the bank borrower for a simple reason: the operation typically involves a second bank, and thus takes us into the analysis of banking at the aggregate level.

The business of banking at the aggregate level is understood to lead to money creation, following a process known by students of economics throughout the world as the money multiplier. The money multiplier process is presented in figure 3 under two guises. Most textbooks use the version in panel (a), which assumes that individual banks behave as described in figure 1. As is well-known, the initial deposit of currency into "Bank 1" gives rise to a sequence of loans and currency deposits across multiple banks in the system which are geometrically decreasing in size. Once the process runs its course, the total amount of new bank deposits in the system equals the amount of currency initially deposited times the inverse of the banks' reserve ratio - a.k.a. the money multiplier.

[Figure 3]

Panel (b) of figure 3 presents the same process assuming that banks disburse loans in the form of a newly-created bank deposit rather than currency - the banking model of figure 2.4 Once the deposit is spent, bank reserves migrate to another bank just as they did in figure 1, only without passing through the hands of the public in currency form. Note how the final balance sheet of "Bank 1" is the same under both versions of the money multiplier process. As was the case in panel (a), the transfer of reserves to "Bank 2" allows for a further loan, giving rise to the same structure of bank deposit expansions as previously.

It is worth noting that in both versions of the money multiplier process banks can never make loans for an amount larger than their holdings of excess reserves. Indeed, in both versions reserves are

_

⁴ See Mishkin (2016, p. 393) for this less common version of the money multiplier.

bound to fall by the full amount of the loan - immediately in the case of panel (a), and once the proceedings of the loan are spent in the case of panel (b).

My argument in this section is that the money multiplier process, whether based on lending by currency transfer or directly by deposit creation, is a very poor description of banking at the aggregate level. Its inadequacy is obvious once pointed out: the process assumes that banks in the economy must make loans sequentially. The second bank in the chain cannot make a loan before the first one, as it needs the reserves which this first one will provide in order to proceed. The same applies to all subsequent banks in the process.

In reality, of course, all banks in the economy make loans simultaneously - and they do not need to wait for reserves to be in place before the loans are made. The process is described in figure 4, which starts from the account of banking at the individual level presented in figure 2 and considers the consequence of all banks in the economy behaving in this way at the same time. Panels (a) and (b) of this figure simply reproduce the corresponding panels of figure 2; there are N banks instead of one, but this is of no consequence up to this point since banks have not yet interacted with each other. Without loss of generality, all banks are assumed to be equal.

[Figure 4]

The crucial part of the process comes in panel (c), as borrowers spend the money received. When they do so, the bank deposit created for their purpose is transferred to another bank in the system, together with an equal amount of bank reserves. As we assume perfect symmetry, each bank will transfer bank deposits to other banks in the system and receive bank deposits from them for the same amount, leading to a net change in bank deposits equal to zero. The same is of course true for bank reserves. As a consequence, the balance sheets of banks do not change between panels (b) and (c) when we look at their overall holdings of reserves or bank deposits - but the identity of the deposit holders does change. In panel (b), all new loans were matched by new deposits belonging to the people who had acquired the loans - bank debtors and bank creditors were the same person. In panel (c), we have reached the recognizable stage where bank debtors and bank creditors are different people – the loan given to Mr. A may be said to be funded by the deposits held by Mrs. X, Y and Z.

Banks, in other words, have found a source of funds for the loans they have made, but this has taken place after the granting of the loans, not before. Bank loans are ultimately funded by the public's willingness to accept bank liabilities as money.

Figure 4 brings to light a number of important features about the banking business which are only visible at the aggregate level. The business of banking consists in financing borrowers via money creation, but this has to be followed by a capacity to attract funds once the money created has been spent by the borrowers. This last assertion does not undermine the fact that banks increase the aggregate amount of funds available whenever they grant a loan, but they still need to compete with the rest of the banks for the funds in existence. Individual banks have funding needs even though their loans are initially funded via money creation: the money they create will leave their books and will need to be gained back against other suitors. The aggregate banking system, on the other hand, has no such financing worries.⁵

Figure 4 assumes that all banks expand their lending portfolio in the same proportion and that, given their share of the banking business, they are all able to fund their loans with incoming deposits. In reality, some banks may choose to run ahead of the industry average and lend more if they believe they will be able to capture an equally large proportion of the deposits being created. Failure to capture enough deposits will lead to the loss of reserves, and eventually to the cost of borrowing reserves from other banks in order to stay within the limits set by prudence or by regulatory authorities. In all cases, however, banks are not constrained to make loans not exceeding in value their current holdings of bank reserves - contrary to what was assumed in both panels of figure 3. Banks can perfectly well make loans for a value in excess of their reserves, provided that subsequent reserve inflows will compensate the reserve outflows this will produce. Notice that, in figure 4, nothing is said about the magnitude of the new loans made with respect to existing reserves.

-

⁵ I shall add that, although bank loans and bank deposits expand by the same magnitude initially, the public may exchange some of the deposits created for non-monetary liabilities issued by banks later on. Such a mechanism would explain the divergence between bank loans and broad money in most developed nations over the post-war period (Schularick and Taylor 2012, p. 1036). This consideration does not invalidate any of the arguments in the present paper since such actions can only take place later on, once borrowers have spent the money created for them.

This last point leads us into considering the limit to the amount of money the banking system is able to create when granting loans. In figure 4, assuming a required ratio of reserves to bank deposits of 10%, banks would not be able to grant new loans beyond the point where aggregate bank deposits equal 10 times aggregate reserve holdings. An individual bank may attract reserves from its peers in order to create more money, but at the aggregate level the total amount of money creation will be constrained by this limit.

This limit to the amount of money that banks may create for a given level of aggregate reserves happens to coincide with the total amount of money created under the money multiplier process. This may be reassuring, but it should not be taken as an argument in favour of the money multiplier as a teaching device. The money multiplier introduces an imaginary structure to the expansion of loans and deposits which simply is not there. Figure 4, on the other hand, advances that bank lending can expand by anything between zero and the maximum level that bank reserves allow for, and that this expansion can take place immediately rather than through a protracted sequence of increasingly small steps.

The money multiplier process assumes that bank lending will proceed in an automatic way - banks will inevitably want to lend the currency or reserves received, as failure to do so would imply they hold an asset whose rate of return does not cover the cost of the corresponding liability. Figure 4, on the other hand, makes clear there is nothing automatic about the expansion of bank lending. Bank lending creates a new asset and a new liability, and banks will proceed with the operation only if the expected return from the asset exceeds the cost of the corresponding liability.

Banks may well decide not to go ahead with new lending even when reserves are plentiful, as policy makers found out to their apparent surprise following the Global Financial Crisis of 2008. Incidentally, the last 10 years of monetary policy experience offer as stark a rejection of the empirical validity of the money multiplier process as we are ever likely to observe.

The money multiplier mechanism is a rather complicated and fictitious process which does students of economics a disservice in their attempts to understand how banks create money. Figure 4 offers a representation of the process of money creation by banks which is both simpler and far more accurate.

4. Commercial Banks and the Central Bank

The last section has confirmed that, in accordance with conventional views, the maximum amount of money which the banking system is able to create is a multiple of the aggregate amount of reserves in the system. As these reserves are nothing but liabilities of the Central Bank, either currency or Central Bank deposits, textbooks dealing with the subject have traditionally concluded that the money supply of the nation is ultimately under the control of the Central Bank. Students of economics are told the Central Bank increases or decreases these reserves via open market operations - purchases or sales of assets held by commercial banks. When bank reserves increase, the money supply increases - in the standard story this happens automatically via the money multiplier process, but the same endpoint would be reached if we adopt the framework of figure 4 and assume that banks exhaust their lending capacity. The principle is then incorporated into most standard macroeconomic models to which students are likely to be exposed - within these models, the money supply is a variable exogenously set by the Central Bank.

Oddly enough, and as is well-known by teachers of the above material, no Central Bank in the world sets up as its policy objective any given level or growth rate of the money supply. Instead, Central Banks define a desired level of the short-term interest rate on the interbank market for Central Bank reserves, and take the necessary steps to see it realized. Open market operations may well be part of these necessary steps, but they are tools used to reach a certain goal rather than a goal in themselves.

As noted long ago by Charles Goodhart, a Central Bank setting the aggregate level of bank reserves has rarely, if ever, been more than a theoretical possibility. If implemented, the policy would give rise to extreme instability in the interest rate on reserves: on each day on which reserve requirements are computed, the overnight rate will fall to zero if aggregate reserves exceed the total amount required by all banks given existing deposits, and will rise to some penalty rate in the opposite case (Goodhart 1994). Central Banks have concluded, sensibly, that such instability would be undesirable.

Even more important, and as the preceding sections have made clear, banks do not sit passively waiting for the arrival of new deposits which, by increasing their excess reserves, would allow them to make additional loans. Instead, banks make new loans by creating new deposits - pushing their ratio

of reserves to deposits downwards in the process. An individual bank may try to stop this ratio from falling by obtaining additional reserves from other banks, but this option is not available at the aggregate level. If the aggregate banking system is extending its loan portfolio, it will need to buy additional reserves from the Central Bank. Under such circumstances, how is the Central Bank to impose any quantitative limit on bank reserves? Denying additional reserves to commercial banks does not appear to be a viable option, as it would stop all new lending. Overdrafts and approved lines of credit would not be able to proceed; economic activity would be compromised. Understandably, no Central Bank wishes to take responsibility for such a policy choice - nor is it likely that it would be allowed to follow it.

As a consequence, Central Banks target the price on bank reserves rather than their quantity: they stand ready to supply the banking system with reserves on demand at a certain price, which then determines the price that banks will charge each other. The relationship between the money supply and bank reserves thus works in the exact opposite direction to that assumed in most textbooks. The private banking system and bank borrowers jointly decide on how much money will be created by entering into loan agreements with each other, which creates a need for additional bank reserves. These will be supplied by the Central Bank on demand, as failure to do so would render its own interest rate policy unworkable. Causality runs from the private banking system to the Central Bank - albeit, to be sure, the target interest rate imposed by the Central Bank is one among several factors influencing bank lending. This degree of influence notwithstanding, it is far more accurate to say that bank reserves endogenously adapt to the requirements of the private sector rather than the other way around.⁶

The above sequence of events is incorporated into the framework of figure 4 by simply adding one additional operation - giving us figure 5. The first three panels of figures 5 and 4 are essentially identical: figure 5 only makes explicit the initial composition of assets and liabilities for all banks, as we will now be concerned with the ratio of reserves to bank deposits. Following the granting of bank loans via deposit creation, and the spending of these deposits, banks in panel (c) find themselves with insufficient reserves: the ratio of reserves to bank deposits is below 10% for all of them. Left to themselves, they would all fall foul of financial regulations, and their willingness to avoid the

_

⁶ For more discussion on this point see Goodhart (2009, 2017). The point has been made numerous times by Central Bank officials, see Holmes (1969, p. 73), King (1994, p. 264) and McLeay et al. (2014, p. 21) for some examples.

corresponding penalties would drive the interest rate on reserves skyward. The banks thus approach the Central Bank in order to exchange some of the assets they hold for additional reserves, and the Central Bank agrees. The resulting balance sheets are shown in panel (d), with all banks now holding reserves in excess of the minimum required by law.

[Figure 5]

Figure 5 completes our description of the business of banking at the aggregate level. In contrast to the essentially passive role ascribed to them in the money multiplier process, banks are on the driving seat of the process of money creation from start to finish. Banks expand the money supply by creating bank deposits when they make loans, and back up these new deposits with additional reserves they acquire using some of the assets they hold - assets which were themselves bought via money creation.

As we see, the capacity of commercial banks to create money is not under the direct influence of the Central Bank - the Central Bank exercises only an indirect influence via interest rates. Banks are constrained in their money creation abilities not so much by public policy, but by two other forces which derive from the analysis above. First, at the level of the individual banking institution, banks are constrained by the need to keep in line with the rest of their peers or, what comes to the same thing, by the need to find sources of funds when the deposits they have created leave their books. This was well understood by John Maynard Keynes when, in his *Treatise on Money*, he affirmed that "there is no limit to the amount of bank money which the banks can safely create *provided that they move forward in step*" (Keynes 1930, p. 23; italics in the original). Second, however, and notwithstanding Keynes, money creation is also constrained at the level of the aggregate banking system by the willingness of the public to acquire more debt. Banks will not be able to expand their lending if they cannot find takers for their loans. The amount of money in circulation is thus endogenous to the decisions taken by the private sector - private banks and the borrowing public - and far from being something which the Central Bank or the government can control.

5. Banks and Economic Models

When it comes to banking, macroeconomic models over the post-war period have fallen into two vast categories. First, most models from the 1950s to the 1980s simply do not feature banks - consider

the IS/LM and AS/AD models of undergraduate courses, the standard DSGE framework of postgraduate courses, and pretty much all of growth theory. This type of models continues to be prominent after the 1980s, but by the end of that decade it is joined by a second category: macroeconomic models that do feature banks, but do not feature money. The financial accelerator literature (Bernanke and Gertler 1989, Kiyotaki and Moore 1997) is the most prominent example of this second case.

The explanation for the first approach lies in the view of banks as financial intermediaries. An intermediary only carries out the orders set to him by other agents, and as such may be considered of secondary importance. Macroeconomic models with no banks implicitly assume that modelling the saving decisions of households and the investment decisions of firms renders banks largely irrelevant.

The second approach arises once we assume there exist some imperfections in the process of financial intermediation – imperfections which are substantial enough for banks to be modelled explicitly. Unfortunately, the modelling of banks is rendered problematic by the long-standing assumption that real economy models are the tools of choice to analyse economic phenomena at a fundamental level. Financial intermediation – the transfer of money from savers to borrowers – cannot be represented in a model without money. Macroeconomists have avoided this impasse by interpreting banks as intermediaries in real resources: institutions whose actions ultimately lead to a transfer of real resources from people who want to save to people who want to borrow. Is either of these two approaches justifiable in the light of our discussion over the preceding sections of this paper?

From the outset, we may say that the first approach – ignoring banks – is simply not tenable. As anticipated in the introduction, banks are not mere financial intermediaries but rather institutions in the business of financing borrowers via money creation. They do not take their orders from the public or from bank depositors, but carry out lending according to their own economic interests. The interests of banks are not perfectly aligned with those of borrowers or the general public, which is why their behaviour needs to be analysed explicitly within macroeconomic models.

In order to assess the second approach, we need to consider the implications of banking on real resources at the aggregate level. Could financing via money creation lead to a transfer of real resources from savers to borrowers?

To be sure, borrowers receiving a loan are being granted access to real resources from the economy. It does not follow, however, that these resources are taken away from bank deposit holders, who do not lose access to any money during the process. Since borrowers are financed via money creation, their funds do not come at the expense of anybody else. Something other than a transfer of resources between bank depositors and bank borrowers must be taking place.

If we chose to think of aggregate resources as exogenously given, an attractive interpretation for the above phenomenon would be that the resources to which borrowers gain access to come at the expense of the public at large or, to be more precise, at the expense of all money holders. With real resources given, more money in circulation would translate into higher prices - thus reducing the consumption capacity of holders of money and freeing up resources for bank borrowers. In this case, banking gives rise to a transfer of real resources, but from the general public to borrowers.

More realistically, we may think of aggregate resources as endogenously determined by aggregate demand. Bank lending increases aggregate demand in nominal terms as it gives borrowers the capacity to buy goods and services without reducing anyone else's capacity to do the same (once again, money creation). Aggregate production would be expected to increase under these conditions, which in turn implies that at least part of the resources which borrowers gain access to do not come at the expense of the non-borrowing public. If such is the case, the flows of real resources resulting from bank lending do not correspond to a transfer at all.

Crucially, in neither of the two scenarios just outlined are banks mere intermediaries in real resources. Instead, banks approve the access to real resources by borrowers without the consent of the general public. The public may benefit from the actions of banks if they lead to increased investment and activity and may suffer from them if they lead to speculation and turmoil; but they have no direct influence over them. Banks do not act as agents of the public, and do not transfer resources allocated to them.

To conclude, our current approach to modelling banks is as much in need of revision as our current approach to their teaching. Banks should not be ignored in macroeconomic models and, when present, should not be incorporated as intermediaries in real resources. The next section illustrates the negative consequences of modelling banks as intermediaries in real resources by considering an area of research which has received a vast amount of attention since the Global Financial Crisis of 2008: the relationship between banking and economic recessions.

6. Banks and Economic Recessions

The Global Financial Crisis of 2008 took most economists by surprise. The main tool for thinking about how the financial sector and the real economy interact was, and to a large extent still is, the financial accelerator literature. Within this literature, imperfections in the intermediation process mean that banks may act as amplifiers of disturbances taking place elsewhere in the economy. In the aftermath of the events of 2008, this feature rendered such models questionable: disturbances originating outside the financial system were hard to identify. Not only that, but most commentators would place the financial system itself at the origin of the problems leading to the Global Financial Crisis - a far cry from the mere amplifying role proposed by the academic profession.

Like most real economy models, models from the financial accelerator literature put emphasis on the supply side of the economy to explain disturbances in aggregate activity. Any effects of banking on aggregate demand will be difficult to envision precisely because banks are conceived as intermediaries in real resources. If banks reallocate real resources, the loss in demand from savers will be, to a first approximation, compensated by the increase in demand from borrowers. The repayment of bank loans would result in flows of real resources in the opposite direction, with equally muted consequences. Departures from this baseline result are possible, but require substantial elaboration.⁷ As a consequence, most of the theoretical work on the interaction between banking and economic activity during recessions abstracts from aggregate demand effects, at a time when some empirical findings strongly support the importance of this precise channel.⁸

⁷ For an example, see Eggertsson and Krugman (2012).

⁸ See, in particular, Mian and Sufi (2010, 2014).

In stark contrast, the effects of banking on economic activity via aggregate demand fluctuations arise almost effortlessly if we view banks as institutions financing borrowers via money creation, and analyse their actions in a model featuring money. As advanced in the previous section, bank lending increases aggregate demand by giving borrowers access to funds without reducing the funds available to the rest of the public. This process, however, leaves something else behind: debt. The rapid accumulation of private debt was perhaps the most salient characteristic of the build-up towards the Global Financial Crisis of 2008; a phenomenon which, some few exceptions aside, left most economists unconcerned. Debt is not a problem in models which view banks as intermediaries of real resources, but it becomes a major risk if we understand how debt flows give rise not only to money creation - but also to money destruction.

The repayment of bank debts reduces money in circulation as it triggers the balance sheet changes described in figure 2 but in reverse: a bank loan is eliminated from the assets side of a bank's balance sheet together with an equal amount of bank deposits. As the debtor needs to abstain from consumption in order to make these bank deposits available for this purpose, his contribution to aggregate demand decreases. And since no member of the public receives the money being repaid, which simply disappears from circulation, the overall effect on aggregate demand is negative.

Under normal circumstances, new bank lending exceeds repayments to existing bank debt so that the net effect on aggregate demand is positive. The accumulation of debt, however, creates the potential for vast declines in aggregate demand if new bank lending slows down or comes to a halt while repayments to existing debt are due. The accumulation of bank debt is thus an important risk factor on the stability of aggregate demand, and therefore on economic activity. The drag on aggregate demand would not wear out in a year or two, but have a protracted time pattern as medium and long-term loans become due. In other words, the economy could see several years of negative or weak growth while it deals with the consequences of excessive bank lending. At the same time, the prospect of several years of slow growth would only reaffirm banks' decision to cut out new lending in the first place. Bank lending and the subsequent process of debt deleveraging would thus produce a pattern of boom and bust in economic activity, which clearly parallels the experience of multiple economies before and after a crisis.

Finally, under the analysis above, the implications for economic policy in response to a debt-fuelled financial crisis are far-reaching. With bank debt viewed as a transfer of real resources debt repayments have no negative macroeconomic consequences - and should be enforced on moral grounds, as it is only fair to return the resources initially ceded by the public. If bank debt is the outcome of money creation, however, a different perspective emerges. Bank lending does not necessarily transfer real resources from the public to borrowers, as aggregate production may itself increase due to bank lending. And debt repayments do not return anything to the public - they take money out of circulation. Debtors should of course be expected to pay back their debts under normal circumstances, but the macroeconomic consequences of such actions should be factored in during periods of aggregate debt deleveraging. Insisting on prompt debt repayments is not necessarily wise if it comes at the expense of aggregate economic activity. Softening conditions on debtors via debt restructuring and partial bailouts, accompanied with adequate measures to sustain the viability of the banking system, is thus a valid policy option under the specific economic environment of high indebtedness and low demand. By and large, this has not been the policy of choice when dealing with the aftermath of the crisis in the years since 2008.

7. Concluding Remarks

The conflict between the two views about the nature of banking here described runs deep in the history of economics, and closely relates to the conflict between the two views about the nature of money, as discussed by Goodhart (1998). Towards the end of his monumental History of Economic Analysis, Joseph A. Schumpeter reflects on these two approaches to banking as they existed towards the middle of the 20th century. While in most other areas of economics Schumpeter would abstain from expressing too marked preferences, he makes an exception at this point and chooses his words rather uncompromisingly:

"It is much more realistic to say that the banks 'create credit', that is, that they create deposits in their act of lending, than to say that they lend the deposits that have been entrusted to them. [...] The theory of 'credit creation' not only recognizes patent facts without obscuring them by artificial constructions; it also brings out the peculiar mechanism of saving and investment that is characteristic of fullfledged capitalist society and the true role of banks in capitalist evolution."

Schumpeter (1954, p. 1114)

In the decades since this opinion was expressed, it has only become more and more pertinent.

Banks should not be omitted from any discussion of the economy at the aggregate level which concerns the financing of borrowers, as most borrowers in most economies are financed by money creation from banks, rather than by money transfers from financial intermediaries. And, for this same reason, an economic analysis of banking cannot take place within a model economy where money is absent.

All models simplify reality, but good models choose wisely which aspects of reality may be left out, and which shouldn't. Trying to think about banks without thinking about money is not only ill-advised; to a large extent it is self-contradictory. The capacity to issue liabilities which function as money is, more than anything else, what defines the modern bank.

References

Aldasoro, I. and Unger, R., 2017, External financing and economic activity in the euro area – Why are bank loans special? Discussion Paper N. 4/2017, Deutsche Bundesbank.

Angeles, L., 2019, Four Phases in the History of Money, Discussion Paper 2019-01, University of Glasgow - Adam Smith Business School.

Bernanke, B. and Gertler, M., 1989, Agency costs, net worth, and business fluctuations, American Economic Review 79 (1), 14-31.

Bogaert, R., Kurgan-van Hentenryk, G. and van der Wee, H., 1994, A History of European Banking, Antwerp: Fonds Mercator Paribas.

De Roover, R., 1974, Business, Banking, and Economic Thought in Late Medieval and Early Modern Europe, Chicago: The University of Chicago Press.

Eggertsson, G. B. and Krugman, P., 2012, Debt, develoraging, and the liquidity trap: A Fisher-Minsky-Koo approach, Quarterly Journal of Economics 127 (3), 1469-1513.

Fisher, I., 1922, The purchasing power of money, its determination and relation to credit, interest and crises, New York: The Macmillan Press.

Goodhart, C. A. E., 1994, What should Central Banks do? What should be their macroeconomic objectives and operations?, The Economic Journal 104, 1424-1436.

Goodhart, C., 1998, The two concepts of money: implications for the analysis of optimal currency areas, European Journal of Political Economy 14, 407-432.

Goodhart, C. A. E., 2007, The Continuing Muddles of Monetary Theory: A Steadfast Refusal to Face Facts, Economica 76, Supplement 1, 821-830.

Goodhart, C. A. E., 2017, The Determination of the Money Supply: Flexibility versus Control, The Manchester School 85 (S1), 33-56.

Harris, W. V., 2008, The nature of Roman money, in: Harris, W. V. (ed.), The monetary systems of the Greeks and Romans, Oxford: Oxford University Press.

Holmes, A., 1969, Operational constraints on the stabilization of money supply growth, Proceedings of the conference "Controlling Monetary Aggregates", Federal Reserve Bank of Boston, June 1969.

Jakab, Z. and Kumhof, M., 2015, Banks are not intermediaries of loanable funds – and why this matters, Working Paper N. 529, Bank of England.

Keynes, J. M., 1930 [reprinted 1971], A Treatise on Money, The collected writings of John Maynard Keynes, volume V, London: The Macmillan Press.

King, M., 1994, The transmission mechanism of monetary policy, Bank of England Quarterly Bulletin 1994 Q3, 261-267.

King, M., 2016, The End of Alchemy: money, banking and the future of the global economy, New York: W. W. Norton.

Kiyotaki, N. and Moore, J., 1997, Credit cycles, Journal of Political Economy 105 (2), 211-248.

McLeay, M., Radla, A. and Thomas, R., 2014, Money creation in the modern economy, Bank of England Quarterly Bulletin 2014 Q1, 14 -- 27.

Mian, A. and Sufi, A. 2010, Household Leverage And The Recession Of 2007 To 2009, IMF Economic Review 58, 74-117.

Mian, A. and Sufi, A. 2014, What Explains the 2007-2009 Drop in Employment?, Econometrica 82 (6), 2197-2223.

Mishkin, F., 2016, The Economics of Money, Banking and Financial Markets, 11th edition, Harlow: Pearson Education.

Schularick, M. and Taylor, A. M., 2012, Credit booms gone bust: monetary policy, leverage cycles, and financial crises, 1870-2008, American Economic Review 102 (2), 1029-1061.

Schumpeter, J. A., 1934, The theory of economic development, London: H. Milford.

Schumpeter, J. A., 1954, History of Economic Analysis, London: Allen & Unwin.

Smith, A., 1776 [reprinted 1976], An Inquiry into the Nature and Causes of the Wealth of Nations, Chicago: University of Chicago Press.

Turner, A., 2016, Between Debt and the Devil: money, credit, and fixing global finance, Princeton: Princeton University Press.

Unger, R., 2016, Traditional banks, shadow banks and the US credit boom – credit origination versus financing, Discussion Paper N. 11/2016, Deutsche Bundesbank.

Usher, A. P., 1934, The Origins of Banking: The Primitive Bank of Deposit, 1200-1600, The Economic History Review 4 (4), 399 - 428.

Von Reden, S., 2010, Money in Classical Antiquity, Cambridge: Cambridge University Press.

Figure 1
Bank lending via money transfer

Panel (a): Initial situation



Panel (b): Reception of currency

Bank				
Reserves (currency)	100	Bank deposit (of Mrs. Z.)	100	

Panel (c): Lending of currency to borrowers

Bank					
Reserves	10	Bank deposit	100		
(currency)		(of Mrs. Z.)			
Bank loan (to Mr. A.)	90				

Figure 2
Bank lending via money creation

Panel (a): Initial situation

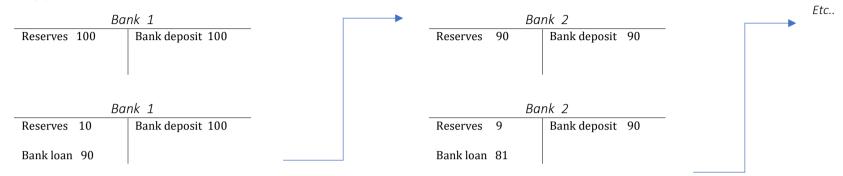
Bank					
Assets	XX	Liabilities	XX		
(incl. reserves)					

Panel (b): Making loans via money creation

Bank				
Assets (incl. reserves)	XX	Liabilities	XX	
Bank loan (to Mr. A)	100	Bank deposit (of Mr. A)	100	

Figure 3 The money multiplier mechanism: two versions

Panel (a): standard version



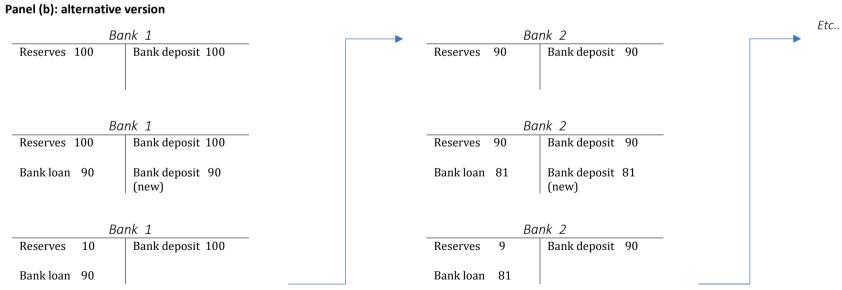


Figure 4
The business of banking at the aggregate level

Panel (a): Initial situation

Ва	nk 1			Ва	nk 2		
Assets xx (incl. reserves)	Liabilities	xx	Assets (incl. reser	xx ves)	Liabilities	XX	

Panel (b): Making loans via money creation

Bank 1		Bank 2		
Assets xx	Liabilities xx	Assets xx	Liabilities xx	
(incl. reserves)		(incl. reserves)		
Bank loan 100 (to Mr. A.)	Bank deposit 100 (of Mr. A.)	Bank loan 100 (to Mr. B.)	Bank deposit 100 (of Mr. B.)	

Panel (c): Spending of the loans

Bank 1		Bank 2		
Assets xx	Liabilities xx	Assets xx	Liabilities xx	
(incl. reserves)		(incl. reserves)		
Bank loan 100 (to Mr. A.)	Bank deposit 100 (of Mr/Mrs. X., Y., Z.)	Bank loan 100 (to Mr. B.)	Bank deposit 100 (of Mr/Mrs. X., Y., Z.)	

Bank N				
Assets	XX	Liabilities	XX	
(incl. reser	ves)			

Bank N				
Assets xx	Liabilities xx			
(incl. reserves)				
Bank loan 100 (to Mr. N.)	Bank deposit 100 (of Mr. N.)			

Bank N					
Assets xx Liabilities xx					
(incl. reserv	es)				
Bank loan 1	00	Bank deposit 100 (of Mr/Mrs. X., Y., Z.)			
(to Mr. N.)		of Mr/Mrs.	. X., Y., Z.)		

Figure 5
The business of banking at the aggregate level: obtaining bank reserves.

Panel (a): Initial situation

B	ank 1	_	Ва	nk 2
Reserves 50 Other assets 450	Bank deposits 500	Reserves Other assets	50 3 450	Bank deposits 500

Panel (b): Making loans via money creation

Bank 1		Bank 2		
Reserves 50	Bank deposits 500	Reserves 50	Bank deposits 500	
Other assets 450		Other assets 450		
Bank loan 100 (to Mr. A.)	Bank deposit 100 (of Mr. A.)	Bank loan 100 (to Mr. B.)	Bank deposit 100 (of Mr. B.)	

Panel (c): Spending of the loans

Bank 1		Bank 2		
Reserves 50 Other assets 450	Bank deposits 500	Reserves 50 Other assets 450	Bank deposits 500	
Bank loan 100 (to Mr. A.)	Bank deposit 100 (of Mr/Mrs. X., Y., Z.)	Bank loan 100 (to Mr. B.)	Bank deposit 100 (of Mr/Mrs. X., Y., Z.)	

Panel (d): Obtaining more reserves

Bank 1		Bank 2		
Reserves 100	Bank deposits 500	Reserves 100	Bank deposits 500	
Other assets 400		Other assets 400		
Bank loan 100 (to Mr. A.)	Bank deposit 100 (of Mr/Mrs. X., Y., Z.)	Bank loan 100 (to Mr. B.)	Bank deposit 100 (of Mr/Mrs. X., Y., Z.)	

Reserves 50 Bank deposits 500
Other assets 450

Reserves 50 Bank deposits 500
Other assets 450
Bank loan 100 Bank deposit 100 (of Mr. N.)

Reserves 50 Bank deposits 500
Other assets 450
Bank loan 100 Bank deposit 100
(to Mr. N.) (of Mr/Mrs. X., Y., Z.)

Reserves 100 Bank deposits 500

Other assets 400

Bank loan 100 Bank deposit 100 (of Mr/Mrs. X., Y., Z.)