BANK OF ENGLAND

Staff Working Paper No. 761

Banks are not intermediaries of loanable funds — facts, theory and evidence

Zoltan Jakab and Michael Kumhof

June 2019

This is an updated version of the Staff Working Paper originally published on 26 October 2018

BANK OF ENGLAND

Staff Working Paper No. 761

Banks are not intermediaries of loanable funds — facts, theory and evidence

Zoltan Jakab[(1)] and Michael Kumhof[(2)]

Abstract

In the loanable funds model, banks are modelled as resource-trading intermediaries that receive deposits of physical resources from savers before lending them to borrowers. In the financing model, banks are modelled as financial intermediaries whose loans are funded by *ex-nihilo* creation of ledger-entry deposits that facilitate payments among nonbanks. The financing model predicts larger and faster changes in bank lending and greater real effects of financial shocks. Aggregate bank balance sheets exhibit very high volatility, as predicted by financing models. Alternative explanations of volatility in physical savings, net securities purchases or asset valuations have almost no support in the data.

Key words: Banks, financial intermediation, loanable funds, money creation, bank lending, bank financing, money demand.

JEL classification: E41, E44, E51, G21.

(1) International Monetary Fund. Email: zjakab@imf.org
(2) Bank of England. Email: michael.kumhof@bankofengland.co.uk

The views expressed in this paper are those of the authors and should not be attributed to the Bank of England or the International Monetary Fund. This is a comprehensively revised draft of our 2015 *Bank of England Staff Working Paper No. 529*, 'Banks are not intermediaries of loanable funds — and why this matters'. The authors are grateful for many comments and helpful suggestions from seminar participants at Autorité de Contrôle Prudentiel et de Résolution, Banco Central do Brasil, Banco de España, Bank of Finland, Bank of Italy, Bank of Korea, Bank of Lithuania, Banque de France, BIS, Bundesbank, Central Bank of Ireland, Central Bank of the Russian Federation, CEPR-IMFS Macro-Modeling Conference, CER-ETH Zürich, Czech National Bank, ECB, Econometric Society World Congress, Her Majesty's Treasury, Institute of International and European Affairs, International Labor Organization, Kingston University, Money Macro and Finance Conference, National Bank of Poland, Österreichische Nationalbank, Oxford-New York Fed Conference on Monetary Economics, Princeton JRCPPF Conference, Royal Economic Society Annual Conference, Sciences Po, Sheffield University, Stockholm School of Economics, Universidad Carlos III, University of Essex, University of Geneva, University of Leeds, University of Nottingham, University of Saint Andrews, University of Surrey and University of Zürich.

The Bank's working paper series can be found at www.bankofengland.co.uk/working-paper/staff-working-papers

Bank of England, Threadneedle Street, London, EC2R 8AH
Telephone +44 (0)20 3461 4030 email publications@bankofengland.co.uk

ISSN 1749-9135 (on-line)

1. Introduction

In the wake of the 2007/8 global financial crisis (GFC) policymakers have paid more attention to the macroeconomic role of banks than at any time since the Great Depression. Macroeconomics was initially only able to offer limited analytical support because banks, with few exceptions, had been ignored in theoretical and empirical work. This has now changed, with a new literature that has made many valuable contributions. However, the conceptual and modeling framework that is almost universally used in this literature is the intermediation of loanable funds (ILF) model, and a critical feature of this model is that it adopts a shortcut that describes banks as intermediaries of physical resources akin to warehouses. We will demonstrate below that this is indeed the only possible interpretation of the ILF model, and we will argue that this shortcut is unnecessary, and that it should be avoided because it is counterfactual, and unrealistic in its implications. We then propose, supported by a long and growing list of central bank publications, a more realistic framework, the financing through money creation (FMC) model. This model describes banks as financial institutions that create digital purchasing power for their borrowers, with all physical resource trading taking place outside the banking sector. We show that a simple but fundamental change to an otherwise standard New Keynesian DSGE model can be used to formalize the logic of the FMC model. This change involves the budget constraints of banks' customers rather than the optimization problems of banks themselves, and as a result FMC models can directly incorporate the many advances that have recently been made in the modelling of banks. However, the differences in non-banks' budget constraints have a strong effect on the ability of banks to grow their balance sheets, and thereby lead to very sizeable changes in the dynamic properties of the model, including changes in the expected effects of policy interventions.

For the *baseline ILF model*, the key mechanism for bank balance sheet growth is that banks need to first collect deposits of physical commodities (loanable funds), from one group of nonbanks, savers, before they can make loans of physical commodities (intermediation) to another group of nonbanks, borrowers. In a variant of this model, which we will refer to as the *asset-trading ILF model*, this mechanism is augmented, but not replaced, by the intermediation of physical capital, either directly or in the form of securities that represent claims to physical capital, but this second type of intermediation will be shown to account for only a very small share of observed variability in aggregate bank balance sheet data. Both of these models therefore represent banks as resource-trading intermediaries that, wholly or primarily, store, borrow and lend physical commodities. Furthermore, because their deposits are created via the physical accumulation of commodities they are, to the extent that their role in economic exchange is acknowledged, commodity money. For the *FMC model*, the key mechanism for bank balance sheet growth is that banks fund new loans by creating new ledger-entry deposits for their borrowers, with deposits subsequently circulating as the economy's money. Loans and deposits are therefore purely financial or bookkeeping transactions that create gross balance sheet positions, as opposed to the physical transactions and net balance sheet positions of the ILF model. This means that loans and deposits are driven by non-banks' preferences over gross financial stocks rather than over net physical flows, and as a result such models can exhibit much greater balance sheet volatility in response to shocks.

Our preference for FMC models over the alternatives is driven by a desire to obtain a realistic representation of the lending process that is consistent with its description by major central banks, and to rule out, *a priori*, representations that are clearly at variance with the actual lending process. This is a consideration that should precede any empirical comparisons. But the stylized facts presented in Section 5 also provide support for our preference for FMC models.[1] The resulting approach sees

[1] Bayesian estimation of a larger New Keynesian FMC model and its alternatives, based on the seminal paper of

the creation and destruction of deposits by banks, and therefore the determination of the size of bank balance sheets, as the outcome of the simultaneous solution of the profit maximization problems of banks and their customers, with no independent role for either physical saving or (under inflation targeting regimes) central bank reserves. As such it is consistent with the credit mechanics approach that has recently been reviewed and advocated by Decker and Goodhart (2018).[2]

To demonstrate that the accumulation of deposits in baseline ILF models cannot represent financial (in the sense of non-physical) transactions, we study a representative financial transaction. Assume that A, who has an account at Bank A, has performed a service for B, who has an account at Bank B. In return A receives a check drawn on B's account, and deposits it in Bank A. The critical observation is that this check only has value *because the deposit already exists* - in Bank B. The transaction simply moves an existing deposit to a different account within the banking system, it does not create a new deposit for the banking system. Furthermore, it does not give Bank A additional funds to lend, because by double-entry bookkeeping the new deposit is automatically lent to Bank B at the moment it is received, in the form of an accounts receivable claim for the collection of funds in the form of central bank reserves. The same logic applies to any deposit of private financial instruments into bank accounts - they are not loanable funds, they are not new funds, no new deposits are created. Central bank money deposited in banks does not represent loanable funds either, because central bank reserves cannot be lent to non-banks, only to other banks, while physical cash is never disbursed against new bank loans, only against preexisting electronic deposits.[3] Furthermore, cash is a minor and non-constitutive element of modern financial systems. The accumulation of deposits in baseline ILF models must therefore represent the only remaining alternative, the accumulation of physical resources. This can also be seen directly in the (stylized) budget constraints of banks' depositors in baseline ILF models, where changes in deposits Δd_t equal the difference between physical incomes and physical expenditures of savers (superscript s) $inc_t^s - exp_t^s$, a difference that represents the physical accumulation, and then intermediation by way of bank loans, of physical resources. A similar budget constraint applies to borrowers (superscript b), who experience changes in loans $\Delta \ell_t$ and who can in addition accumulate physical capital k_t at market price q_t. We have

$$\begin{aligned} \Delta d_t &= inc_t^s - exp_t^s \ , \\ -\Delta \ell_t + \Delta \left(q_t k_t \right) &= inc_t^b - exp_t^b \ , \end{aligned} \tag{1}$$

where the change in banks' balance sheet is given by $\Delta \ell_t = \Delta d_t + \Delta e_t$, and where e_t is bank equity. The right-hand side of the first equation shows that savers can *only* increase the stock of bank deposits through a decision to accumulate flows of additional physical commodities. The volatility of such savings is limited by curvature in savers' preferences over net physical flows. The resulting smoothing implies that changes in aggregate bank balance sheets cannot be very volatile, while in the data such changes are extremely volatile. As we will show in Section 5.1, saving is indeed far smoother than changes in aggregate bank balance sheets, and often has a very different magnitude. We will at times and for brevity refer to the saving channel as SAV.

Asset-trading ILF models do not solve this problem. Asset-trading ILF models are variants of ILF models that allow the size of bank balance sheets to *also* change with changes in either banks'

Christiano et al. (2014), is the subject of ongoing work. Deleidi and Levrero (2019) study the US money creation process using VAR and VECM methods and find strong evidence for the FMC mechanism.

[2] This approach was popular in Germany in the 1920s-1960s, and several of its leading proponents became high-ranking Bundesbank officials. Key references include Hahn (1920), Gestrich (1936), Lautenbach (1952), Rittershausen (1956) and Stützel (1958).

[3] Overdrafts are not an exception. An overdraft permits the creation of new bank deposits on demand. A newly created deposit can then be instantaneously withdrawn in the form of a check or cash.

holdings of or the value of physical capital, or of securities that represent claims to physical capital. The first of these channels emphasizes the net acquisition, by banks from nonbanks, of physical capital (or the corresponding securities), which has its counterpart in the net creation of new deposits for these nonbanks. We will refer to this channel as direct financing substitution or for brevity as DFS. The second channel emphasizes valuation effects on banks' existing asset portfolios, such as changes in the price of capital, which has its counterpart in changes in bank equity. We will for brevity refer to this channel as VAL. In this model class it is often assumed that banks (superscript bk) do not make loans but instead directly own capital, in which case we have the stylized budget constraint of savers

$$\Delta d_t + \Delta \left(q_t k_t^s\right) = inc_t^s - exp_t^s \ , \tag{2}$$

while the change in banks' balance sheet is given by $\Delta \left(q_t k_t^{bk}\right) = \Delta d_t + \Delta e_t$, with $k_t = k_t^s + k_t^{bk}$. This shows that the DFS and VAL channels do allow for more rapid changes in the size of bank balance sheets, and this does correspond to actually observed financial mechanisms. However, it also shows that the default mechanism for deposit creation remains the accumulation of physical commodities, with this mechanism only augmented but not replaced by DFS and/or VAL effects. Furthermore, as we will show in Section 5.1, DFS and VAL effects only account for extremely small shares of empirically observed variations in aggregate bank balance sheets.[4] To explain the data using these models, the accumulation of physical resources must therefore remain the main mechanism for deposit creation. That however leaves us far short of a complete explanation for how real-world banking systems grow or shrink their balance sheets.

That explanation can however readily be found in a long and growing list of central bank publications. These explain that the key function of the banking system is the creation of new monetary purchasing power through loans, for a single agent that is both borrower and depositor. The clearest explanations of this fact, which make exactly the same arguments as in this paper, come from two reports by the Bank of England (McLeay et al. (2014a,b)), from numerous papers by the Bank for International Settlements (see e.g. Borio and Disyatat (2011, 2015)), and from reports by the Bundesbank (Bundesbank (2017)) and the Reserve Bank of Australia (Doherty et al. (2018)). A Supplementary Literature Review, available from the authors upon request, provides a comprehensive list of other citations on this topic, many of them from the world's leading central banks. FMC models are now also starting to be incorporated into the DSGE models of central banks and policy institutions, including the International Monetary Fund (Benes et al. (2014a,b)), Central Bank of Ireland (Lozej et al. (2017), Lozej and Rannenberg (2017)), Lithuanian Central Bank (Ramanauskas and Karmelavičius (2018)), Norges Bank (Kravik and Paulsen (2017)) and People's Bank of China (Sun and He (2018)).[5]

To demonstrate that the function of banks is indeed monetary financing as described in the FMC model, we study the procedure that a bank uses to make a new loan to a customer X.[6] The bank simultaneously creates a new loan entry, in the name of X, on the asset side of its balance sheet,

[4] An additional role for physical capital is as loan collateral, which represents contingent rather than direct claims to capital by banks. However, the connection between the size of bank balance sheets and the quantity of collateral is also very weak. First, banks frequently change the quantity of financing at an unchanged quantity of underlying collateral. Second, a large share of bank lending, including business loans, is uncollateralized. For example, according to the Fed's flow of funds, over the last two decades only around 40% of all US commercial and industrial loans have been collateralized.

[5] There is also an important literature from leading legal scholars that explains the FMC nature of banking. See e.g. Desan (2014), Ricks (2018) and Hockett and Omarova (2017). The latter make the same points made in this paper, but of course from a legal rather than an economic perspective.

[6] Werner (2014a) documents this procedure for an actual individual loan disbursement at a small bank.

which represents its right to receive future installments and interest on the loan, and a new and equal-sized deposit entry, also in the name of X, on the liability side of its balance sheet, which represents its obligation to deliver current funds. The key observation is that in the case of banks this newly-created "accounts payable" liability (IOU) to deliver current funds can immediately be used *as* current funds, as money. Only banks have this ability, because only banks are perceived to be able to credibly commit to honouring their IOUs universally (that is, vis-à-vis any subsequent holder of the IOU), thereby making these IOUs acceptable as a universal medium of exchange, or money. We will discuss the reasons below. A representation of the FMC model in terms of a stylized budget constraint for a representative non-bank agent (superscript r) is

$$\Delta d_t - \Delta \ell_t + \Delta (q_t k_t) = inc_t^r - exp_t^r \ , \tag{3}$$

where the change in banks' balance sheet is again given by $\Delta \ell_t = \Delta d_t + \Delta e_t$.[7] In this case savers can *only*[8] increase the stock of bank deposits through a decision to obtain an increase in the stock of loans. Such a loan request can be accommodated by a bank, provided that it considers the proposition profitable, through purely financial means, namely through offsetting loan and deposit ledger-entries. Except for general equilibrium feedback effects, this has no effects on physical savings $inc_t^r - exp_t^r$, which therefore equal the increase in the value of physical capital. The latter is of course the national accounts identity (in a closed economy and at the world level) that saving equals investment. Deposit accumulation is therefore not constrained by non-banks' preferences over net physical flows, which call for intertemporal smoothing, but rather by their preferences over gross financial stocks, which can call for rapid jumps. Deposits can therefore change instantaneously while savings can only accumulate gradually over time, and as a result the volatility of the stock of bank deposits can be much higher than in the ILF model. This is key to understanding the very different behavior of national (physical) saving and aggregate bank balance sheets, which we will study in Section 5.1.

Because banks' liability to deliver current funds *is* current funds, banks create their own funding in the act of lending. And because both entries are in the name of the same customer, a new loan involves no intermediation. No real resources need to be diverted from other uses, by other agents, in order to be able to lend to the new customer.[9] Instead, new purchasing power is injected into the economy independently of the economy's pre-loan income flow. In turn this new purchasing power, by mobilizing resources that would otherwise have remained idle, triggers increases in the economy's post-loan income flow. In other words, ex–ante spending plans are not constrained by prior income, but by prior income and net new credit, while ex-post realized spending must of course equal realized income. What is needed in this from third parties is not a deposit of physical resources but the acceptance of the new purchasing power in payment for physical resources. This is never in question as long as bank deposits remain the universally accepted medium of exchange. This distinction between income and purchasing power can be traced back to many leading economists of the past, including Schumpeter (1934), Keynes (1939) and Kaldor (1989). The tradition has continued in Post-Keynesian economists' emphasis on the ability of commercial banks to create money, and the importance of this mechanism for monetary and financial stability, see e.g. Minsky (1977), Moore (1979), Lavoie (2014) and Keen (2014, 2015). FMC models capture this distinction.

[7] Because in all banking models bank equity is ultimately owned by households, Δe_t could have been added to the left-hand side of (3). This would emphasize that all bank balance sheet items, including equity, are ledger entries and unrelated to physical saving.

[8] In this case the word "only" follows directly from the bank balance sheet identity.

[9] The only "resource" that banks require to make a loan is a keyboard or, in earlier times, a pen. This was pointed out long ago by Milton Friedman (1971, p. 2): "The correct answer for [the question of the origin of] both Euro-dollars and liabilities of US banks is that their major source is a bookkeeper's pen."

By contrast, in ILF models the debt-financed increase in spending power of the borrower, who gains access to additional physical resources, is offset by the diminished spending power of the depositor, who gives up access to part of his physical resources. The much greater impact of bank financing on aggregate purchasing power in FMC models is the main reason why they exhibit significantly larger real effects following financial shocks.

To avoid misunderstandings, while FMC models reflect the fact that banks face no *technical* limits to expanding their balance sheets instantaneously and by large amounts, they also reflect the fact that banks face *economic* limits. However, the latter do not include physical limits such as the availability of sufficient savings. Rather, the most important limits are their own and their customers' potentially volatile expectations regarding future profitability and solvency. Monetary and regulatory policies can influence these expectations, but they do not fully determine them.

In the FMC model of this paper, a single representative household holds bank deposits because of a transactions cost technology. As in all models that use such technologies (or the alternatives of cash-in-advance constraints or money in the utility function), the actual inter-agent transactions that require the use of bank deposits are not explicitly modeled. The only reason for this assumption is model economy. As shown in Kumhof and Wang (2018), it is possible, albeit more complex, to build FMC models with explicit inter-sectoral transactions. In that paper, deposits need to be used to make all of the economy's payments inside each period, in a sequence of deposits-in-advance constraints.[10] First, loans create deposits for firms. Second, firms make payments for labor and other services to workers so that production of commodities can take place. Third, workers make payments to firms so that sales of commodities can take place, thereby returning the deposits to firms. Fourth, firms use the deposits to extinguish their loans. The process repeats in the following period. Because of the intra-period nature of this sequence, it remains compatible with the model used in this paper, which merges workers and firms into a single representative household in the spirit of Lucas (1990) and Schmitt-Grohé and Uribe (2004), and where the multiple deposits-in-advance constraints, whereby deposits are a prerequisite for real activity, are replaced by a single transactions cost technology, whereby deposits make real activity cheaper.

ILF and FMC models attach very different importance to macroeconomic flows and stocks. In contrast to ILF models, in FMC models physical *flows* of (national accounts) saving do not play a significant role in the determination of lending/financing, because new loans are funded by the *ex nihilo* creation of new ledger-entry deposits. The fact that a newly created deposit may subsequently be withdrawn from a particular bank and deposited at another bank does not change this fact, it merely changes the identity of the deposit holder and may, if the deposits and withdrawals of a particular institution do not net out on any given day, be settled through interbank lending or borrowing of central bank reserves, which does not affect the aggregate quantity of deposits vis-à-vis non-banks. In this context, it is sometimes argued that even though banks create deposits according to the FMC model, ex-post the withdrawn deposit will be kept by the recipient and is therefore saving, so that the ILF model is still an adequate description of reality. There are several problems with this argument. First, even for the special case of a new loan for physical investment purposes the ILF model, by not accurately capturing the underlying mechanism, reverses cause and effect. For this case a new loan and the corresponding deposit creation lead to additional investment that could otherwise not have occurred, because the investor did not have access to the necessary purchasing power. They must therefore lead to additional saving by definition, specifically as a result of the national accounts identity between saving and investment. The direction of causation is therefore from financing to investment to saving. In other words, saving does not

[10] The following exposition is simplified but captures the essence of the argument in Kumhof and Wang (2018).

finance investment, financing does.[11] The reversal of causation matters because, as explained above, when saving precedes lending the quantity of lending is far more constrained than when lending precedes saving. The reason is that in the former case deposit creation is constrained by non-banks' smoothing preferences over net physical flows, while in the latter case it is constrained by their portfolio preferences over gross financial stocks. Second, while it is possible that recipients intend to become savers that hold on to their deposits, it is far more likely, given the observed volume of payment flows relative to changes in the volume of saving, that they merely intend to be holders of purchasing power, who both continue to circulate their deposits through further payments[12] and at the same time continue to receive deposits from other depositors. Any theory of banking that is built on the notion of depositors as savers is therefore certain to miss a major reason for changes in bank balance sheets, namely the creation of purchasing power for purposes that are unrelated to saving. In fact, as we will see in Section 5.1, in the data this is the dominant reason by far. In this context it is important to observe that in most economies physical investment loans are a quantitatively small component of overall lending volumes.[13]

In contrast to ILF models, in FMC models financial *stocks*, specifically the ledger entries that create financial balance sheets, play a critical role in the determination of lending/financing. Banks in the ILF model record nonzero net non-financial (physical) transactions, which means that *loans are conditional on deposits* that are created through the saving of physical resources, and where loans represent an exchange of two physical bank assets, existing physical resources against a repayment claim for future delivery of physical resources. This notion is one of barter of different resources against each other, and as such it originates in thinking of banks analogously to non-financial firms. Banks in the FMC model record nonzero gross, but zero net, financial (monetary) transactions, which do not require prior saving of physical resources, and where loans represent the creation of new financial assets that are funded by the simultaneous creation of new financial liabilities, new deposits against a repayment claim for future delivery of deposits. In other words, *deposits are conditional on loans* that are created through ledger entries. The bank's customer experiences no impact change in his net balance sheet position vis-à-vis the bank, but such transactions are nevertheless essential for the functioning of the real economy, because the new bank liability represents an addition to the economy's stock of money. But while this money is essential in facilitating purchases and sales of physical resources outside the banking system, it is not itself a physical resource, and can be created at negligible resource cost.

We have argued that in FMC models deposits are conditional on loans. An alternative terminology is that loans come before deposits, but when we employ that terminology (or its opposite in the ILF model) this is understood in a causal rather than a temporal sense. The distinction is important because in the specific models used here the creation of loans and deposits happens inside the same time period. In the ILF model, loans cannot be made *unless* a non-bank saves physical resources. In the FMC model, deposits cannot be created *unless* a bank makes a loan (or acquires an existing asset). However, in both models a temporal dimension can arise due to an assumed intra-period sequence of actions. Kumhof and Wang (2018) is an FMC example.

The foregoing discussions illustrate that accounting is an indispensable tool for understanding the economics of money and banking. The activity of banks consists almost entirely of the profit-

[11] See also Lindner (2012, 2013).

[12] In other words, there is no reason why the velocity of circulation of new deposits should be equal to one.

[13] In most major economies well under half of variations in total credit to the nonfinancial private sector are accounted for by loans to nonfinancial business, of which in turn the major portion relates to the financing of activities that are not related to saving, such as working capital loans and loans related to the acquisition of existing assets. See Bezemer et al. (2017) for details.

maximizing and risk-minimizing creation and destruction of digital ledger entries. To understand how these ledger entries are created, one has to understand accounting, because accounting is the technology of banking.

The ability to create their own funding in the act of lending is what distinguishes banks from shadow banks, or non-bank financial institutions (NBFIs). In the words of Zoltan Pozsar, "central banks create the ultimate means of settlement (cash and reserves), banks create the means of settlement (deposits) for players at lower levels of the hierarchy, dealers (the ultimate shadow banks) make markets with the deposit money they have (repo, swaps, securities)."[14] This means that NBFIs, in order to lend, need to first borrow to obtain bank deposits. This activity does intermediate loanable funds, but in the form of bank deposits, not of physical resources. Therefore, such intermediaries of loanable funds can only exist in symbiosis with banks, which are not intermediaries of loanable funds. As discussed by Werner (2014b), in the UK context, when monies are deposited with (rather than lent to) NBFIs, the latter are subject to Client Money Rules, which require them to hold these monies in trust or off-balance sheet, and therefore not as their liabilities. Banks on the other hand are not subject to Client Money Rules and keep deposits on their balance sheet, as their liabilities. Depositors are therefore not the legal owners of their bank deposit, with the bank holding it in trust for them. Rather, they are one of the creditors of the bank.

It is important to emphasize that when we refer to banks, bank loans and bank deposits, both conceptually and in our formal model, we are thinking of the institutions, assets and liabilities of the entire financial system, which includes both banks and NBFIs. Banks are the central actors in the overall financial system, because their ability to create new money implies that the system as a whole has the ability to create new money, even if a subset of institutions, NBFIs, only has the ability to use but not to create such money. From this perspective it is therefore legitimate to model the overall financial system as consisting entirely of banks, and this explains why we will frequently refer to the financial sector rather than the banking sector. We note that we abstract from movements of assets and liabilities between different parts of the financial sector, because those are not of first-order importance for the questions studied in this paper.

As discussed above, the FMC nature of banks has always been very well understood by central banks. It was also, in the wake of the Great Depression and until the 1950s, almost universally understood among leading macroeconomists - Schumpeter (1954) contains an excellent summary of the intellectual history as seen from the 1950s. Unfortunately this understanding was gradually lost in the wake of Gurley and Shaw (1955, 1956) and Tobin (1963). Gurley and Shaw (1955, 1956) replace the critical distinction between banks and NBFIs with the much less important distinction between intermediated and direct debt. They treat banks as another form of intermediary, and bank liabilities as another form of debt. This work was strongly criticized by monetary theorists of that time (Culbertson (1958), Smith (1959)).[15] Tobin (1963) accepts that banks can initially create new deposit money through loans, but argues that this money creation will eventually be either neutralized or reversed. He offers two main arguments.

First, he argues that depositors will remove their newly created deposits from banks, whose liabilities he equates with money, to NBFIs, whose liabilities he equates with non-money. This argument does not contradict the mechanism whereby the financial sector has the ability to fund new credit by creating new money, because even if new deposits are immediately withdrawn from banks and redeposited with NBFIs, the size of the aggregate financial sector balance sheet has increased by

[14] From a private communication with Zoltan Pozsar. For more details, see Pozsar (2014).
[15] See the Supplementary Literature Review for details of the critiques.

the exact amount of the newly created deposit.[16] More importantly, the argument is misleading concerning banks' ability to create new money, because the liquidity or "moneyness" of different assets is not a binary characteristic but a continuum. An asset's location on this continuum can be identified by reference to the so-called convenience yield, which equals the difference between the risk-free rate and the financial return on the monetary asset. This has been emphasized in the long literature on divisia indices of money (Barnett et al. (1984), Anderson et al. (2018)), and is part of our model. The creation of new bank deposits might therefore move the aggregate economy along this continuum, but both old and new deposits serve as money.

Second, Tobin (1963) argues that newly created deposits will be used by their recipients to repay loans, thereby destroying the deposits, and that it would therefore be incorrect to make the claim that banks possess the same "widow's cruse" as the central bank, whose printing press can expand the money supply exogenously and at will. However, the FMC model does not make that claim. Instead it argues that equilibrium money creation is endogenous and finite, being determined by the interaction of the convex optimization problems of banks and their customers. Applied to such an environment, Tobin's implicit assumption appears to be that money-creation shocks will be accompanied by exactly offsetting money-destruction shocks, an assumption that economists do not make for any other shocks. For example, if some agents receive additional loans and deposits to make payments to other agents, this does not imply that those other agents must want to use those deposits to repay loans, just like a shock that increases the consumption demand of some agents does not imply that other agents must want to reduce their consumption in an offsetting fashion. In fact, in both cases, it implies the opposite. In the case of the money-creation shock, the reason is that the additional money creation, by making additional transactions feasible or at least cheaper, mobilizes additional resources and thus increases incomes. Additional incomes stimulate additional spending, and additional spending requires higher deposit balances rather than loan repayments. This mechanism is a key ingredient of our FMC model.

Recent models of monetary economies have increasingly omitted monetary aggregates altogether, especially since Woodford (2003), instead adopting the so-called cashless limit assumption whereby the transmission of monetary policy can be thought of exclusively in terms of interest rates. By contrast, the preceding model generation, which is sometimes referred to as Sidrauski-Brock (SB) models[17], and which is still commonly used today, had a central role for money. FMC models are best classified as a further development of SB models. In this context two typical features of the original SB models are important. First, private agents hold money because of a money demand function that is generated by either a cash-in-advance constraint, money in the utility function or a transactions cost technology. Second, the only money in these models is exogenous government fiat money, while in the wake of Gurley and Shaw (1955, 1956) banks are either omitted altogether or modeled as resource traders. A large and seminal literature has since arisen to improve upon the first feature (see e.g. Gu et al. (2013), Lagos et al. (2017) and the many references cited therein). This literature often retains the assumption of exogenous government fiat money, in combination with private real credit and private commodity money, but typically not with FMC-type deposit money. We mention this literature only briefly, because it generally has a different focus from the financial cycle and business cycle issues that concern us. Our work can instead be seen as an attempt to improve upon the second feature of SB models. Our motivation is that in modern economies government fiat money accounts for only a very small fraction (3% in the case of the UK) of the broad money supply, with the liabilities of private financial institutions accounting for

[16]Note also that if the recipient NBFIs are subject to Client Money Rules, this transaction does not represent a withdrawal of deposits from banks in the first place.

[17]See Sidrauski (1967) and Brock (1975).

the remainder (97%).[18] The key feature of FMC models relative to SB models is to have these 97% rather than the 3% enter the money demand function, and to present the optimizing calculus according to which banks and their customers decide on the creation of that money. Because in our model bank deposits are effectively riskless in nominal terms, they serve the same functions as government fiat money in older SB models. The difference is therefore not in the function or risk properties of money, but only in the manner of its creation.

This approach makes it very straightforward to develop FMC models from their ILF counterparts, and given that the latter almost completely dominate the models currently used in policymaking institutions, this is of considerable practical interest. In the simplest cases, as in this paper, arriving at the FMC model only requires the dropping of one equation and one variable from the ILF model. However, money demand and the convenience yield play a fundamentally different role in the two models. In FMC models a single representative household interacts with banks, whose service to the household consists of the creation of deposits through loans. For the household this service has a cost, the interest rate spread between loans and deposits. For a positive demand for deposits to exist, it is therefore essential that this cost is offset by the benefit of a monetary convenience yield. On the other hand, in ILF models the service of banks is the intermediation of physical resources. In such models a convenience yield on deposits is not essential. It is nevertheless added to our ILF model to maintain the symmetry of structural equations and steady states across models.

Universal acceptability as a medium of exchange requires that bank deposits must be perceived as nominally (nearly) risk-free. Part of the literature has argued that banks are able to issue risk-free money because of their ability to diversify and thereby construct optimal portfolios of liabilities that are perceived to be safe, or at least much safer than the alternatives. This private mechanism indeed plays an important role, but we would argue that public support for banking systems is even more important, especially during crisis times, when the quality of banks' assets often only determines the fiscal cost of public support operations. In fact, the entire system of central banks and regulatory agencies was constructed with this as its original objective, because trust in the nation's main medium of exchange has always been considered critical for the performance of the overall economy (Goodhart (1988)). In our model bank deposits are therefore safe because of effective official regulations.

This emphasis on the important role of public institutions is not only motivated by current institutional arrangements but also by monetary history. The explanation of monetary phenomena as having their origin in optimal private arrangements has a history that goes back to Menger (1892). In this story trade was initially based on barter, which was later followed, in order to overcome the double coincidence of wants problem in private exchange, by coinage and then credit. We now know that this "creation myth" of money was in fact not based on actual historical research, and we have very extensive historical and anthropological evidence[19] which shows that the sequence among the major known civilizations was in fact the opposite, with the emergence of sophisticated temple-based (meaning government-based) credit systems among the Mesopotamian civilizations coming first, and being constructed with a view not to private interests but to the interests of state administration and of economic interests closely connected with the state. Such systems predated the emergence of coinage by thousands of years, and barter only ever occurred upon the breakdown

[18] Furthermore, under inflation targeting and away from the zero lower bound (ZLB) on nominal interest rates, the quantity of government fiat money is itself endogenous rather than exogenous. The reasons are discussed below.

[19] See Graeber (2011), Hudson (2017) and the very large literature cited therein. Anthropologist Caroline Humphrey (1985) observes: "No example of a barter economy, pure and simple, has ever been described, let alone the emergence from it of money ... all available ethnography suggests that there has never been such a thing."

of such civilizations, in other words during and after periods of war. Money was therefore from the beginning a creature of the law and of public administration, and this has been recognized throughout a long intellectual history (see Zarlenga (2002)). It is evident to us that to a very significant extent this is still true today.

A key contribution of our paper is a discussion of the differences in simulation properties between New Keynesian ILF and FMC models. In both models bank lending is subject to the costly state verification friction of Bernanke et al. (1999), and bank deposits are demanded because of the transactions cost technology friction of Schmitt-Grohé and Uribe (2004). The two models have identical functional forms, parameter calibrations, first-order conditions and deterministic steady states, including completely identical optimization problems for banks, and identical interpretations of standard analytical balance sheet ratios such as leverage ratios. The only difference is in the size and speed of transitional adjustments of bank balance sheets to shocks. This in turn is solely due to the fact that the ILF model features two separate budget constraints for banks' customers, one for a representative depositor household and another for a representative borrower entrepreneur (see equation (1)), while the FMC model features a single budget constraint for a representative borrower-cum-depositor household who takes over all of the functions of the ILF model's entrepreneur (see equation (3)). Relative to the ILF model, the FMC model therefore has fewer equations (the entrepreneur budget constraint is removed) and fewer variables (entrepreneur net worth is removed). This single difference accounts for all of the differences in the models' properties.

The simulation properties of ILF and FMC models exhibit very sizeable differences. Compared to ILF models, and following identical shocks, FMC models predict changes in aggregate bank balance sheets that are far larger and far faster, and that have greater effects on the real economy, while the adjustment process involves smaller changes in lending spreads. Compared to ILF models, the empirical predictions of FMC models are that changes in the size of financial sector balance sheets relative to real variables are large and rapid, that the leverage of aggregate banking systems is procyclical, and that financial crashes have a large credit rationing component. As shown in Section 5, these predictions are qualitatively in line with the data.

Our analysis has important implications for policies that are designed to address financial frictions. In ILF models financial frictions are physical bottlenecks in the flow of physical resources to their most productive use, while in FMC models they are liquidity bottlenecks that have tax-like distortionary effects on the allocation of physical resources. The two models therefore call for potentially very different policy interventions, for example in the case of policies designed to prevent excessive lending booms.[20] This is an important area for future research.

We have so far discussed the FMC model entirely in relation to the ILF model. However, discussions of banking still frequently appeal to a third model class, the deposit multiplier model. It argues that the size of bank balance sheets is a multiple of the policy-determined quantity of central bank money, with this quantity being either a constraint on credit creation or on repeated re-lending by banks. While this is, unlike the ILF model, a monetary model, it has several shortcomings as a description of the money creation process. First, it ignores that central bank money either cannot be (reserves) or is not (cash) lent to nonbanks. Second, it ignores that, away from the ZLB, modern central banks target interest rates, and are committed to supplying as much central bank money as banks demand at that rate.[21] The quantity of reserves is therefore at most times a consequence, not

[20] See Mian and Sufi (2014), who argue that such policies might have prevented the GFC of 2008.

[21] As shown by Kydland and Prescott (1990), the availability of central bank reserves also did not constrain bank lending during the period when the Federal Reserve officially targeted monetary aggregates. The same is true, for

a cause, of private money creation. This argument has been made repeatedly in publications of the world's leading central banks, and also, supported by recent experience, in Decker and Goodhart (2018). The Supplementary Literature Review provides a comprehensive list of citations.

The rest of the paper is organized as follows. Section 2 reviews the modern theoretical literature on banks and relates it to the models studied in this paper. Section 3 develops the ILF and FMC models. Section 4 studies the differences between these models by way of impulse responses. Section 5 presents the stylized facts as they relate to the predictions of these models. Section 6 concludes.

2. Literature Review

In this section we review the recent theoretical macro literature on banking.[22] For convenience we divide our review into two parts, New Keynesian DSGE models and other macroeconomic theory models. Because this literature is large, the list is necessarily incomplete.

2.1. New Keynesian DSGE Models

In relation to our paper, recent DSGE models with financial frictions can be divided into three groups. In the first group, all lending is direct. Because banks are absent, this literature is distant from the topic of this paper. It includes Iacoviello (2005), which uses the borrowing constraint first introduced by Kiyotaki and Moore (1997), and Jermann and Quadrini (2012). In the second and third groups, banks are present and modelled as ILF banks. In the second group, banks' net worth and balance sheets play no role in the analysis, typically because all lending risk is diversifiable and the emphasis is on loan pricing. This group includes Christiano et al. (2014), which uses the costly state verification mechanism first introduced into macro models by Bernanke et al. (1999), as well as Cúrdia and Woodford (2010), de Fiore et al. (2011) and Boissay et al. (2016). In the third group, banks' balance sheets and net worth do play a role, either through an incentive constraint under moral hazard or through a regulatory constraint. This group includes Gerali et al. (2010), Gertler and Karadi (2011), Gertler and Kiyotaki (2011), Adrian and Boyarchenko (2012), Clerc et al. (2015), Nelson et al. (2015), Justiniano et al. (2015), Benes and Kumhof (2015), Eggertsson et al. (2017) and Nuño and Thomas (2017). This paper's ILF model belongs to the third group, with net worth subject to a regulatory constraint.

2.2. Other Macroeconomic Theory Models

We classify models in this group by the three main groups of mechanisms for aggregate bank balance sheet growth that have been identified in the literature. The first mechanism is the creation or destruction of deposits through loans (FMC), the second is the acquisition, sale or revaluation of existing physical assets or of securities that represent claims to such assets (DFS or VAL, here combined and denoted by SEC), and the third is the saving of physical resources (SAV). Other than Benes and Kumhof (2012), Jakab and Kumhof (2015)) and the related work at central banks cited in Section 1, to our knowledge only three papers, Goodfriend and McCallum (2007), Faure and Gersbach (2017) and Clancy and Merola (2017), are based exclusively on the FMC mechanism. In Goodfriend and McCallum (2007) loans create deposits that enter a cash-in-advance constraint

different reasons, under quantitative easing at the ZLB.

[22] An older literature on the credit channel view of monetary policy is summarised in Kashyap and Stein (1993) and Kashyap et al. (1993). This paper will not discuss partial equilibrium models of banking.

on consumption. The paper does not emphasize the major differences between this approach and the ILF model, which would however be muted in their model because of the assumption of a real loan production function that depends on collateral values and labor effort. Faure and Gersbach (2017), in a 2-period model, show that in the absence of uncertainty SAV and FMC models imply identical allocations. This is related to the result in this paper that the deterministic steady states of the two model classes are identical. Finally, Clancy and Merola (2017) study macroprudential policies in small open economies in the variant of the FMC model of Benes et al. (2014a,b). The 3-period model of Donaldson et al. (2018) is based on a combination of the FMC and SAV mechanisms for balance sheet growth. As in FMC models, banks issue book money ("fake receipts") through risky lending. But as in SAV models, banks also need to function as warehouses that issue commodity money ("commodities receipts"), because warehoused commodities are required as collateral. The paper has at its heart an excellent explanation of the ability of banks to create money through lending. However, physical warehousing is almost completely absent in modern financial systems.[23] The model could be turned into an exclusively FMC model by assuming either risk-free uncollateralized loans, commodities collateral without physical warehousing, or other and more common forms of collateral such as real estate. Piazzesi and Schneider (2018) is based on a combination of all three mechanisms for bank balance sheet growth. Bank deposits are created through non-banks' physical saving and the purchase of Lucas trees by banks from non-banks. But households can also obtain liquidity through intraday loans, which are close relatives of the fake receipts of Donaldson et al. (2018). Chari and Phelan (2014) is based on a combination of the SEC and SAV mechanisms, with bank deposits that are used as a medium of exchange for consumption purposes, but created through a combination of physical asset purchases from and the accumulation of physical savings by households. Because, as discussed in Section 5.1, the SEC mechanism plays only a very small role in the data, the practically most relevant ingredient of this model class remains the SAV mechanism. The bank run model of Gertler and Kiyotaki (2015) is also based on a combination of the SEC and SAV mechanisms, with fire sales of securities and valuation losses playing a major role in the evolution of bank balance sheets during runs. But in other respects this model does not fit neatly into our classification. The reason is that these authors divide their economy into shadow banks, who experience runs, and a combination of commercial banks and nonbanks, who run, while the aggregate financing of nonbank borrowers, the main object of interest of our paper, remains constant throughout. This choice is appropriate for the study of financial crises, where securities transactions between different parts of the financial sector have played a critical role even though, as discussed in Section 5.1, securities transactions between the financial and nonfinancial sectors and valuation losses have only played a minor role for the aggregate financial sector. Gertler et al. (2015) builds on Gertler and Kiyotaki (2015) but divides the economy into three groups, shadow banks, commercial banks and households. In this model changes in the size of the aggregate financial sector's balance sheet are due to changes in direct holdings of securities (capital) by households. Bigio and Weill (2016) is based exclusively on a variant of the SEC mechanism. In this three-period two-state model banks buy high-risk illiquid assets from producers in exchange for issuing low-risk liquid deposits, thereby allowing producers to hire additional workers who will not accept to be paid in high-risk assets. At the final stage however bankers settle their deposit contracts using the physical returns on their assets, so that banks are still intermediaries of physical resources, and workers still buy the output of producers using physical resources. This could be changed by allowing for an FMC mechanism, with workers spending their deposits to buy output from producers, and producers using these deposits to repurchase the illiquid assets. Another difference is that Bigio and Weill (2016) emphasize risk on the asset side of banks, while our paper exclusively emphasizes the liability side, with loans assumed to be

[23] Custodial services do not qualify because custodial items are held in trust and not on bank balance sheets.

risk-free. This permits a significantly more tractable (and infinite horizon) model. Brunnermeier and Sannikov (2016) is based on the SEC mechanism, with bank liabilities that serve as a low-risk store of value, rather than as a medium of exchange as in the papers cited so far. Bianchi and Bigio (2014) is based exclusively on the SAV mechanism, with banks intermediating real savings between households and firms. While in this model loans and deposits can experience large jumps, the reason is the assumption of an infinitely elastic supply of physical savings. This simplifying assumption facilitates a clearer focus on the main topic of the paper, the interaction between monetary policy and the interbank market.

3. The Models

3.1. Overview

In this section we develop a pair of simple canonical New Keynesian ILF and FMC models. The models are identical in all functional forms, parameter calibrations, optimality conditions and steady states. The model economy is closed, and experiences constant positive technology growth $x = T_t/T_{t-1}$, where T_t is the level of labour augmenting technology.[24] When the model's nominal variables, say Z_t, are expressed in real normalized terms, we divide by the GDP deflator P_t and by T_t. We use the notation $\check{z}_t = Z_t/\left(T_t P_t\right) = z_t/T_t$, with the steady state of $\check{z}_t$ denoted by $\bar{z}$.

3.2. Banking Sector

The banking sector consists of three subsectors, a retail deposit bank that determines the terms of deposit contracts, a retail lending bank that determines the terms of loan contracts, and a wholesale bank that ensures compliance with macroprudential regulations. Nominal and real policy interest rates, which equal the rates on one-period government debt, are denoted by i_t and r_t, where $r_t = i_{t-1}/\pi_t^p$ and $\pi_t^p = P_t/P_{t-1}$. Wholesale lending rates are $i_{\ell,t}$ and $r_{\ell,t}$, retail lending rates, which add a credit risk spread to wholesale rates, are $i_{r,t}$ and $r_{r,t}$, and deposit rates are $i_{d,t}$ and $r_{d,t}$.

3.2.1. Retail Deposit Banks

Retail deposit banks have unit mass and are indexed by j, where individual banks differ by the deposit variety they offer. Retail deposit banks create deposit money $d_t(j)$ to purchase wholesale deposits $o_t(j)$ and government bonds $b_t(j)$. Because wholesale deposits are perfect substitutes for government bonds in the creation of deposit money, their interest rate is arbitraged with the policy rate i_t. Retail depositors require a CES composite d_t of different deposit varieties with elasticity of substitution σ, so that retail deposit banks act as monopolistic competitors vis-a-vis depositors. Letting $s = \sigma/(\sigma+1)$, we have the optimal interest rate condition

$$i_{d,t} = s \cdot i_t \ . \tag{4}$$

Real normalized profits of retail deposit banks are $\check{\Pi}_t^R = \left(\frac{r_t}{x} - \frac{r_{d,t}}{x}\right)\check{d}_{t-1}$, where we have dropped the index j because in equilibrium all retail deposit banks behave identically. Furthermore, in equilibrium government debt equals zero at all times, so that wholesale loans are equal to retail deposits. We will therefore from now on, to simplify the exposition, set $\check{o}_t = \check{d}_t$.

[24] Steady state growth is needed for the model to generate realistic dynamics of financial stocks and flows when steady state real interest rates are calibrated at their historical average values.

3.2.2. Wholesale Banks

Wholesale banks have unit mass and are indexed by j, where individual banks differ by the size of their net worth. Wholesale banks' nominal and real normalized stock of loans between periods t and $t+1$ is given by $L_t(j)$ and $\check{\ell}_t(j)$, while their stock of deposits is $D_t(j)$ and $\check{d}_t(j)$, and net worth is $N_t^b(j)$ and $\check{n}_t^b(j)$. Their balance sheet, in real normalized terms, is therefore given by

$$\check{\ell}_t(j) = \check{d}_t(j) + \check{n}_t^b(j) \ . \tag{5}$$

The model economy does not feature central bank reserves or cash, only bank deposits. The reason is that central bank reserves do not quantitatively constrain the ability of banks to extend loans (see the discussion in Section 1) and that cash is dominated in return by bank deposits.

Banks are assumed to face pecuniary costs of falling short of official minimum capital adequacy ratios (MCAR). This regulatory framework introduces a discontinuity whereby in any given period a bank either remains sufficiently well capitalized or it falls short of capital requirements and must pay a penalty. The cost of such an event, weighted by the probability of its occurrence, is incorporated into banks' optimal capital choice. An assumption of homogenous banks would lead to unrealistic outcomes where all banks simultaneously either pay or do not pay the penalty. We therefore assume that banks are exposed to idiosyncratic shocks. This implies that there is a continuum of ex-post capital adequacy ratios across banks, and a time-varying small fraction of banks that have to pay penalties in each period. This idiosyncratic risk can reflect a number of individual bank characteristics, such as differing success at raising non-interest income and minimizing non-interest expenses, where the sum of the two equals zero over all banks.

Specifically, we assume that at the beginning of period $t+1$ each individual wholesale bank draws a shock ω_{t+1}^b such that the idiosyncratic return on its loan book equals $r_{\ell,t+1}\omega_{t+1}^b$, where ω_{t+1}^b is a unit mean lognormal random variable distributed independently over time and across banks. The standard deviation of $\ln(\omega_{t+1}^b)$ equals $\left(\sigma^b\right)^2$, and the density function and cumulative density function of ω_{t+1}^b are denoted by $f^b(\omega_{t+1}^b)$ and $F^b(\omega_{t+1}^b)$. The regulatory framework stipulates that banks have to pay a real penalty of $\chi\check{\ell}_t(j)$ at time $t+1$ if the difference between gross returns on the wholesale loan book and gross payments on the wholesale deposit book, plus retail deposit profits and net retail loan profits, is less than a fraction γ of the gross return on their loan book. In other words, a penalty is payable if at time $t+1$ net worth is less than γ times the value of assets, so that γ can be interpreted as the Basel MCAR. Banks' actually maintained capital adequacy ratio γ_t^a remains considerably above the minimum requirement γ, because by maintaining an optimally chosen buffer banks protect themselves against the risk of penalties while minimizing the cost of excess capital. The penalty cutoff condition for bank j is given by

$$r_{\ell,t+1}\check{\ell}_t(j)\omega_{t+1}^b - r_{t+1}\check{d}_t(j) + \check{\Pi}_{t+1}^R(j)x - \check{\Lambda}_{t+1}^\ell(j)x < \gamma r_{\ell,t+1}\check{\ell}_t(j)\omega_{t+1}^b \ , \tag{6}$$

where the terms $\check{\Pi}_{t+1}^R(j)$ and $\check{\Lambda}_{t+1}^\ell(j)$ represent bank j's shares of the profits of retail deposit banks and the net losses of retail lending banks, and where the shares are lump-sum and pro-rated in proportion to each bank's net worth. We denote the cutoff idiosyncratic shock to loan returns below which the MCAR is breached ex-post by $\bar{\omega}_t^b$. The formula for $\bar{\omega}_t^b$ follows directly from (6). Banks choose the volume of loans to maximize their pre-dividend net worth:

$$\underset{\check{\ell}_t(j)}{Max}\ E_t\left[r_{\ell,t+1}\check{\ell}_t(j)\omega_{t+1}^b - r_{t+1}\check{d}_t(j) + \check{\Pi}_{t+1}^R(j)x - \check{\Lambda}_{t+1}^\ell(j)x - \chi\check{\ell}_t(j)F^b\left(\bar{\omega}_{t+1}^b\right)\right] \ ,$$

where the last term equals expected penalties due to violations of the MCAR. We arrive at post-dividend net worth by deducting dividends $\delta^b \check{n}_t^b$, which are paid out to households in a lump-sum fashion, a specification that can be obtained by applying the "extended family" approach of Gertler and Karadi (2011). In the ILF model this approach will also be applied to entrepreneurs. The optimality condition for wholesale loans is

$$E_t \left\{ r_{\ell,t+1} - r_{t+1} - \chi \left(F^b\left(\bar{\omega}_{t+1}^b\right) + f^b\left(\bar{\omega}_{t+1}^b\right) \left(\frac{r_{t+1} + \frac{\check{\Pi}_{t+1}^R x}{\check{n}_t^b}}{(1-\gamma)\, r_{\ell,t+1} \frac{\check{\ell}_t}{\check{n}_t^b}} \right) \right) \right\} = 0 \,. \tag{7}$$

In this condition the index j has been dropped because in equilibrium all wholesale banks face the same prices and risks and receive the same lump-sum distributions relative to their net worth. Equation (7) therefore features unique and identical loans-to-net-worth ratios, and therefore also unique and identical deposits-to-net-worth ratios, across all banks. The equation states that banks' wholesale lending rate is at a premium over the policy rate, by a margin that depends on the size of the MCAR γ, the penalty coefficient χ for breaching the MCAR, and expressions $F^b\left(\bar{\omega}_{t+1}^b\right)$ and $f^b\left(\bar{\omega}_{t+1}^b\right)$ that reflect the expected riskiness of banks and therefore the likelihood of a breach. Banks' retail lending rate, whose determination is discussed in the next subsection, is at another premium over the wholesale lending rate, to compensate for the bankruptcy risk of borrowers. A sensible interpretation of the wholesale rate is therefore as the rate that a bank would charge to a hypothetical borrower (not present in the model) with zero default risk.

3.2.3. Retail Lending Banks

Borrowers of retail lending banks have unit mass and are indexed by j, where individual borrowers differ by the size of their internal funds. Each borrower uses an optimally chosen combination of bank loans $\check{\ell}_t(j)$ and internal funds to purchase physical capital $\check{k}_t(j)$ at the market price q_t. The financial return to capital is given by $ret_{k,t} = \left(q_t\,(1-\Delta) + r_t^k\right)/q_{t-1}$, where Δ is the physical depreciation rate and r_t^k is the rental rate of capital. After the asset purchase, at the beginning of period $t+1$, each individual borrower draws a shock ω_{t+1}^k such that his idiosyncratic return to capital equals $ret_{k,t+1}\omega_{t+1}^k$, where ω_{t+1}^k is a unit mean lognormal random variable distributed independently over time and across borrowers. The standard deviation of $\ln(\omega_{t+1}^k)$, σ_{t+1}^k, is itself a first-order autoregressive stochastic process that we will refer to as the borrower riskiness shock. The density function and cumulative density function of ω_{t+1}^k are given by $f^k(\omega_{t+1}^k)$ and $F^k(\omega_{t+1}^k)$.

Each borrower j receives a loan contract from the bank. This specifies a nominal loan amount $L_t(j)$, a gross nominal retail rate of interest $i_{r,t}$, payable in $t+1$ as long as ω_{t+1}^k is sufficiently high to avoid default, and the fraction κ_t of the value of capital against which the bank is willing to lend (we will refer to this as pledged capital), because it expects to be able to seize and sell it in a bankruptcy. Both $i_{r,t}$ and κ_t are identical across all loans. The fraction κ_t is a first-order autoregressive stochastic process, and we will refer to it as the willingness-to-lend shock. An important difference between our model and those of Bernanke et al. (1999) and Christiano et al. (2014) is that the interest rate $i_{r,t}$ is assumed to be pre-committed in period t, rather than being determined in period $t+1$ after the realization of time $t+1$ aggregate shocks.[25] The latter assumption ensures zero ex-post profits for banks at all times, while under our debt contract banks make zero expected profits, but realized ex-post profits generally differ from zero. A borrower j who draws ω_{t+1}^k below a cutoff level $\bar{\omega}_{t+1}^k(j)$ cannot pay the interest rate $i_{r,t}$ and declares bankruptcy. He must hand over all his

[25] See Bernanke et al. (1999): "... conditional on the ex-post realization of R_{t+1}^k, the borrower offers a (*state-contingent*) *non-default* payment that guarantees the lender a return equal in expected value to the riskless rate."

pledged capital, which excludes the fraction $(1-\kappa_t)$ against which banks did not lend, to the bank. The bank can only recover a fraction $(1-\xi)$ of the asset value of such borrowers. The remaining fraction represents monitoring costs. Retail lending banks' ex-ante zero profit condition for loans to borrower j, in real terms, is therefore given by

$$E_t\left\{r_{\ell,t+1}\check{\ell}_t(j)-\left[\left(1-F^k(\bar{\omega}^k_{t+1}(j))\right)r_{r,t+1}\check{\ell}_t(j)+(1-\xi)\int_0^{\bar{\omega}^k_{t+1}(j)}\kappa_t q_t\check{k}_t(j)ret_{k,t+1}\omega^k f^k(\omega^k)d\omega^k\right]\right\}=0\ .$$

The ex-post cutoff productivity level is determined by equating, at $\omega^k_t(j)=\bar{\omega}^k_t(j)$, the gross interest charges payable by the borrower in the event of continuing operations $r_{r,t}\check{\ell}_{t-1}(j)$, in other words the cost of not defaulting, to the gross idiosyncratic return on the borrower's asset that needs to be handed over to the bank in the event of not continuing operations, $\kappa_{t-1}ret_{k,t}q_{t-1}\check{k}_{t-1}(j)\bar{\omega}^k_t(j)$, in other words the cost of defaulting. We denote the lender's gross share in pledged capital earnings by $\Gamma_{t+1}(j)=\Gamma(\bar{\omega}^k_{t+1}(j))$, and the lender's monitoring costs share in pledged capital earnings by $\xi G_{t+1}(j)=\xi G(\bar{\omega}^k_{t+1}(j))$. The borrower is left with a share $1-\kappa_t\Gamma_{t+1}(j)$ of total asset earnings. Then we can express the zero profit condition of banks in a way that determines the retail lending rate:

$$E_t\left\{\left(1-F^k(\bar{\omega}^k_{t+1})\right)\frac{r_{r,t+1}}{r_{\ell,t+1}}+(1-\xi)\,\kappa_t G_{t+1}\frac{ret_{k,t+1}}{r_{\ell,t+1}}\frac{q_t\check{k}_t}{\check{\ell}_t}\right\}=1\ . \tag{8}$$

In this condition the index j has been dropped because in equilibrium all borrowers face the same prices and risks. Equation (8) therefore determines unique and identical balance sheet ratios across all borrowers. Equation (8) says that the bank will set the unconditional lending rate such that its expected earnings are sufficient to cover the opportunity cost of the loan plus monitoring costs. The remainder of the analysis is similar to Bernanke et al. (1999), except for the fact that the lending rate is not conditional on period $t+1$ shock realizations, and for the presence of the willingness-to-lend coefficient κ_t, which is fixed at 1 in Bernanke et al. (1999). Specifically, the borrower selects the optimal level of investment by maximizing $E_t\left\{(1-\kappa_t\Gamma_{t+1}(j))\,q_t\check{k}_t(j)ret_{k,t+1}\right\}$ subject to (8). The details of the derivations differ depending on whether the borrower is a household, as in the FMC Model, or an entrepreneur, as in the ILF Model, but the final optimality conditions are identical. Retail lending banks make net loan losses if a larger than anticipated number of borrowers defaults, so that ex-post banks find that they have set their pre-committed retail lending rate at an insufficient level. Banks' ex-post loan losses, again after dropping the index j, are given by $\check{\Lambda}^\ell_t x=r_{\ell,t}\check{\ell}_{t-1}-\kappa_{t-1}q_{t-1}\check{k}_{t-1}ret_{k,t}\left(\Gamma_t-\xi G_t\right)$.

3.3. Manufacturers and Unions

Manufacturers have unit mass and are indexed by j, where individual manufacturers differ by the goods variety that they produce and sell. They purchase an aggregate of labour services $h_t(j)$ from unions, at the aggregate real producer wage rate $\check{w}^{pr}_t$, and capital services $k_{t-1}(j)$ from households (in the FMC model) or entrepreneurs (in the ILF model), at the rental rate r^k_t. Their differentiated output $y_t(j)$ is sold, at price $P_t(j)$, to be used for consumption, investment, government spending, monitoring activities and monetary transactions costs. In each case, demand is for a CES aggregate over individual output varieties, with elasticity of substitution θ_p, and thus with a gross mark-up $\mu_p=\theta_p/(\theta_p-1)$. The production function of an individual manufacturer is given by

$$y_t(j)=\left(a_tT_th_t(j)\right)^{1-\alpha}k_{t-1}(j)^{\alpha}\ , \tag{9}$$

where the productivity shock a_t is a first-order autoregressive stochastic process. Optimality conditions for cost minimization are standard. Each manufacturer faces price adjustment costs that are quadratic in changes in the rate of price inflation. In real normalized terms, they are given by $\check{G}_{P,t}(j) = \frac{\phi_p}{2}\check{y}_t\left(\pi_t^p(j)/\pi_{t-1}^p - 1\right)^2$. The optimality condition for price setting is a standard New Keynesian Phillips curve.

Unions have unit mass and are indexed by j, where individual unions differ by the labour variety that they sell. Unions purchase homogenous labour services from households, at the real household wage rate $\check{w}_t^{hh}$, and sell differentiated labour varieties to manufacturers, at the real union-specific producer wage rate $\check{w}_t^{pr}(j)$. Manufacturers demand a CES aggregate over individual labour varieties, with elasticity of substitution θ_w, and thus with gross mark-up $\mu_w = \theta_w/(\theta_w - 1)$. Each union faces wage adjustment costs that are quadratic in changes in the rate of wage inflation. In real normalized terms, they are given by $\check{G}_{W,t}(j) = \frac{\phi_w}{2}\check{w}_t^{pr}h_t\left(\pi_t^w(j)/\pi_{t-1}^w - 1\right)^2$. The optimality condition is a standard New Keynesian Phillips curve for wage setting.

3.4. The ILF Model: Households and Entrepreneurs

3.4.1. Households

Households have unit mass and are indexed by j. Households consume $c_t(j)$ and supply labour hours $h_t(j)$. We assume that each household holds identical initial stocks of all physical and financial assets, and receives identical lump-sum dividends from banks, entrepreneurs, manufacturers and unions. Then the index j can be dropped except when stating the lifetime utility function, where individual consumption $c_t(j)$ and aggregate consumption c_t need to be distinguished because of external habit persistence:

$$Max \quad E_0\sum_{t=0}^{\infty}\beta^t\left\{(1-\frac{v}{x})\log(c_t(j) - vc_{t-1}) - \psi\frac{h_t(j)^{1+\frac{1}{\eta}}}{1+\frac{1}{\eta}}\right\} . \tag{10}$$

Households hold real deposit money balances $\check{d}_t^c$ and $\check{d}_t^i$ that lower the cost of transactions, as in Schmitt-Grohé and Uribe (2004). Specifically, defining velocities as $v_t^c = \check{c}_t/\check{d}_t^c$ and $v_t^i = \check{I}_t/\check{d}_t^i$, we have transactions costs of

$$s_t^z = A^z v_t^z + \frac{B^z}{v_t^z} - 2\sqrt{A^z B^z} \ \ , \ \ z \in \{c,\, i\} \ . \tag{11}$$

At the beginning of each period the household splits into two groups, consumers/workers/capital holders (CWC) and capital goods producers. Inside each period capital goods producers purchase the depreciated capital stock $q_t\check{k}_{t-1}(1-\Delta)$, investment goods $\check{I}_t$, and resources to pay for monetary transactions costs and real investment adjustment costs $s_t^i\check{I}_t + \check{G}_{I,t}$, where $\check{G}_{I,t} = (\phi_i/2)\,\check{I}_t\left(\left(\check{I}_t/\check{I}_{t-1}\right) - 1\right)^2$. They then sell the sum of old and new capital $q_t\left(\check{k}_{t-1}(1-\Delta) + \check{I}_t\right)$. Households receive the dividends of entrepreneurs $\delta^k\check{n}_t^k$ plus a lump-sum net income stream $\check{\Omega}_t$ that consists of the dividends of manufacturers, unions and banks, plus adjustment cost payments, minus government lump-sum taxes. Household j's budget constraint, in real normalized form and cancelling terms involving $\check{k}_{t-1}(1-\Delta)$, is then given by

$$\check{d}_t = \frac{r_{d,t}}{x}\check{d}_{t-1} + \left(q_t - 1 - s_t^i\right)\check{I}_t - \check{G}_{I,t} - \check{c}_t\left(1 + s_t^c\right) + \check{w}_t^{hh}h_t + \check{\Omega}_t + \delta^k\check{n}_t^k \ . \tag{12}$$

The first-order condition for hours worked is standard. The two first-order conditions for deposit money contain monetary wedges whereby the intertemporal marginal rate of substitution equals

$(1-s_t^{z'}(v_t^z)^2)/r_{d,t+1}$, $z \in \{i,c\}$. The first-order conditions for investment and consumption are given by:

$$q_t = \left(1 + s_t^i + s_t^{i'} v_t^i\right) + \phi_i \left(\frac{\check{I}_t}{\check{I}_{t-1}}\right)\left(\frac{\check{I}_t}{\check{I}_{t-1}} - 1\right) - E_t \beta \frac{\check{\lambda}_{t+1}}{\check{\lambda}_t} \phi_i \left(\frac{\check{I}_{t+1}}{\check{I}_t}\right)^2 \left(\frac{\check{I}_{t+1}}{\check{I}_t} - 1\right) \;, \tag{13}$$

$$\frac{(1-\frac{v}{x})}{\check{c}_t - \frac{v}{x}\check{c}_{t-1}} = \check{\lambda}_t \left(1 + s_t^c + s_t^{c\prime} v_t^c\right) \;. \tag{14}$$

These conditions contain monetary wedges $1 + s_t^z + s_t^{z'} v_t^z$, $z \in \{c,i\}$, that we will refer to as the effective prices of investment and consumption. These are higher than the market prices, which equal one in real terms, with mark-ups that are due to monetary frictions, and that are formally equivalent to distortionary capital income and consumption tax rates $\tau_{k,t}^{\ell iq} = s_t^i + s_t^{i'} v_t^i$ and $\tau_{c,t}^{\ell iq} = s_t^c + s_t^{c\prime} v_t^c$. For investment, we can see this equivalence by noting that the steady state gross return to capital in the presence of proportional capital income taxes τ_k (under the simplifying assumption that $\Delta = 0$) would be given by $1 + r^k(1-\tau_k)/(1+\tau_k^{\ell iq})$. For consumption, we can see the equivalence by noting that the steady state ratio of the marginal utilities of consumption and wealth in the presence of proportional consumption taxes τ_c would be given by $(1+\tau_c)\left(1+\tau_c^{\ell iq}\right)$. The effects of changes in the size of bank balance sheets, and therefore in deposit creation, will thus be transmitted to the real economy in a formally equivalent way to changes in distortionary tax rates. The distortion is a shortage, relative to the Friedman rule, of bank deposits, a shortage that can never be completely eliminated because the cost of the creation of bank deposits can never go to zero. As we can see in equations (4) and (7), this is due to a combination of regulation, precautionary capital buffers and market power.

3.4.2. Entrepreneurs

Entrepreneurs have unit mass and are indexed by j, where individual entrepreneurs differ by the size of their net worth. The entrepreneur's objective function is to maximize

$$q_t \check{k}_t(j) \frac{ret_{k,t+1}}{x} \left(1 - \kappa_t \Gamma_{t+1}(j)\right) \;, \tag{15}$$

subject to the bank's zero profit condition (8). The optimality condition for capital is (except for the presence of κ_t) identical to Bernanke et al. (1999) and Christiano et al. (2014):

$$E_t \left\{ \frac{ret_{k,t+1}}{r_{\ell,t+1}} \left(1 - \kappa_t \Gamma_{t+1}\right) + \tilde{\lambda}_{t+1} \left(\frac{ret_{k,t+1}}{r_{\ell,t+1}} \kappa_t \left(\Gamma_{t+1} - \xi G_{t+1}\right) - 1 \right) \right\} = 0 \;, \tag{16}$$

where we have dropped the index j by exploiting the fact that all entrepreneurs have identical balance sheet ratios. The variable $\tilde{\lambda}_{t+1}$ equals $\Gamma_{t+1}^{\omega} / \left(\Gamma_{t+1}^{\omega} - \xi G_{t+1}^{\omega}\right)$, where Γ_{t+1}^{ω} and G_{t+1}^{ω} are the partial derivatives of Γ_{t+1} and G_{t+1} with respect to $\bar{\omega}_{t+1}^k$, and where $\tilde{\lambda}_{t+1}$ represents an indicator of the tightness of bank lending conditions. Aggregate entrepreneur net worth is $\check{n}_t^k = q_t \check{k}_t - \check{\ell}_t$. Its evolution, in real normalized form, is given by

$$\check{n}_t^k = q_{t-1} \check{k}_{t-1} \frac{ret_{k,t}}{x} \left(1 - \kappa_{t-1} \Gamma_t\right) - \delta^k \check{n}_t^k \;. \tag{17}$$

3.5. The FMC Model: Representative Household

In the FMC model, a single representative household, with identical preferences to (10), takes over the functions of the entrepreneur of the ILF model, and both borrows from and holds deposits with banks. The household budget constraint, in real normalized terms, is given by

$$\begin{aligned} & q_t\check{k}_t + \check{d}_t - \check{\ell}_t \\ = \; & q_{t-1}\check{k}_{t-1}\frac{ret_{k,t}}{x}\left(1-\kappa_{t-1}\Gamma_t\right) + \frac{r_{d,t}}{x}\check{d}_{t-1} \\ & + \left(q_t - 1 - s_t^i\right)\check{I}_t - \check{G}_{I,t} - \check{c}_t\left(1+s_t^c\right) + \check{w}_t^{hh}h_t + \check{\Omega}_t \; . \end{aligned} \tag{18}$$

This consolidates the budget constraints of households and entrepreneurs of the ILF model, and this represents the only difference between the two model classes. The optimality conditions for consumption, investment and bank deposits are identical to those of the representative household in the ILF model. The optimality condition for capital is identical to that of the entrepreneur in the ILF model.

3.6. Government and Market Clearing

Monetary policy follows a conventional inflation-forecast-based interest rate rule

$$i_t = \left(i_{t-1}\right)^{m_i}\left(\bar{\imath}\right)^{(1-m_i)}\left(\frac{\pi_{4,t+3}}{\bar{\pi}}\right)^{(1-m_i)m_\pi} \quad , \tag{19}$$

where $\pi_{4,t} = \left(\pi_t\pi_{t-1}\pi_{t-2}\pi_{t-3}\right)^{\frac{1}{4}}$, and where the specific expression for the steady state nominal interest rate, $\bar{\imath} = (x(1-\bar{s}^{c'}(\bar{v}^c)^2)\bar{\pi})/\beta s$ follows from the optimality conditions for consumption deposits of households and of retail deposit banks. Government spending equals a fixed fraction s_g of steady state output $\bar{y}$, $\check{g}_t = s_g\bar{y}$, and the government budget is balanced in each period by way of lump-sum taxes, $\check{\tau}_t = \check{g}_t$. Assuming that initial government debt equals zero, government debt therefore remains at zero at all times.

The goods market clearing condition is given by

$$\check{y}_t = \check{c}_t + \check{I}_t + \check{g}_t + \check{\mathcal{M}}_t^b + \check{\mathcal{M}}_t^k + \check{\mathfrak{T}}_t^c + \check{\mathfrak{T}}_t^i \quad , \tag{20}$$

where $\check{\mathcal{M}}_t^b$ are bank regulatory penalties, $\check{\mathcal{M}}_t^k$ are loan monitoring costs, and $\check{\mathfrak{T}}_t^c + \check{\mathfrak{T}}_t^i$ are monetary transactions costs, $\check{\mathfrak{T}}_t^c = \check{c}_t s_t^c$ and $\check{\mathfrak{T}}_t^i = \check{I}_t s_t^i$. Finally, $g\check{d}p_t = \check{c}_t + \check{I}_t + \check{g}_t$.

3.7. Calibration

The calibration of the model is based on U.S. historical data and on the DSGE literature insofar as it relates to the United States. For balance sheet and spreads data, we focus on the period prior to the Great Recession. One period corresponds to one quarter.

We calibrate both the steady state real growth rate x and the steady state inflation rate $\bar{\pi}$ at 2% per annum, and we fix β to set the model's risk-free real interest rate at 3% per annum. The parameter α is calibrated to obtain a steady state labour share of 60%. This is in line with Bureau of Labor Statistics (BLS) data for the U.S. business sector. The private investment to GDP ratio is set to 20% of GDP, roughly its average in pre-crisis U.S. data. The implied depreciation rate Δ, at around 8% per annum, is in the middle of the range of values typically used in the literature.

The government spending to GDP ratio is set to its approximate historical average of 18% of GDP, by fixing the parameter s_g.

For household preferences, we set habit persistence at $v = 0.75$ and the inverse of the labour supply elasticity at $\eta = 1$. The parameter ψ is calibrated to normalize steady state labour supply to 1. The steady state price and wage mark-ups of monopolistically competitive manufacturers and unions are fixed, in line with many papers in the New Keynesian literature, at 10%, or $\mu_p = \mu_w = 1.1$. Wage and price stickiness parameters are calibrated as $\phi_w = 300$ and $\phi_p = 300$, which corresponds to average contract lengths of 6.25 quarters in a Calvo (1983) model with full indexation to past inflation. The investment adjustment cost parameter is calibrated at $\phi_i = 2.5$, following Christiano et al. (2005). The monetary policy rule is calibrated at $m_i = 0.7$ and $m_\pi = 3.0$.

For interest rate spreads, we fix the parameter $\bar{\sigma}^k$ to calibrate the steady state spread of the retail lending rate $i_{r,t}$ over the policy rate i_t at 1.5%, consistent with the evidence in Ashcraft and Steindel (2008). We fix the parameter χ to calibrate the steady state spread of the wholesale lending rate $i_{\ell,t}$ over the policy rate at 0.5%, consistent with the approximate historical spread of the 3-month US$ LIBOR over the 3-month treasury bill rate, and we fix the parameter s to calibrate the steady state spread of the policy rate over the deposit rate $i_{d,t}$ at 1%, again consistent with the evidence in Ashcraft and Steindel (2008). For consistency with the specifications of Bernanke et al. (1999) and Christiano et al. (2014) we set the steady state value of the willingness-to-lend parameter $\bar{\kappa}$ to 1. We fix the parameter ξ to calibrate a quarterly loan default rate of 1.5%, consistent with the evidence for non-financial listed US companies in Ueda and Brooks (2011).

The parameters A^i and A^c determine the overall demand for bank liabilities and therefore the size of bank balance sheets. We use them to calibrate the steady state share of investment-related deposits in overall deposits at 50%, and the steady state borrower leverage ratio, defined as loans divided by the difference between the value of physical capital and loans, at 100%. The latter is consistent with the evidence in Ueda and Brooks (2011). We set the money demand parameters B^j, $j \in \{c, i\}$, to obtain steady state interest semi-elasticities of money demand of 0.05, based on the estimates in Ball (2001). For the banking sector, we calibrate bank riskiness σ^b such that the percentage of banks violating the minimum capital adequacy ratio equals 2.5% of all banks per quarter. The parameter determining the Basel minimum capital adequacy ratio is set to 8% of assets, $\gamma = 0.08$. The parameter δ^b is calibrated to be consistent with the assumption that banks maintain an average actual capital adequacy ratio of 10.5%, which means that they maintain a capital conservation buffer of 2.5%, as envisaged under Basel III. Together with our assumptions about household leverage and money demand, this implies an overall volume of bank lending for physical investment purposes equal to 120% of GDP. This is broadly in line with the data. Using Flow of Funds data we find that in 2006 total credit market debt of non-financial businesses with maturities of more than one year reached around 60% of GDP, while residential mortgages[26] reached around 80% of GDP. For our simulations, the first-order autoregressive coefficients of borrower riskiness shocks, willingness-to-lend shocks and technology shocks are set to 0.9, 0.9 and 0.95.

[26] In the data investment in fixed capital also includes residential housing investment.

4. Model Comparison: Impulse Responses

Impulse responses for three shocks are presented in Figures 1-3. In each case the size of shocks is identical across models, with the dashed and solid lines representing the ILF and FMC models. All variables are shown in deviations from trend. The retail lending spread is defined as the annualized percentage point difference between the retail lending rate and the policy rate. The bank leverage ratio is defined as the ratio of bank assets to bank net worth.

4.1. Credit Boom due to Higher Willingness to Lend

Figure 1 shows impulse responses for a sequence of unanticipated shocks whereby, over a period of six quarters, the share of the value of capital against which banks are willing to lend, κ_t, increases by 30 percentage points, and then gradually returns to its original value.[27] With this shock the driver of the cycle is a change in the terms of banks' lending policies rather than, as for borrower riskiness shocks in Figure 2 below, a change in the fundamentals of borrowers' business. At banks' existing balance sheet and pricing structure, their anticipated profitability immediately following these shocks is significantly improved. They therefore respond through a combination of higher lending volumes and lower lending spreads. The main difference between the two models is that there is a much larger change in lending volumes in the FMC model and a larger change in lending spreads in the ILF model. In each case the shock, which directly mainly affects capital investment, has a stimulative effect on aggregate demand through an investment boom. Inflation increases and peaks at the end of the second year, accompanied by an increase in the real policy rate.

ILF banks cannot quickly change their lending volume because, in order to lend more to borrowers, savers have to make sufficient deposits of physical savings. Savers, due to their desire to smooth consumption and labor supply, are only willing to do so gradually over time. Banks therefore mostly respond to the favorable shock by reducing their lending spread, by almost 90 basis points at the end of the first year. In the FMC model on the other hand the quantity of loans can increase rapidly irrespective of physical saving. The main constraint is banks' and their borrowers' expectation of the gains from additional loan and deposit creation, which improves with each successive increase in the parameter κ_t. Banks therefore increase their lending volume very significantly, by 13% after six quarters, compared to only around 1% in the ILF model. They fund the additional lending by creating additional deposits, and they do so almost one-for-one, as initial movements in net worth are comparatively small.[28] The much greater increase in lending in the FMC model leads to higher loan-to-value ratios among bank borrowers. This in turn implies a *ceteris paribus* increase in lending risk, which in this simulation completely offsets the effect on lending risk and spreads of the favorable shocks. As a result, 100% of banks' initial response to the willingness-to-lend shock consists of quantity rather than price changes, while in the ILF model the opposite is true. Even by the end of the sixth quarter FMC banks have only reduced their lending spread by 20 basis points, compared to 90 basis points for ILF banks. These differences are also evident in the behavior of bank leverage. In the ILF model bank leverage is acyclical because neither bank net worth nor loans change significantly on impact. In the FMC model leverage is highly procyclical, due to a rapid increase in lending at initially nearly constant net worth.

[27] Note that $\kappa_t > 1$ does not imply that loans exceed the value of physical capital, because in steady state, where $\bar{\kappa} = 1$, loans equal 50% of the value of physical capital.

[28] Shocks to κ_t mostly affect future loan contracts, and unlike shocks to σ_t^k have few effects on the performance of existing loans, and therefore on banks' profits and net worth.

As for the effects on the real economy, in the FMC model investment increases by almost 8% by the sixth quarter, compared to 5% in the ILF model. Furthermore, while consumption decreases by 0.4% in the ILF model, it increases by 0.6% in the FMC model. As a result, the 1.8% expansion in GDP in the FMC model by the end of the sixth quarter is more than twice as large as in the ILF model. There are two reasons for these differences. The first is that in the ILF model a drop in consumption helps, along with an increase in labor supply, to generate the physical savings that allow deposits and therefore lending to rise. This channel is absent in the FMC model. The second reason is that the large differences in the creation of liquidity by banks lead to large differences in the effective prices of investment and consumption. In the ILF model deposit creation is slower than the growth in real activity, so that the effective prices of investment and consumption increase slightly, while in the FMC model deposit creation is far faster than the growth in real activity, so that the effective price of investment drops by almost 4% by the end of the sixth quarter. This is roughly equivalent to a very sizeable 4 percentage point drop in a capital income tax rate, which has large effects because this drop is also quite persistent. Relative to the ILF model, the very large stimulus effect of the additional liquidity more than offsets the much smaller drop in lending spreads and the much larger increase in the policy rate, which increases by around 120 basis points by the end of the sixth quarter (with further increases thereafter) while in the ILF model it only increases by around 20 basis points. More importantly liquidity, once created, thereafter circulates to wherever it is most effective at facilitating trade. In our model, liquidity can be used not only for investment purchases but also for consumption purchases. There is therefore also a large stimulus to consumption, whose effective price drops by almost 1.5% by the end of the sixth quarter, equivalent to a roughly 1.5 percentage point drop in a consumption tax rate. This is the main reason why consumption increases in the FMC model while it decreases in the ILF model.

We emphasize that the circulation of bank-created liquidity, together with its strong procyclicality in FMC models, offers a natural transmission channel that generates, despite the complete absence of bank lending for consumption purposes, a positive comovement of consumption and investment, one of the most robust empirical features of business cycles. Because the response of deposit creation in ILF models is so much slower, this transmission channel is missing, and the comovement of consumption and investment is much weaker, and in fact in our simulation is negative.[29]

4.2. Credit Crash due to Higher Borrower Riskiness

Figure 2 shows impulse responses for a shock whereby the standard deviation of borrower riskiness, σ_t^k, increases by 15% on impact and then gradually returns to its original value. Banks' profitability and net worth immediately following this shock worsen significantly at their existing balance sheet and pricing structure, with banks losing around 4% of their net worth on impact, and with a larger share of banks violating the MCAR. In order to protect their profitability and capital adequacy, banks respond through a combination of lower lending volumes and higher lending spreads. The main difference between the two models is again a much larger change in lending volumes in the FMC model and a larger change in lending spreads in the ILF model. In each case, the tightening of bank lending leads to a significant contraction in investment, which is the main reason for the contraction in GDP. Lower demand leads to lower inflation, which in turn leads to a reduction in the real policy rate, but this effect comes later and is much weaker in the ILF model.

[29]This comovement can be positive in ILF models, as shown in Christiano et al. (2014). However, in otherwise identical FMC models it will always be much stronger.

For ILF banks, deposits and therefore loans can only decline gradually through lower physical saving, by households reducing their labour supply and, despite the economic contraction, increasing their consumption. Banks therefore continue to lend to borrowers that have become much riskier, and to compensate for this risk they increase the retail lending spread by almost 200 basis points on impact. In the FMC model the quantity of loans can immediately drop irrespective of physical saving, with households using some of their deposits to repay outstanding loans while borrowing costs are high. There is therefore a large and discrete contraction of banks' balance sheets, of around 5% on impact in a single quarter (with almost no initial change in the ILF model), as deposits and loans shrink simultaneously. Because, *ceteris paribus*, this cutback in lending reduces borrowers' loan-to-value ratios and therefore the riskiness of the remaining loans, banks only increase their lending spread by around 50 basis points on impact, compared to almost 200 basis points in the ILF model. In the ILF model bank leverage increases on impact, because the immediate net worth losses dominate the gradual decrease in loans. In the FMC model leverage drops, because the rapid cutback in lending is even larger than the drop in net worth. In the FMC model bank leverage is therefore procyclical, while in the ILF model it is countercyclical.

As for the effects on the real economy, the contraction in GDP in the FMC model is almost 50% larger than in the ILF model, and this is entirely due to the behavior of consumption. The large and highly persistent drop in investment is in fact slightly larger in the ILF model, as the much larger increase in the lending spread in that model dominates the much larger increase in the effective price of investment in the FMC model (equivalent to a 1.8 percentage point increase in a capital income tax rate on impact) that is due to the rapid contraction in bank-created liquidity. But consumption increases by 0.1% on impact in the ILF model, while it decreases by around 0.3% in the FMC model, thereby comoving with investment. The main reason is the large increase in the effective price of consumption in the FMC model (equivalent to a 0.65 percentage point increase in a consumption tax rate on impact), compared to a slight decrease in the ILF model. Inflation and the policy rate drop by 20-30 basis points more in the FMC model compared to the ILF model, which further reduces the overall increase in interest rates compared to the ILF model.

4.3. Contractionary Technology Shock

Figure 3 shows impulse responses for a shock whereby the level of labor-augmenting technology, a_t, drops by 7.5% on impact and then gradually returns to its original value. This shock depresses consumption, investment and the value of loan collateral. The combination of the resulting contraction in bank net worth and the contraction in real activity, and therefore in the demand for the liquidity provided by bank deposits, causes banks to cut back on lending and deposit creation.

In the ILF model, the slowdown in bank lending, which requires a decumulation of physical savings, is significantly slower and smaller than the contraction in real activity. As a result, liquidity becomes more abundant relative to output, and the effective prices of investment and consumption drop by the equivalent of 2.0 and 0.75 percentage point cuts in capital income and consumption tax rates. This provides a monetary stimulus that helps to buffer the effects of the contractionary real shock. In the FMC model, the slowdown in bank lending is unrelated to physical savings, and is therefore much faster and, over the first three years, more than 50% larger than in the ILF model. This is closer to the drop in real activity, so that the effective prices of investment and consumption drop by significantly less. The stimulus from more abundant liquidity is therefore smaller, and the overall contraction in real activity is larger, although the difference is small compared to the direct GDP effects of the technology shock. As for inflation, the smaller drop in the effective prices of investment and consumption in the FMC model implies a larger contraction in aggregate demand relative to aggregate supply, and therefore a smaller increase in inflation and the policy rate. The initial drop in the real policy rate is explained by interest rate smoothing.

5. Stylized Facts

In this section we present stylized facts that relate to the FMC model's most important predictions from Section 4. First, changes in the size of aggregate financial sector balance sheets are large and rapid, and relative to this the accumulation of physical savings by the nonfinancial private sector (SAV), the net acquisition of securities or direct financing substitution (DFS) and asset valuation effects (VAL) are very different on average (SAV) and extremely small (DFS and VAL), as well as far less volatile. Second, financial sector book value leverage tends to be procyclical. Third, quantity rationing of credit plays an important role in the financial sector's response to shocks. To study the first two predictions we examine the financial sector data of the US, the eurozone (EUR), Germany (GER) and France (FRA). We emphasize that because of the scope of the data, with one partial exception, movements of balance sheet exposures between different parts of the financial sector will not be visible because they disappear upon consolidation, leaving only changes in exposures between the financial and nonfinancial sectors, the quantity of interest in most macroeconomic models. Specifically, US data represent the unconsolidated sum of regulated banks (depository institutions, credit unions) and shadow banks (finance companies, security brokers and dealers), while EUR, GER and FRA data represent the consolidated financial sector. The sample for GER and FRA starts in 2003Q2 due to data availability issues.

5.1. Large and Rapid Changes in Financial Sector Balance Sheets

We show that financial sector balance sheets exhibit large and rapid changes, at all points of the business cycle, and for all countries in our sample. As we have seen, FMC models naturally have this property, because they assign an important role to quantity adjustments to shocks. ILF models need to rely on a combination of three alternative mechanisms. The first, SAV, has its data counterpart in net private saving from the national accounts. The second, VAL, has its data counterpart in the difference, in flow of funds data, between net changes in financial sector asset stocks, which include valuation effects, and net new flows into assets, which mostly exclude them. The third, DFS, has its data counterpart in the net acquisition, in flow of funds flow data, of nonfinancial private sector securities by the financial sector.

We will study two concepts of changes in financial sector balance sheets, changes in total assets (Figures 4-5) and changes in total asset-side exposures to the nonfinancial private sector (Figures 6-10), which includes both loans and holdings of securities, and which will be denoted by d(LOANS&SEC). We note that LOANS&SEC represents a major part of total assets, with the difference consisting mostly of external assets, which is in large part cross-border interbank lending, holdings of government securities, and remaining assets, which is mainly financial derivatives. We note that while d(LOANS&SEC) is very volatile, changes in total assets are even more volatile. Changes in financial sector balance sheets can be presented either in terms of changes in stocks or in terms of flows. When we study the contributions of VAL and DFS effects to total effects (Figures 7, 8 and 9) we show changes in stocks because this is the relevant total for a comparison with VAL effects. For the remaining results we show flows.

A key requirement for a fully satisfactory decomposition into DFS and VAL effects is detailed data in three dimensions, namely, first, both stocks and flows of securities, second, both debt and equity securities, and third, a breakdown of securities that identifies the issuer as the nonfinancial business sector, because in asset-trading ILF models (and in FMC models) banks' securities exposures are to this sector.[30] This detail is fully available for the US but not for Europe, except for the very short sample period, starting in 2014Q1, covered by the EUR whom-to-whom tables. The problem for Europe is that over the longer sample the first and second requirements can only be met for a broad category of issuers called other residents, which includes not only nonfinancial business but also securitization vehicles. Because the latter merely represents a repackaging of loans that for our purposes are properly classified as loans rather than securities, this tends to overstate the importance of DFS and VAL effects. We nevertheless present the corresponding results in Figures 7 and 8, which show that even with this bias DFS and VAL effects are very small compared to d(LOANS&SEC). In the whom-to-whom tables VAL effects are very similar to Figure 7 while DFS effects are much smaller than in Figure 8 and extremely close to zero. Furthermore, the longer sample of European data is sufficient to compute, for nonfinancial business sector issuers, the sum of DFS and VAL effects, which only requires data on stocks of securities. The only limitation is that this can only be computed for the category of securities other than shares, which represent a little over half of total securities holdings by the financial sector. We present these results in Figure 9, which shows that the sum of DFS and VAL effects is even smaller than in Figures 7 and 8 and very close to zero, as in the US, with the latter furthermore illustrating that the addition of shares does not change this result.

5.1.1. Financial Sector Balance Sheets

Figure 4 is motivated by the work of Adrian et al. (2013), who show that the U.S. financial sector has responded to shocks mainly through one-for-one changes in its assets and debt, rather than in its assets and equity, where the latter is defined as total assets minus debt. This is true for both aggregate and micro-level data, for both commercial and shadow banks, and during both booms and recessions. Adrian et al. (2013) illustrate their results using scatter plots, based on data from FDIC call reports and SEC 10Q filings, that plot dollar changes in debt and equity of U.S. commercial banks against the corresponding dollar changes in assets. The large majority of their assets-debt pairs lie along a line with a slope of approximately one, while the assets-equity pairs are far more dispersed. We repeat a similar exercise in Figure 4[31], with three differences. First, our sample is different, we study quarter-on-quarter balance sheet changes of the 200 largest U.S. commercial banks, in the single quarter 2009Q4, a turbulent period with large changes in balance sheets. Second, we study percent changes in addition to dollar changes, because dollar changes give rise to outliers for very large institutions. Third, we also examine histograms of the distributions of the elasticities of debt and equity with respect to assets.

The exposition in terms of dollar changes in the top left panel of Figure 4 shows a very similar pattern to Adrian et al. (2013). The top right panel shows that for percent changes the assets-debt pairs continue to lie very close to a line with a slope of approximately one, while the assets-equity pairs continue to be much more widely dispersed. The histograms show that the elasticity of debt with respect to assets is clearly centered near 1, with few outliers, while the median and mean of

[30]The household sector is not a quantitatively significant issuer. See below for comments on securitized mortgages.

[31]The reason why Figures 4 and 5 show changes in total assets rather than d(LOANS&SEC), apart from consistency with Adrian et al. (2013), is that for d(LOANS&SEC) the allocation of equity between included and excluded asset classes would necessarily be arbitrary.

the elasticities of equity with respect to assets are far lower, and the elasticities are far more widely dispersed. Finally, for a large number of institutions the percent changes in assets and debt are very large, in some cases equalling 10% or more in this single quarter. In less turbulent quarters the magnitudes are somewhat smaller but the pattern remains the same.

Figure 5 performs a similar exercise using aggregate financial sector time series data of percent changes in assets, equity and debt for the US, EUR, GER and FRA. The pattern is similar to Figure 4, with the elasticity of debt with respect to assets very close to one across all countries, the elasticity of equity with respect to assets far lower on average and far more widely dispersed, and not infrequent changes in assets and debt of 4% or more in a single quarter. This shows that banking systems as a whole frequently exhibit simultaneous expansions or contractions of assets and debt of very large amounts. If total assets are replaced by loans, the magnitude of changes is somewhat smaller but the pattern remains the same.

5.1.2. Physical Saving (SAV)

Before studying the data, we reiterate that an emphasis on physical saving in explaining changes in financial sector balance sheets forces us to accept the notion that banks are traders of physical resources, a notion that should be ruled out on a priori grounds. Figure 6 demonstrates that in the data the relationship between the savings-to-GDP ratio SAV/GDP (and also the investment-to-GDP ratio INV/GDP) and two measures of financial sector lending, the change in credit to nonfinancial business relative to GDP and d(LOANS&SEC)/GDP, is very weak. The reasons were discussed in Section 1.

We focus our comments on the relationship between SAV/GDP and d(LOANS&SEC)/GDP. In the US, in the decades prior to the GFC, SAV/GDP fluctuated in a relatively narrow range of 4% to 9%, while d(LOANS&SEC)/GDP was much larger on average, as the size of the financial system grew relative to the economy, and fluctuated in a far wider range of 2% to 18%. The GFC itself exhibited an extremely large decline in d(LOANS&SEC)/GDP, from around +18% to -10%, and this was accompanied by a sizeable increase in SAV/GDP, from 4% to 9% over the same time period. The behavior of net private saving was therefore clearly disconnected from, and generally far smoother than, that of financial sector balance sheets. Figure 6 shows that the other three economies in our sample exhibit similar patterns to the US, with massive contractions in bank balance sheets at the onset of the GFC, accompanied by much smaller movements, and in two cases increases, in physical saving. The conclusion is that models that rely on physical saving as an explanation for observed variations in financial sector balance sheets have difficulties in explaining the data.

5.1.3. Valuation Effects (VAL)

The quantitative significance of VAL effects can be studied by two methods. The first, based on Figure 5, is indirect, and studies whether changes in financial sector assets have their principal counterpart in changes in equity, as would occur with valuation effects, rather than in debt. The second, based on Figure 7, directly computes valuation effects as the difference between asset stock changes and asset flows, and then compares them to changes in total asset stocks.[32]

In the data underlying Figure 5, approximately half of the variation in US financial sector leverage ratios (assets/equity) is due to percent changes in assets and the rest is due to percent changes in equity, and a similar decomposition obtains for EUR, GER and FRA. Given that assets are far

[32] VAL effects equal the sum of net provisions and write-offs on loans and market price fluctuations on securities.

larger than equity, the absolute changes in assets (and debt) are therefore an order of magnitude larger than those in equity. This, in turn, implies that valuation effects cannot be the dominant or even an important source of changes in total financial sector assets.

Figure 7 plots the time series data for d(LOANS&SEC)/GDP and VAL/GDP. We observe that, despite the bias in the European data discussed above, valuation effects only account for a very small share of changes in financial sector balance sheets, except during a few periods when both fluctuate around zero. In fact, as shown in Table 1, average VAL effects are negative and small[33] in each of the four economies in our sample, while average changes in financial sector balance sheets are positive and (with the exception of Germany over our sample) large. The conclusion is that models that rely on VAL effects as an explanation for observed variations in financial sector balance sheets have difficulties in explaining the data.

5.1.4. Direct Financing Substitution (DFS)

Figure 8 demonstrates that DFS effects also account for a very small share of changes in financial sector balance sheets. There are several reasons why this is not unexpected. First, for countries with predominantly bank-based rather than securities-market-based financial sectors (EUR, GER, FRA) there are fewer possibilities for DFS, yet the volatility of financial sector balance sheets in these economies is almost as large as in the US. Second, the large nonfinancial borrowers that can issue tradable bonds and equity only account for a small fraction of the economy's overall financing needs even in the US, and average holdings of such securities as a share of the financial sector's aggregate balance sheet are, in our sample, only 5.3% in the US, 5.7% in EUR, 5.2% in GER and 5.7% in FRA. Third, a large share of trading in securities is concentrated either within the financial sector, and therefore disappears upon its consolidation, or between nonfinancial firms and households, and therefore never appears in financial sector data.

DFS effects in asset-trading ILF models concern net purchases, by the financial sector from savers, of nonfinancial sector securities, accompanied by the net creation of deposits for savers. Figure 8 plots the time series data for d(LOANS&SEC)/GDP and DFS/GDP. We observe that DFS is extremely small in the US. In Europe, despite the bias in the data discussed above, DFS effects account for a share of changes in financial sector balance sheets that, while larger than for VAL effects, is nevertheless very small, except for Germany, which exhibited very low loan growth throughout our sample. This is confirmed by Table 1, which shows that average DFS effects are positive but much smaller than the growth in the sum of loans and securities. The conclusion is that models that rely on DFS effects as an explanation for observed variations in financial sector balance sheets have difficulties in explaining the data.

Figure 9 plots the time series data for d(LOANS&SEC)/GDP and (VAL+DFS)/GDP. For Europe this allows us to change the issuer from other residents, which includes securitization vehicles, to nonfinancial business while changing the securities category from all securities to securities other than shares. We observe that the removal of the issuer bias implies that total securities-related effects are now far smaller than before and, similar to the US, fluctuate around an average of just below zero (see Table 1), even for Germany. We cannot know how the addition of shares would change the results for Europe, but the US results, which are almost identical for corporate bonds and the sum of corporate bonds and shares, allow us to conjecture that it would not change them significantly.

[33] The consistent negative sign is not surprising because, while market valuation risks for securities are two-sided, valuation risks for loans only exist on the downside.

Another securities-related phenomenon concerns net sales of new securities by nonfinancial business, which are sometimes interpreted as a substitute for bank financing that can explain a simultaneous net reduction in the size of financial sector balance sheets.[34] We argue that this explanation, which would imply a negative correlation between changes in the stock of credit to nonfinancial business and changes in the stocks of bonds and equity issued by nonfinancial business, can be misleading. It can be shown that for the period 1990Q1 - 2016Q4 the correlation between changes in credit and bonds is clearly positive (0.31), and even higher (0.52) during the period following the GFC. By contrast, the correlation between credit and equity is highly negative (-0.67) over the full sample, and this dominates the correlation between credit and the sum of bonds and equity (-0.48). This mostly represents share buybacks. But correlation is silent on magnitudes. Figure 10 shows the relevant detail, and illustrates that the size of share buybacks has at most times been far smaller than fluctuations in credit.[35] Furthermore, net issuance of securities does not mechanically lead to a net decline in bank credit, because the direct effect of securities issuance is only to transfer existing deposits from purchasers to issuers. The issuers then have two options, they can use the deposits to repay loans or they can spend the deposits. In the latter case the result is increased circulation of existing bank deposits, which has no effect on aggregate financial sector balance sheets. The data in Figure 10 suggest that (net) securities issuance was often not of the right sign or not nearly large enough to account for the large observed changes in credit even if it had been used to repay loans. The main exception is the significantly debt-financed share buybacks in the immediate pre-GFC years.

The GFC illustrates some of the above points especially well. Well over 50% of the sharp decline in US financial sector balance sheets was due to lower credit to households and to non-financial non-corporate business, who do not have a significant option of substituting towards direct financing. And for the non-financial corporate business sector, as shown in the working paper version of this paper, the decline in bank financing far exceeded any increase in securities financing. The dominant feature is therefore the large decline in bank financing, not substitution towards non-bank financing. Relatedly, Chrétien and Lyonnet (2016) find that the large sales of securities by US shadow banks following the GFC were not made to nonbanks but to commercial banks. The latter simultaneously cut back on lending to SME enterprises, as shown also by Abbassi et al. (2016), while their overall balance sheet remained broadly unchanged. It is therefore these cutbacks in credit that are the main reason for the contraction in the consolidated balance sheet of the financial sector, rather than sales of securities by the financial sector to the non-financial sector.

5.2. Procyclical Bank Leverage

Nuño and Thomas (2017) study the comovement between the cyclical components of U.S. bank leverage and GDP. In Table 2 we follow Nuño and Thomas (2017). We also extend their analysis, by taking account of lags of output[36] and by studying three additional economies. We define the leverage ratio as the ratio of total financial sector book value assets to book value equity. In

[34] For example, Adrian et al. (2013) make this argument for the US at the onset of the GFC.

[35] The sample period for Figure 10 ends in 2010Q4, the end of the sample period studied by Bhatia and Bayoumi (2012), whose data we use for a key subplot. Both nonfinancial business' liabilities to households and net equity purchases by households include direct as well as indirect purchases, the latter via closed-end funds, ETFs, mutual funds, life insurers and pension funds. The most significant securities-holding sector omitted from Figure 10 is the rest of the world.

[36] The rationale for including lags is the work of Ivashina and Scharfstein (2010), who show that at the outset of the Great Recession the leverage of U.S. commercial banks increased for approximately one year before deleveraging started. Reasons include the drawdown of precommitted credit lines and the time lags involved in renegotiations of existing credit lines.

doing so we follow Adrian and Shin (2014), and for identical reasons.[37] For the US, Table 2 shows that leverage in both the regulated and shadow financial sectors (excluding GSEs) is procyclical and statistically significant at lags up to three or four quarters. For EUR and FRA leverage is strongly and significantly procyclical at similar lag lengths. GER exhibits a similar pattern, but here procyclicality is weaker and not always statistically significant. As a robustness check for eurozone data, we have repeated the same analysis using leverage data that filter out the sizeable interbank lending component of financial sector assets. Results are qualitatively similar.

Adrian et al. (2013) note that banking models generally do not generate credit-induced procyclical bank leverage. The reason can be seen in our own ILF model, where the main impact effect of adverse financial shocks is reduced bank equity, while loans and deposits, which represent slow-moving physical savings, initially change very little, thereby implying countercyclical leverage. On the other hand, in FMC models changes in bank loans and deposits are much larger relative to changes in bank equity. As a result, leverage is generally procyclical. Furthermore, this channel does not require VAL or DFS effects, which mostly rely on the behavior of securities portfolios rather than of credit, and which as we have seen are quantitatively of little importance.

5.3. Quantity Rationing Versus Price Rationing of Credit

A substantial empirical literature demonstrates that a major part of banks' response to turning points in the credit cycle consists of adjustments in the quantity of loans rather than in the price of loans. Following Waters (2013a), we refer to these quantity effects as quantity rationing of credit, but we note that this term has a different meaning from the majority of the theoretical literature, which is based on the ILF model of banking.[38] Our argument is not that banks' willingness to lend out loanable funds may be limited by informational asymmetries, or that non-price lending terms may affect the equilibrium quantity of lending independently of interest rates. Rather, we argue that banks' willingness to create new funds through lending always changes along with the price terms that they are prepared to offer to their borrowers.

For the recent period, the empirical evidence of Lown and Morgan (2006), Thies and Gouldey (2010), Waters (2013a,b) and Gilchrist and Zakrajšek (2012) shows that the quantity of credit is a more important driver of real activity than the price of credit.[39] Lown and Morgan (2006) use the Federal Reserve's Survey of Loan Officers Opinion Survey (SLOOS) series to show that lending standards are significantly correlated with the realized quantity of aggregate lending, that they identify fluctuations in credit supply rather than in credit demand, and that they affect real output. Waters (2013a,b) confirms that quantity rationing, as measured by SLOOS, is a significant leading indicator for output. Lown and Morgan (2006) and Waters (2013a,b) find that borrowing costs are a less powerful leading indicator. Gilchrist and Zakrajšek (2012) construct a U.S. corporate credit spread and decompose it into a price rationing component that reflects movements in expected default, and a quantity rationing component, referred to as the excess bond premium (EBP), which

[37] "....Our concern is with the availability of credit through the intermediary, and hence with the lending decisions of the bank. Thus, the appropriate balance sheet concept is that of total assets, rather than enterprise value, since total assets address the issue of how much the bank lends. The corresponding equity concept is book equity, and the appropriate concept of leverage is the ratio of total assets to book equity."

[38] For early theoretical contributions on quantity rationing, see Jaffee and Russell (1976), Stiglitz and Weiss (1981), Blinder and Stiglitz (1983) and Fuerst (1992). See Bellier et al. (2012) for a literature survey.

[39] Recent empirical studies using loan-level or bank-level data, and studies of the Great Depression, have also found evidence for the importance of quantity rationing of credit. For the former see Ivashina and Scharfstein (2010), Khwaja and Mian (2008), Kapan and Minoiu (2013) and Berger et al. (2018), and for the latter see Bernanke (1983), Hamilton (1987) and Baum and Thies (1989).

they interpret as reflecting shifts in the effective supply of funds offered by financial intermediaries. They find that the EBP and the SLOOS series, both indicators of quantity rationing, are highly correlated, and that all of the forecasting ability of their spread for the 1985-2010 period is due to the EBP, in other words to changes in quantity rationing of credit.

Our model simulation in Figure 2 has shown that in ILF models banks' reaction to adverse financial shocks consists almost entirely of changes in lending spreads, while changes in credit are slow and small. By contrast, in FMC models immediate, large and persistent cutbacks in credit also play a very important role.

6. Conclusions

The integration of banks into macroeconomic models continues to be a high priority, both for academic research and for policymakers, who depend on the insights generated by such models for the design and calibration of their policies. The currently dominant modeling framework is what we have labelled the intermediation of loanable funds (ILF) model. We have shown that the ILF representation of the accumulation of bank deposits does not correspond to real-world deposits of private financial instruments (or of central bank money) at banks. The reason is that any financial instrument only has value because the deposit already exists at another bank. We have also shown, using budget constraints, that the ILF model instead represents banks as resource traders that accept deposits of physical resources from savers before lending them to borrowers. But as a modeling shortcut this is unrealistic and has unrealistic implications. Instead, as explained by many central bank publications, banks are financial institutions that create deposits as ledger entries to fund their loans (or, but this is quantitatively far less important, to purchase securities), in transactions that involve a single borrower-cum-depositor and no saver. We have proposed a financing through money creation (FMC) model that reflects this. The FMC model recognizes that neither the quantity of physical savings (as in the ILF model) nor the quantity of central bank money (as in the deposit multiplier model) imposes quantitative constraints on banks' ability to create deposit money. The main constraint is banks' own and their customers' expectations concerning the profitability of additional loans. If these expectations are volatile, then financial sector balance sheets must also be volatile.

We compare the simulation properties of our proposed FMC model to those of an ILF model that, with a single exception, is identical in its structural equations, calibration of structural parameters, optimality conditions and deterministic steady state. The exception is that the budget constraints of the household saver and entrepreneur borrower of the ILF model are merged into the single budget constraint of the household-entrepreneur of the FMC model, and entrepreneur net worth is dropped as a variable. With this single change, following identical financial and real shocks, the FMC model predicts changes in bank loans and deposits that are significantly larger, happen much faster, and that have larger effects on the real economy, while the adjustment process is less dependent on changes in lending spreads.

In the data, changes in the size of aggregate financial sector balance sheets are highly volatile. This is a natural outcome of the money creation or financing mechanism of the FMC model, but not of the physical saving mechanism of the ILF model, because curvature in preferences implies that the accumulation of physical savings should be smooth and gradual. The latter is confirmed by data for net private saving, which is far smoother, and in many economies of completely different average size, than changes in aggregate financial sector balance sheets. More fundamentally, an emphasis on physical saving in explaining changes in financial sector balance sheets forces us to accept the notion that banks are resource traders, a notion that should be ruled out on a priori grounds. Augmenting the ILF model by also allowing for net purchases by banks from nonbanks of physical capital or of securities that represent claims to physical capital, and that are financed by the creation of bank deposits, does little to address this problem. The reason is that while this mechanism does exist and therefore does not need to be ruled out on a priori grounds, its quantitative significance is negligible. The same is true for asset valuation effects, whose principal component is changes in the market value of such securities. This leaves the money creation mechanism of the FMC model to account for virtually the entirety of observed changes in financial sector balance sheets. We show that the FMC model is also able to qualitatively account for two other stylized facts, procyclical financial sector leverage and quantity rationing of credit during downturns.

Our results could potentially have important policy implications for central banks, because they suggest that a quantitative investigation of the effects of monetary and macroprudential policies that is based on FMC models, whose logic is advocated by many central banks themselves, is likely to yield results that differ significantly from those of the existing literature. We are confident that this will generate a very useful research agenda.

References

Abbassi, P., Iyer, R., Peydro, J.-L. and Tous, F. (2016), "Securities Trading by Banks and Credit Supply: Micro-Evidence from the Crisis", *Journal of Financial Economics*, **121**, 569-594.

Adrian, T. and Boyarchenko, N. (2012), "Intermediary Leverage Cycles and Financial Stability", Federal Reserve Bank of New York Staff Reports, No. 567.

Adrian, T., Colla, P. and Shin, H.S. (2013), "Which Financial Frictions? Parsing the Evidence from the Financial Crisis of 2007 to 2009", *NBER Macroeconomics Annual*, **27**, 159-214.

Adrian, T. and Shin, H.S. (2014), "Procyclical Leverage and Value-at-Risk", *Review of Financial Studies*, **27(2)**, 373–403.

Anderson, R.G., Duca, J.V. and Jones, B.E. (2018), "Financial Innovation, Shadow Banking, Risk Premia and Divisia Velocity", Working Paper, Federal Reserve Bank of Dallas and Binghampton University.

Ashcraft, A.B. and Steindel, C. (2008), "Measuring the Impact of Securitization on Imputed Bank Output", Working Paper, Federal Reserve Bank of New York.

Ball, L. (2001), "Another Look at Long-Run Money Demand", *Journal of Monetary Economcics*, **47**, 31-44.

Barnett, W., Offenbacher E. and Spindt, P. (1984), "The New Divisia Aggregates", *Journal of Political Economy*, **92(6)**, 1049-1085.

Baum, C. and Thies, C. (1989), "The Term Structure of Interest Rates and the Demand for Money During the Great Depression", *Southern Economic Journal*, **56(2)**, 490-498.

Bellier, A., Sayeh, W. and Serve, S. (2012), "What Lies Behind Credit Rationing? A Survey of the Literature", Université de Cergy Pontoise, Thema Working Papers, No. 2012-39.

Benes, J. and Kumhof, M. (2012), "The Chicago Plan Revisited", IMF Working Paper WP/12/202.

Benes, J. and Kumhof, M. (2015), “Risky Bank Lending and Countercyclical Capital Buffers”, *Journal of Economic Dynamics and Control*, **58**, 58-80.

Benes, J., Kumhof, M. and Laxton, D. (2014a), “Financial Crises in DSGE Models : A Prototype Model”, IMF Working Papers, WP/14/57.

Benes, J., Kumhof, M. and Laxton, D. (2014b), “Financial Crises in DSGE Models : Selected Applications of MAPMOD”, IMF Working Papers, WP/14/56.

Berger, A., Makaew, T. and Turk-Ariss, R. (2018), “Who Pays for Financial Crises? Price and Quantity Rationing of Different Borrowers by Domestic and Foreign Banks”, IMF Working Papers, WP/18/158.

Bernanke, B. (1983), “Nonmonetary Effects of the Financial Crisis in the Propagation of the Great Depression”, *American Economic Review*, **73(3)**, 257-276.

Bernanke, B., Gertler, M. and Gilchrist, S. (1999), “The Financial Accelerator in a Quantitative Business Cycle Framework”, in: Taylor, J.B., Woodford, M. (eds.), *Handbook of Macroeconomics, Volume 1C*. Amsterdam: Elsevier, pp. 1341-1393.

Bezemer, D., Samarina, A. and Zhang, L. (2017), “The Shift in Bank Credit Allocation: New Data and New Findings”, DNB Working Papers, No. 559.

Bhatia, A. and Bayoumi, T. (2012), “Leverage? What Leverage? A Deep Dive into the U.S. Flow of Funds in Search of Clues to the Global Crisis”, IMF Working Papers, WP/12/162.

Bianchi, J. and Bigio, S. (2014), “Banks, Liquidity Management and Monetary Policy”, NBER Working Papers, No. 20490.

Bigio, S. and Weill, P.-O. (2016), “A Theory of Bank Balance Sheets”, Working Paper.

Blinder, A. and Stiglitz, J. (1983), “Money, Credit Constraints and Economic Activity”, *American Economic Review, Papers and Proceedings*, **73**, 297-302.

Boissay, F., Collard, F. and Smets, F. (2016), “Booms and Banking Crises”, *Journal of Political Economy*, **124(2)**, 489-538.

Borio, C. and Disyatat, P. (2011), “Global Imbalances and the Financial Crisis: Link or No Link?”, BIS Working Papers, No. 346.

Borio, C. and Disyatat, P. (2015), “Capital Flows and The Current Account: Taking Financing (More) Seriously”, BIS Working Papers, No. 525.

Brock, W.A. (1975), “A Simple Perfect Foresight Monetary Model”, *Journal of Monetary Economics*, **1**, 133-150.

Brunnermeier, M. and Sannikov, Y. (2016), “The I-Theory of Money”, Working Paper, Princeton University.

Bundesbank (2017), “The Role of Banks, Non-Banks and the Central Bank in the Money Creation Process”, Monthly Report, April 2017.

Calvo, G.A. (1983), “Staggered Prices in a Utility-Maximizing Framework”, *Journal of Monetary Economics*, **12**, 383-398.

Chari, V.V. and Phelan, C. (2014), “On the Social Usefulness of Fractional Reserve Banking”, *Journal of Monetary Economics*, **65**, 1-13.

Chrétien, E. and Lyonnet, V. (2016), “Traditional and Shadow Banks During the Crisis”, ACPR and Banque de France, Débats Économiques et Financiers, No. 27.

Christiano, L. J., Eichenbaum, M. and Evans, C. L. (2005), “Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy,” *Journal of Political Economy*, **113(1)**, 1–45.

Christiano, L., Motto, R. and Rostagno, M. (2014), “Risk Shocks”, *American Economic Review*, **104(1)**, 27-65.

Clancy, D. and Merola, R. (2017), "Countercyclical Capital Rules for Small Open Economies", *Journal of Macroeconomics*, **54**, 332-351.

Clerc, L., Derviz, A., Mendicino, C., Moyen, S., Nikolov, K., Stracca, L., Suarez, J. and Vardoulakis, A. (2015), "Capital Regulation in a Macroeconomic Model with Three Layers of Default", *International Journal of Central Banking*, 9-63.

Culbertson, J. (1958), "Intermediaries and Monetary Theory: A Criticism of the Gurley-Shaw Theory", *American Economic Review*, **48(1)**, 119-131.

Curdia, V. and Woodford, M. (2010), "Credit Spreads and Monetary Policy", *Journal of Money, Credit and Banking*, **24(6)**, 3–35.

De Fiore, F., Teles, P. and Tristani, O. (2011), "Monetary Policy and the Financing of Firms", *American Economic Journal: Macroeconomics*, **3(4)**, 1112-1142.

Decker, F. and Goodhart, C. (2018), "Credit Mechanics - A Precursor to the Current Money Supply Debate", CEPR Discussion Papers, No. 13233.

Deleidi, M. and Levrero, E. (2019), "The Money Creation Process: A Theoretical and Empirical Analysis for the United States", *Metronomica*, available at https://doi.org/10.1111/meca.12238.

Desan, C. (2014), *Making Money: Coin, Currency and the Coming of Capitalism*, Oxford: Oxford University Press.

Doherty, E., Jackman, B. and Perry, E. (2018), "Money in the Australian Economy", Reserve Bank of Australia Bulletin, September.

Donaldson, J., Piacentino, G. and Thakor, A. (2018), "Warehouse Banking", *Journal of Financial Economics*, **129(2),** 250-267.

Eggertsson, G., Juelsrud, R. E. and Wold, E. G. (2017), "Are Negative Nominal Interest Rates Expansionary?", NBER Working Papers, No. 24039.

Faure, S. and Gersbach, H. (2017), "Loanable Funds versus Money Creation in Banking: A Benchmark Result", CFS Working Paper Series, No. 587, Frankfurt: Goethe University.

Friedman, M. (1971), "The Euro-Dollar Market: Some First Principles", Federal Reserve Bank of St. Louis Review, 16-24.

Fuerst, T. (1992), "Liquidity, Loanable Funds and Real Activity", *Journal of Monetary Economics*, **29**, 3-24.

Gerali, A., Neri, S., Sessa, L. and Signoretti, F. (2010), "Credit and Banking in a DSGE Model of the Euro Area", *Journal of Money, Credit, and Banking*, **42**, 107-141.

Gertler, M. and Karadi, P. (2011), "A Model of Unconventional Monetary Policy", *Journal of Monetary Economics*, **58(1)**, 17-34.

Gertler, M. and Kiyotaki, N. (2011), "Financial Intermediation and Credit Policy in Business Cycle Analysis", in: Friedman, B. and Woodford, M. (eds.), *Handbook of Monetary Economics, Volume 3A*, North Holland: Elsevier.

Gertler, M. and Kiyotaki, N. (2015), "Banking, Liquidity, and Bank Runs in an Infinite Horizon Economy", *American Economic Review*, **105(7)**, 2011-2043.

Gertler, M., Kiyotaki, N. and Prestipino, A. (2015), "Wholesale Banking and Bank Runs in Macroeconomic Modelling of Financial Crises", in: J. Taylor and H. Uhlig, eds., *Handbook of Macroeconomics*.

Gestrich, H. (1936), *Neue Kreditpolitik*, W. Kohlhammer: Stuttgart and Berlin.

Gilchrist, S. and Zakrajšek, E. (2012), "Credit Spreads and Business Cycle Fluctuations", *American Economic Review*, **102(4)**, 1692-1720.

Goodfriend, M. and McCallum, B. (2007), "Banking and Interest Rates in Monetary Policy Analysis: A Quantitative Exploration", *Journal of Monetary Economics*, **54**, 1480-1507.

Goodhart, C. (1988), *The Evolution of Central Banks*, Cambridge, Massachusetts: MIT Press.

Graeber, D. (2011), *Debt: The First 5,000 Years*, Brooklyn, NY: Melville House Printing.

Gu, C., Mattesini, F., Monnet, C. and Wright, R. (2013), "Banking: A New Monetarist Approach", *Review of Economic Studies*, **80**, 636-662.

Gurley, J. and Shaw, E. (1955), "Financial Aspects of Economic Development", *American Economic Review*, **45**, 515-538.

Gurley, J. and Shaw, E. (1956), "Financial Intermediaries and the Saving-Investment Process", *Journal of Finance,* **11**, 257-276.

Hahn, L. (1920), *Volkswirtschaftliche Theorie des Bankkredits*, Tübingen: J.C.B. Mohr.

Hamilton, J. (1987), "Monetary Factors in the Great Depression", *Journal of Monetary Economics*, **19(2)**, 147-170.

Hockett, R. and Omarova, S. (2017), "The Finance Franchise", *Cornell Law Review*, **102(5)**, 1143-1218.

Hudson, M. (2017), "Origins of Money and Interest: Palatial Credit, Not Barter", in Y. Cassis, K. Yago and S. Battilossi, eds., *Handbook of the History of Money and Currency*, Springer Verlag (forthcoming).

Humphrey, C. (1985), "Barter and Economic Disintegration", *Man*, **20**, 48-72.

Iacoviello, M. (2005), "House Prices, Borrowing Constraints and Monetary Policy in the Business Cycle", *American Economic Review*, **95(3)**, 739-764.

Ivashina, V. and Scharfstein, D. (2010), "Bank Lending During the Financial Crisis of 2008", *Journal of Financial Economics*, **97**, 319-338.

Jaffee, D. and Russell, T. (1976), "Imperfect Information and Credit Rationing", *Quarterly Journal of Economics*, **90**, 651-666.

Jakab, Z. and Kumhof, M. (2015), "Banks Are Not Intermediaries of Loanable Funds - And Why This Matters", Bank of England Staff Working Papers, No. 529.

Jermann, U. and Quadrini, V. (2012), "Macroeconomic Effects of Financial Shocks", *American Economic Review*, **102(1)**, 238-271.

Justiniano, A., Primiceri, G. and Tambalotti, A. (2015), "Credit Supply and the Housing Boom", Federal Reserve Bank of New York Staff Reports, No. 709.

Kaldor, N. (1989), *Further Essays on Economic Theory and Policy*, edited by F. Targetti, and A.P. Thirlwall, London: Duckworth.

Kapan, T. and Minoiu, C. (2013), "Liquidity Shocks and the Supply of Credit After the 2007-08 Crisis", Working Paper, Fannie Mae and International Monetary Fund.

Kashyap, A. and Stein, J. (1993), "Monetary Policy and Bank Lending", NBER Working Papers, No. 4317.

Kashyap, A., Stein, J. and Wilcox, D. (1993), "Monetary Policy and Credit Conditions: Evidence from the Composition of External Finance", *American Economic Review*, **83(1)**, 78-98.

Keen, S. (2014), "Endogenous Money and Effective Demand ", *Review of Keynesian Economics*, **2(3)** 271-291.

Keen, S. (2015), "The Macroeconomics of Endogenous Money: Response to Fiebiger, Palley and Lavoie ", *Review of Keynesian Economics*.

Keynes, J.M. (1939), "The Process of Capital Formation", *The Economic Journal*, **49**, 572-573.

Khwaja, A. and Mian, A. (2008), “Tracing the Impact of Bank Liquidity Shocks: Evidence from an Emerging Market”, *American Economic Review*, **98(4)**, 1413-1442.

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

Kravik, E. and Paulsen, K. (2017), “A Complete Documentation of Norges Bank’s Policy Model NEMO”, Working Paper.

Kumhof, M. and Wang, X. (2018), “Banks, Money and the Zero Lower Bound”, Bank of England Staff Working Papers, No. 752.

Kydland, F. and Prescott, E. (1990), “Business Cycles: Real Facts and a Monetary Myth”, *Federal Reserve Bank of Minneapolis Quarterly Review*, **14(2)**, 3-18.

Lagos, R., Rocheteau, G. and Wright, R. (2017), “Liquidity: A New Monetarist Perspective”, *Journal of Economic Literature*, **55(2)**, 371-440.

Lautenbach, W. (1952), *Zins, Kredit und Produktion*, Stützel, W. (ed.), Tübingen: J.C.B. Mohr.

Lavoie, M. (2014), “A Comment on Endogenous Money and Effective Demand: a Revolution or a Step backwards?”, *Review of Keynesian Economics*, **2(3)**, 321-332.

Lindner, F. (2012), “Saving Does Not Finance Investment: Accounting as an Indispensable Guide to Economic Theory”, IMK Working Papers, No. 100.

Lindner, F. (2013), “Does Saving Increase the Supply of Credit? A Critique of Loanable Funds Theory”, IMK Working Papers, No. 120.

Lown, C. and Morgan, D. (2006), “The Credit Cycle and the Business Cycle: New Findings Using the Loan Officer Opinion Survey”, *Journal of Money, Credit and Banking*, **38(6)**, 1575-1597.

Lozej, M., Onorante, L. and Rannenberg, A. (2017), “Countercyclical Capital Regulation in a Small Open Economy DSGE Model”, Central Bank of Ireland, Research Technical Papers, 03/RT/2017.

Lozej, M. and Rannenberg, A. (2017), “The Macroeconomic Effects of the Regulatory LTV and LTI Ratios in the Central Bank of Ireland’s DSGE Model’, Central Bank of Ireland, Economic Letter Series, Vol. 2017, No. 4.

Lucas, R. (1990), “Liquidity and Interest Rates”, *Journal of Economic Theory*, **50**, 237-264.

McLeay, M., Radia, A. and Thomas, R. (2014a), “Money in the Modern Economy: An Introduction”, *Bank of England Quarterly Bulletin*, **54(1)**, 4-13.

McLeay, M., Radia, A. and Thomas, R. (2014b), “Money Creation in the Modern Economy”, *Bank of England Quarterly Bulletin*, **54(1)**, 14-27.

Menger, K. (1892), “On the Origin of Money”, *Economic Journal*, **2**, 238–255.

Mian, A. and Sufi, A. (2014), *House of Debt*, Chicago: University of Chicago Press.

Minsky, H.P.(1977), “The Financial Instability Hypothesis: an Interpretation of Keynes and an Alternative to Standard Theory”, *Nebraska Journal of Economics and Business*

Moore, B.J. (1979), “The Endogenous Money Stock ”, *Journal of Post Keynesian Economics* , 2(1), 49-70.

Nelson, B., Pinter, G. and Theodoridis, K. (2015), “Do Contractionary Monetary Policy Shocks Expand Shadow Banking?”, Bank of England Working Papers, No. 521.

Nuño, G. and Thomas, C. (2017), “Bank Leverage Cycles”, *American Economic Journal: Macroeconomics*, **9(2)**, 32-72.

Piazzesi, M. and Schneider, M. (2018), “Payments, Credit and Asset Prices”, BIS Working Papers, No. 737.

Pozsar, Z. (2014), "Shadow Banking: The Money View", OFR Working Papers.

Ramanauskas, T. and Karmelavičius, J. (2018), "Bank Credit and Money Creation in a DSGE Model of a Small Open Economy", Lietuvos Bankas, Working Paper Series, No. 52/2018.

Ricks, M. (2018), "Money as Infrastructure", *Columbia Business Law Review* (forthcoming).

Rittershausen, H. (1956), *Bankpolitik*, Fritz Knapp Verlag: Frankfurt am Main.

Schmitt-Grohé, S. and Uribe, M. (2004), "Optimal Fiscal and Monetary Policy under Sticky Prices", *Journal of Economic Theory*, **114**, 198-230.

Schumpeter, J.A. (1934), *The Theory of Economic Development: An Inquiry into Profits, Capital, Credits, Interest, and the Business Cycle*, Piscataway: Transaction Publishers.

Schumpeter, J.A. (1954), *History of Economic Analysis*, New York: Oxford University Press.

Sidrauski, M. (1967), "Rational Choice and Patterns of Growth in a Monetary Economy", *American Economic Review Papers and Proceedings*, **47**, 534-544.

Smith, W. (1959), "Financial Intermediaries and Monetary Controls", *Quarterly Journal of Economics*, **73,** 533-553.

Stiglitz, J. and Weiss, A. (1981), "Credit Rationing in Markets with Imperfect Information", *American Economic Review*, **71(2)**, 393-410.

Stützel, W. (1958), *Volkswirtschaftliche Saldenmechanik*, second edition, Mohr Siebeck: Tübingen.

Sun, G. and He, X. (2018), "Zero Lower Bound on Deposit Rates and Transmission Mechanism of the Negative Interest Rate Policy", The People's Bank of China Working Paper Series, No. 2018/2.

Thies, C. and Gouldey, B. (2010), "Direct Measures of Credit Rationing, 2002-2010", *Journal of the Northeastern Association of Business, Economics and Technology*, 45-51.

Tobin, J. (1963), "Commercial Banks as Creators of Money", Cowles Foundation Discussion Papers, No. 159.

Ueda, K. and Brooks, R. (2011), "User Manual for the Corporate Vulnerability Utility: The 4th Edition", International Monetary Fund.

Waters, G. (2013a), "Quantity Rationing of Credit and the Phillips Curve", *Journal of Macroeconomics*, **37**, 68-80.

Waters, G. (2013b), "Quantity versus Price Rationing of Credit: An Empirical Test", *International Journal of Financial Studies*, **1**, 45-53.

Werner, R. (2014a), "Can Banks Individually Create Money Out of Nothing? The Theories and the Empirical Evidence", *International Review of Financial Analysis*, **36**, 1-19.

Werner, R. (2014b), "How Do Banks Create Money, and Why Can Other Firms Not Do the Same? An Explanation for the Coexistence of Lending and Deposit-taking", *International Review of Financial Analysis*, **36**, 71–77.

Woodford, M. (2003), *Interest and Prices*, Princeton and Oxford: Princeton University Press.

Zarlenga, S. (2002), *The Lost Science of Money*, American Monetary Institute Charitable Trust.

Table 1: Sample Averages

		Averages (% of GDP)			
	Sample	*d(LOANS&SEC)*	*VAL*[1]	*DFS*[1]	*VAL*[2]+*DFS*[2]
Measurement		ΔST	ΔST-FL	FL	ΔST
Catogories		Equities+Bonds	Equities+Bonds	Equities+Bonds	Bonds
Issuers		NFB + IS	NFB + IS	NFB + IS	NFB
US	90Q1 - 16Q4	7.1%	-0.3%	0.3%	0.0%
EUR	97Q4 - 16Q4	4.5%	-0.2%	0.9%	-0.2%
GER	03Q2 - 16Q4	0.9%	-0.6%	0.4%	-0.4%
FRA	03Q2 - 16Q4	4.7%	-0.1%	0.7%	-0.4%

(Data: Flow of funds. Quarterly.)

(ΔST = change in stocks, FL = flows, NFB = nonfinancial business, IS = insurers and securitization vehicles)

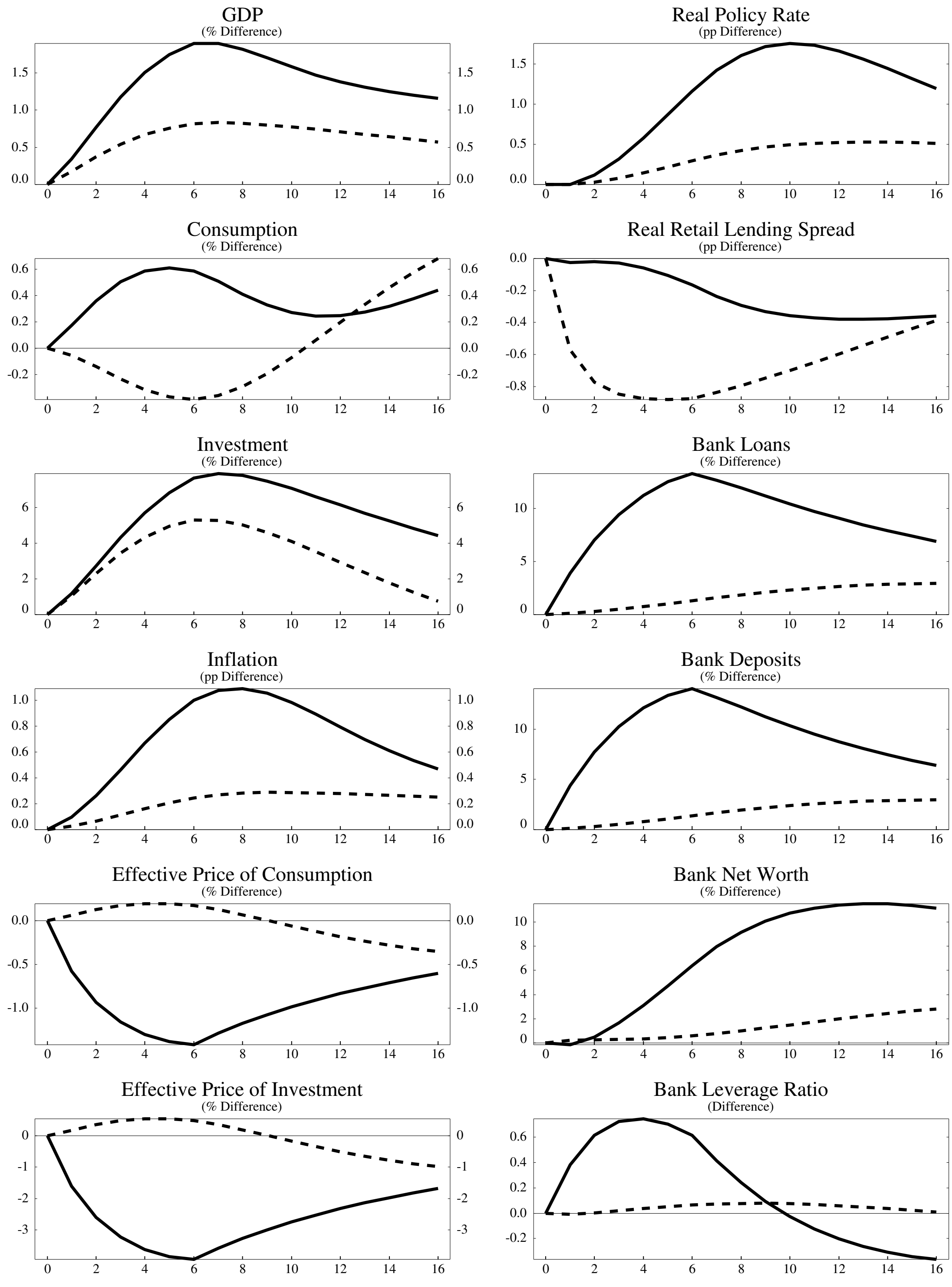
Table 2: Correlation of Financial Sector Leverage and GDP in Four Economies

Cross-correlation between cyclical components of logarithm of lagged GDP and leverage ratio (with assets = cumulated flows)

	US Regulated	**US Shadow**	**US Regulated + Shadow**	**EUR**	**GER**	**FRA**
Lags	90:1 - 16:4	90:1 - 16:4	90:1 - 16:4	97:3 10:3	97:3 10:3	97:3 16:4
0	0.18*	0.66***	0.53***	0.40**	0.24	0.40***
1	0.20*	0.65***	0.50***	0.56***	0.33**	0.54***
2	0.19*	0.50***	0.38***	0.62***	0.33**	0.60***
3	0.16	0.29***	0.23**	0.56***	0.25	0.58***
4	0.15	0.08	0.10	0.41***	0.14	0.48***
5	0.13	-0.07	0.00	0.21	0.03	0.33***

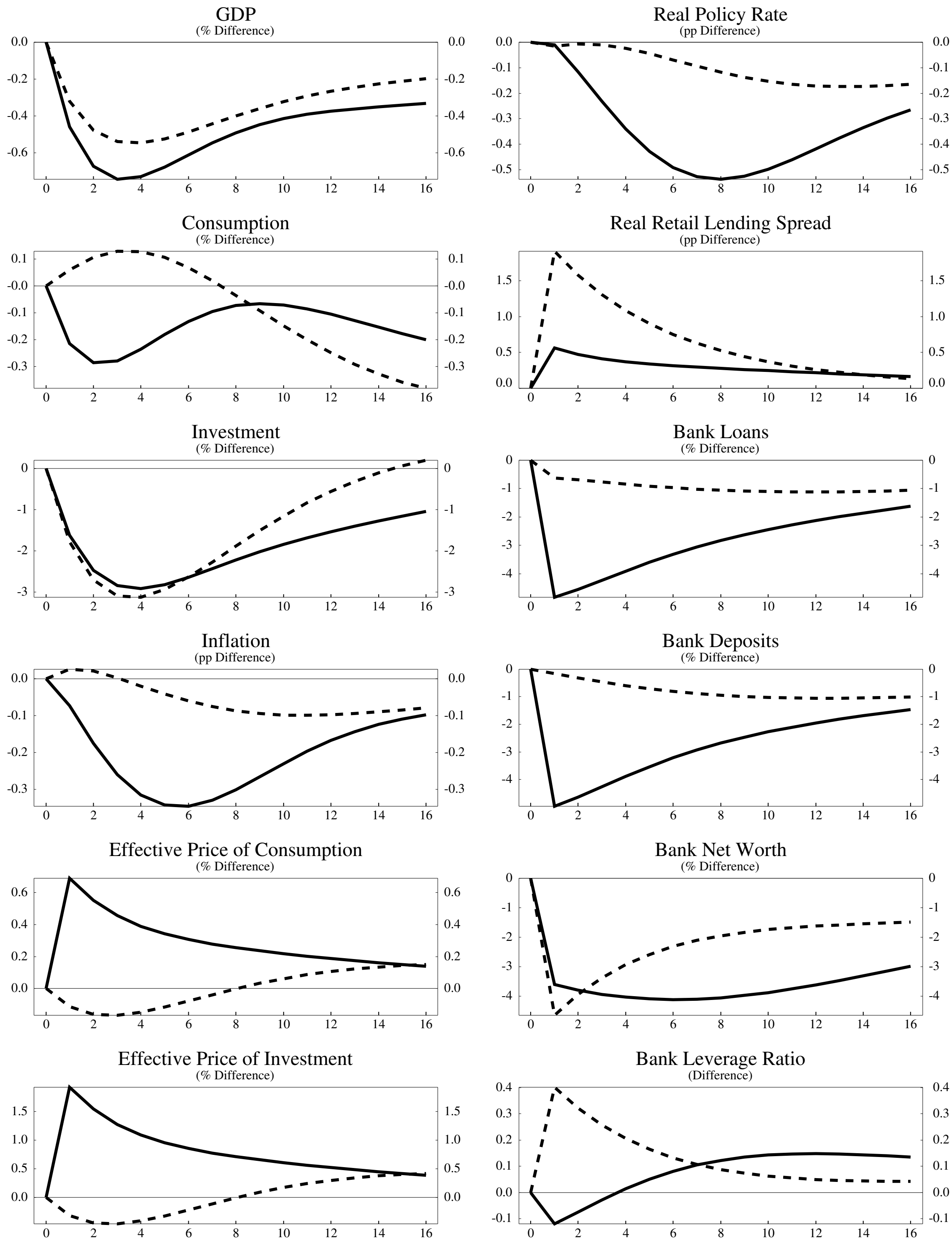
(Data: Flow of funds. Quarterly.)

(* = Significant at 10% confidence level, ** = Significant at 5% confidence level, *** = Significant at 1% confidence level)



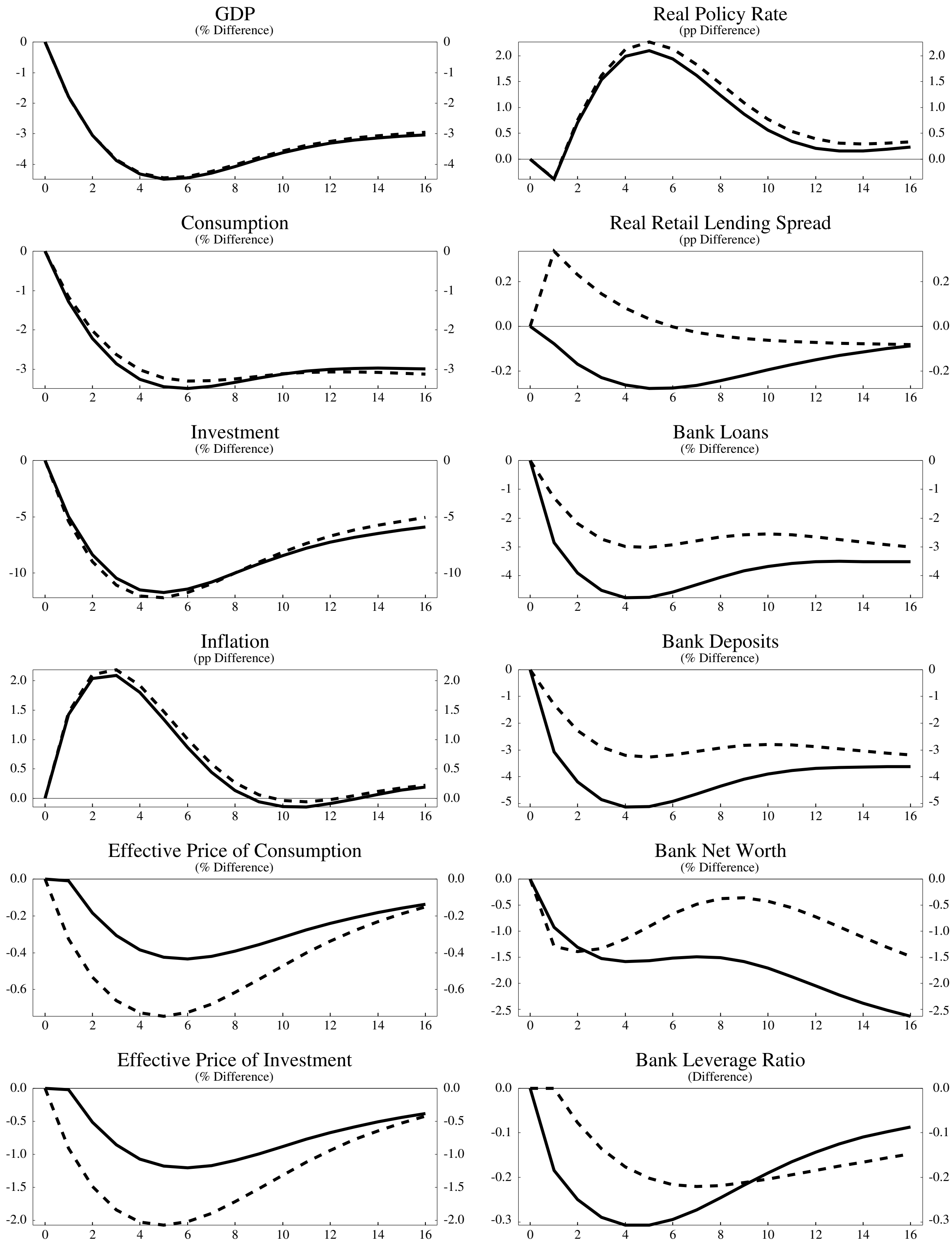
(ILF Model = dashed lines, FMC Model = solid lines)

Figure 1: IRF - Credit Boom due to Higher Willingness to Lend



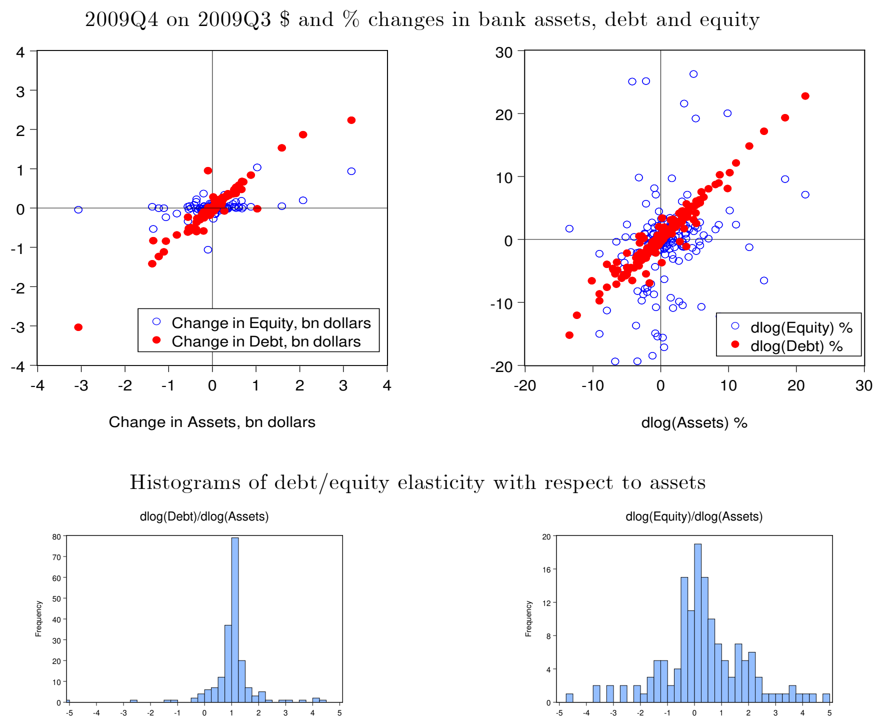
(ILF Model = dashed lines, FMC Model = solid lines)

Figure 2: IRF - Credit Crash due to Higher Borrower Riskiness



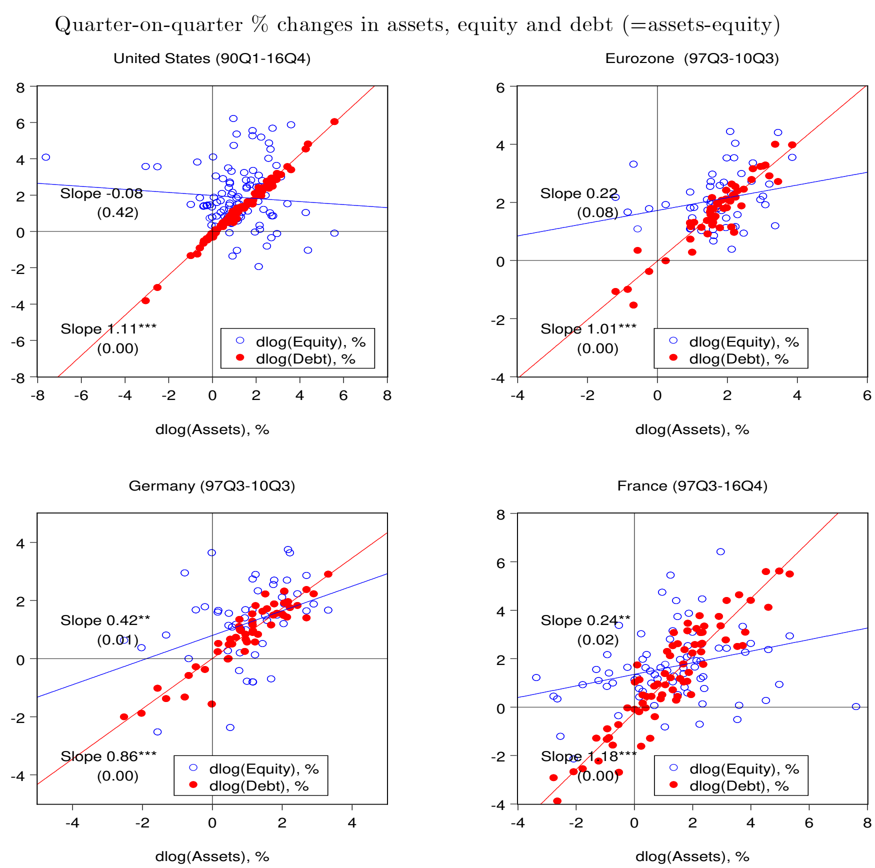
(ILF Model = dashed lines, FMC Model = solid lines)

Figure 3: IRF - Contractionary Technology Shock



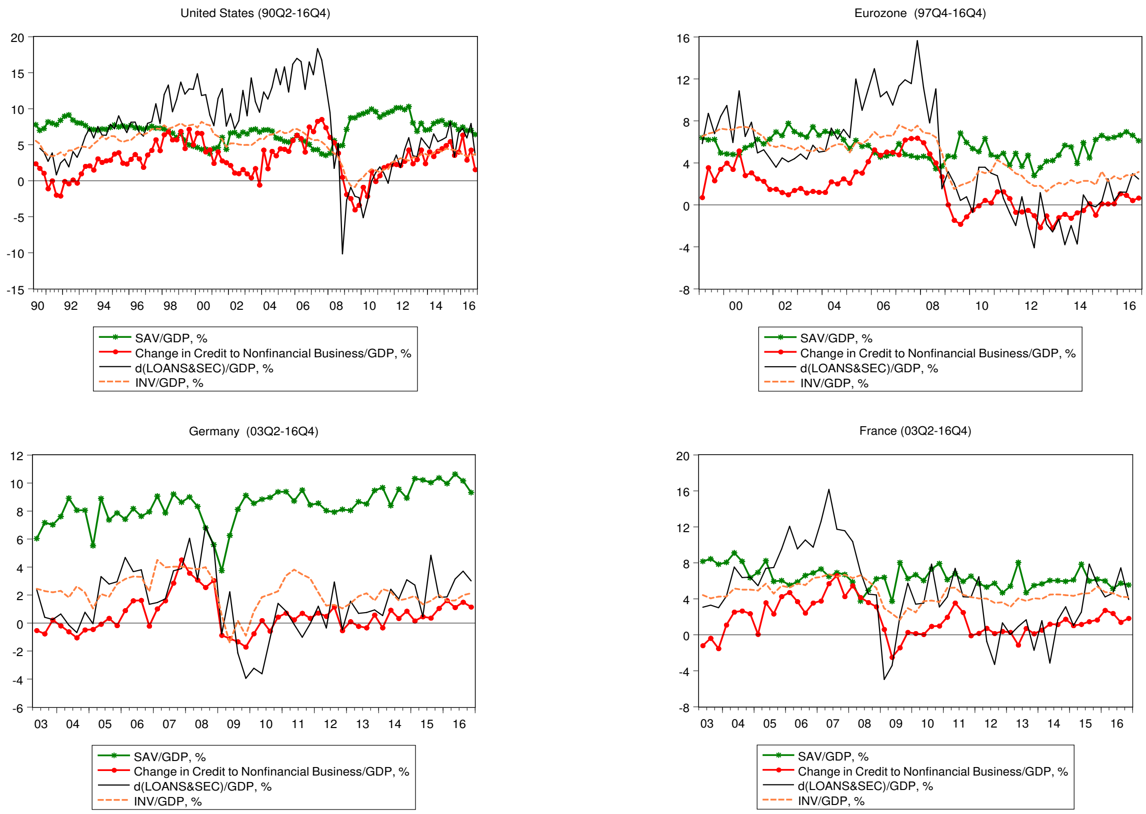
(Data: FDIC call reports. 200 largest U.S. banks by asset size. Based on flow data.)

Figure 4: Bank Balance Sheets, Cross-Sectional Evidence



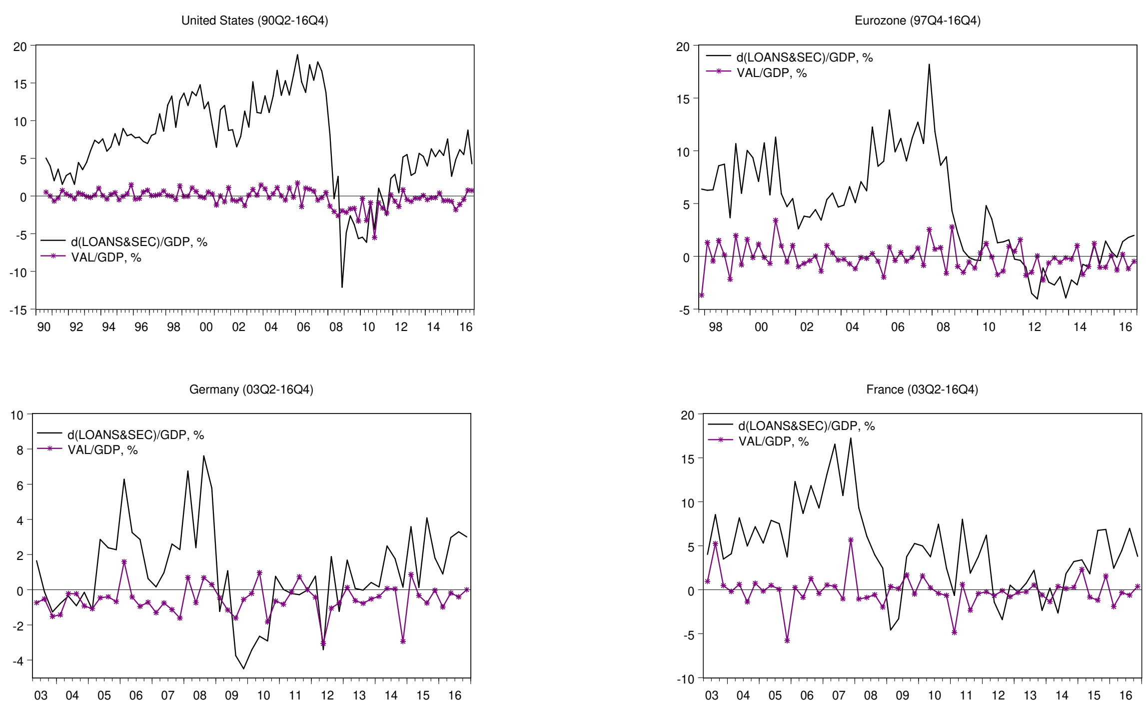
(Data: Flow of funds for aggregate banking systems. Based on flow data.)

Figure 5: Bank Balance Sheets, Time Series Evidence



(Data: Flow of funds and national accounts. Quarterly. Based on flow data for d(LOANS&SEC). All variables divided by the same quarter's GDP.)

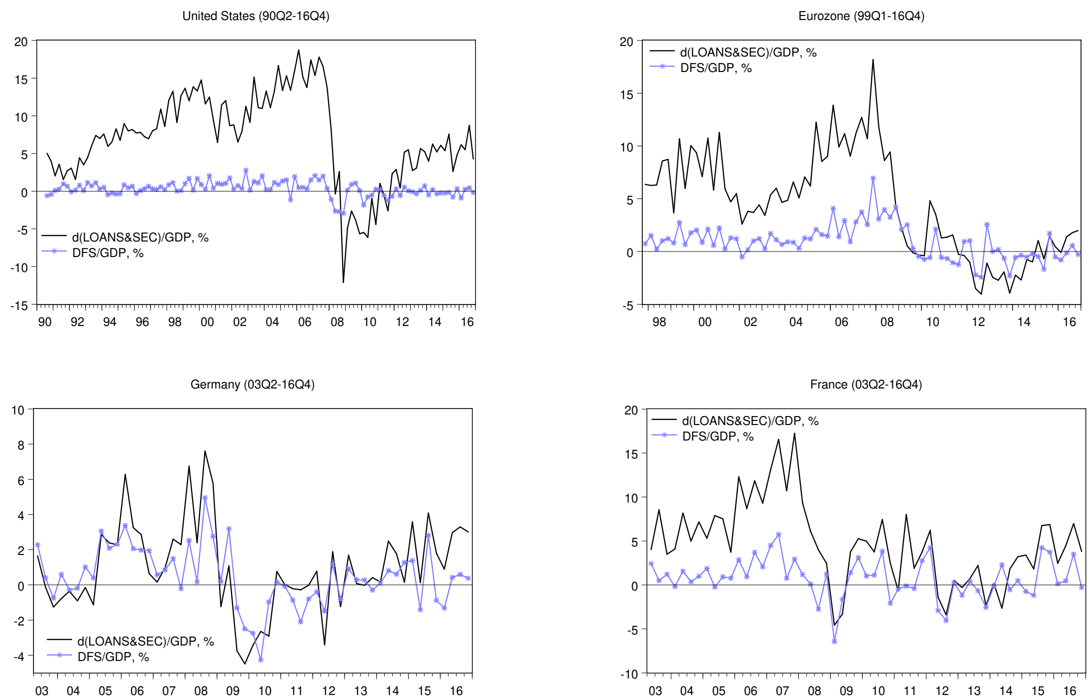
Figure 6: Physical Saving (SAV)



(Data: Flow of funds. Quarterly. Based on stock data for d(LOANS&SEC). All variables divided by the same quarter's GDP.)

(VAL/GDP for US: Securities issued by Nonfinancial Business. VAL/GDP for EUR/GER/FRA: Securities issued by Other Residents.)

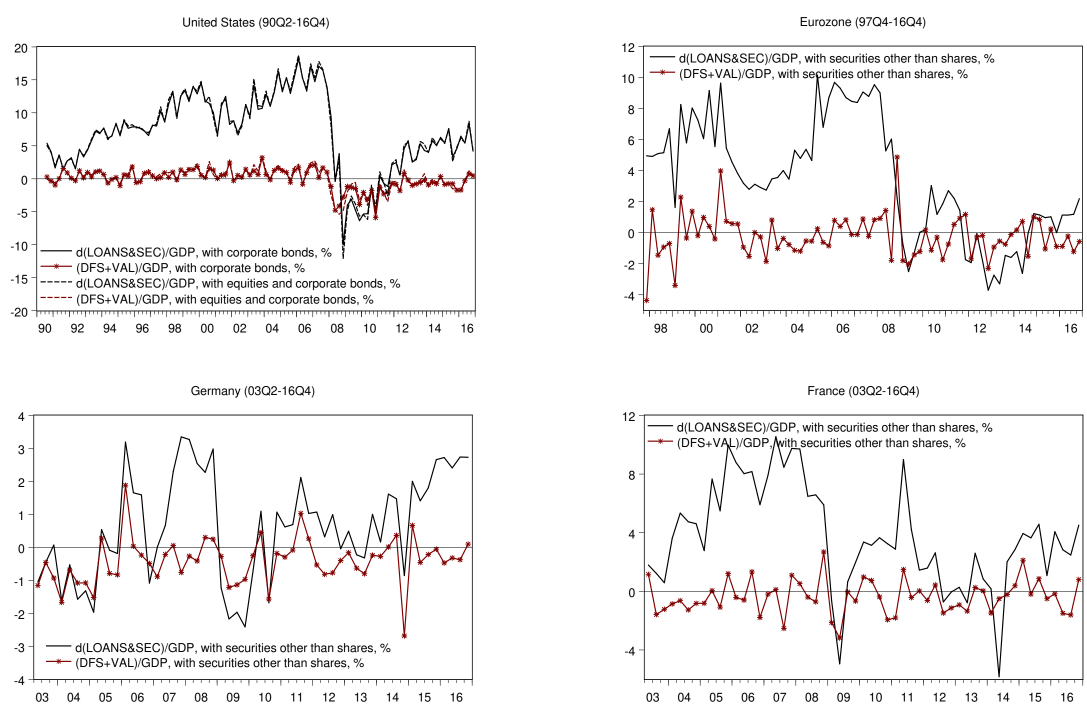
Figure 7: Valuation Effects (VAL)



(Data: Flow of funds. Quarterly. Based on stock data for d(LOANS&SEC). All variables divided by the same quarter's GDP.)

(DFS/GDP for US: Securities issued by Nonfinancial Business. DFS/GDP for EUR/GER/FRA: Securities issued by Other Residents.)

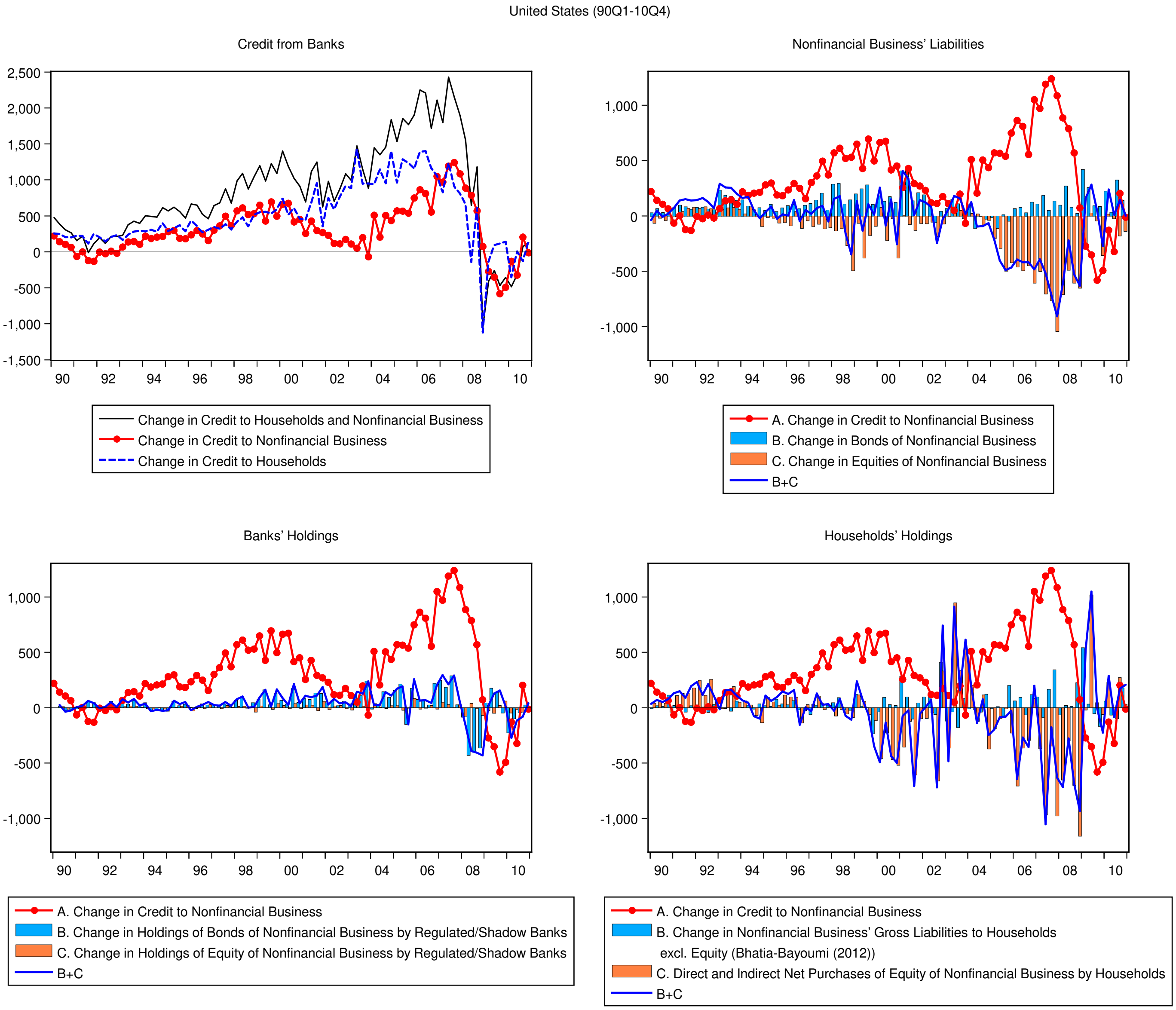
Figure 8: Direct Financing Substitution (DFS)



(Data: Flow of funds. Quarterly. Based on stock data for all series. All variables divided by the same quarter's GDP.)

((VAL+DFS)/GDP: Securities issued by Nonfinancial Business for all countries in the sample.)

Figure 9: VAL+DFS - Nonfinancial Business Issuers



(Data: Flow of funds, Bhatia and Bayoumi (2012) and authors' calculations.

Quarterly. Based on flow data. All variables divided by the same quarter's GDP.)

Figure 10: Net Sales of New Securities - US Detail