The Macro Implications of Narrow Banking:
Financial Stability versus Growth

Sebastian Merkel†
Princeton University

November 12, 2020
Latest Version: [Click Here]

Abstract

This paper builds a model of the inherent instability of a banking system based on money-creating fractional reserve banks and studies the macroeconomic implications of narrow banking, a reform proposal to overcome this instability by prohibiting bank money creation. In an economy without fractional reserve banks, the interaction between the roles of money as a store of value and a medium of exchange is inherently stabilizing: when risk premia rise and induce agents to demand more money for safety reasons, the resulting deflationary pressure makes real transaction media less scarce, lowers the money premium and thus mitigates the overall price reaction. The presence of fractional reserve banks undermines this mechanism, as simultaneous bank losses lead banks to delever, thereby shrinking the supply of deposit money. This makes transaction media scarcer, which generates additional deflationary pressures, a Fisherian deflationary spiral emerges. Introducing narrow banking restores the stabilizing force that would prevail in a world without banks. However, narrow banking also lowers economic growth by increasing the need for government-provided outside money, which crowds out real productive investment.

The calibrated model suggests large quantitative effects: narrow banking reduces economic growth by 1%, almost eliminates the occurrence of banking crises, and generates welfare gains that are equivalent to a permanent 6.7% increase of consumption.

∗This paper has benefitted enormously from numerous discussions on the topic of narrow banking with Markus Brunnermeier and Jonathan Payne. In addition, I would also like to thank Klaus Adam, Mark Aguiar, Daniel Grosshans, Nobuhiro Kiyotaki, Yann Koby, Moritz Lenel, Atif Mian, Karsten Müller, Olaf Posch, Adriano Rampini, Yuliy Sannikov, Andres Sarto, Chris Sims and Gianluca Violante for comments and suggestions.
†Email: smerkel@princeton.edu.
1 Introduction

After major financial crises, some commentators regularly identify fractional reserve banking and private money creation as major culprits for the emergence of such crises and propose narrow banking as a potential cure. Advocates of such a policy claim immense macroeconomic benefits. By restricting private money creation, the financial system could become stable and immune to crises. While many economists have made or endorsed such claims, surprisingly little effort has been devoted to demonstrating these claims in formal economic models. The purpose of this paper is to analyze the macro effects of narrow banking in a dynamic general equilibrium money and banking framework.

While the term “narrow banking” has been used to describe a variety of related policy proposals, in this paper I mean by “narrow banking” an idealized version of what was advocated by Irving Fisher (Fisher, 1935): a policy that requires depository institutions that provide the economy’s medium of exchange to hold only cash and central bank reserves – thereby turning them into “narrow banks” – without imposing restrictions on the debt funding of other (non-narrow) financial institutions. A number of distinct arguments lead to the assertion that narrow banking improves financial stability. Here, I focus on the original argument that was put forward in favor of narrow banking when the policy was proposed after the Great Depression. That argument emphasizes endogenous fluctuations in the quantity of bank-created inside money (deposits) under the current system of fractional reserve banking as a destabilizing force that narrow banking seeks to remove.

In this paper, I construct a model that features this financial instability of fractional reserve banking. In the model, agents can hold two types of assets, capital assets and money. Agents hold money both as a safe store of value and as a medium of exchange. When capital risk premia rise, e.g., due to an exogenous increase in risk, this raises the demand for money as a safe asset and leads to a portfolio reallocation from capital assets to money. Relative price changes resulting from such portfolio reallocation hurt banks because their balance sheet features...

---

1 Other variants of narrow banking are discussed in the context of the related literature in Section 2. The term “narrow banking” is not universally used. Other common labels are “full reserve banking” or “sovereign money”. “Narrow banking” itself is due to Litan (1987) and is occasionally used for a broader, less restrictive class of proposals. See Lainà (2015) and Appendix A for a survey of narrow banking proposals.

2 This presumes the existence of effective legal barriers that prevent debt instruments of these other institutions to circulate as a medium of exchange.

3 The first narrow banking reform proposal in response to the Great Depression was the Chicago Plan by Knight et al. (1933). Many of the arguments for and against narrow banking are made more explicitly in the decade-long academic debate that followed this proposal, compare Section 2. In that section, I also comment on other arguments and how they relate to my analysis of narrow banking.
a mismatch between an asset side comprised of capital assets and a liability side comprised of monetary deposit liabilities. Banks respond by deleveraging and contract their balance sheets in the process.

Relative to the large literature of macro models with financial frictions in banking, the distinguishing feature of my model is that banks issue deposits that are both nominal and serve as a medium of exchange. As a consequence, bank deleveraging in response to bad shocks does not only result in a financial accelerator that works through fire sales – a reduction of the assets held by the banking sector – but it also shrinks the supply of deposit money and creates a scarcity of transaction media. This, in turn, generates deflationary pressures that increase the real value of bank liabilities and thereby amplify initial losses. The presence of such a disinflationary spiral is the key element in Fisher’s (1933; 1935) description of the role of banks in his debt deflation mechanism and the main reason for his advocacy of narrow banking.

In addition to providing a model of this destabilizing force under fractional reserve banking, my paper makes four further contributions. First, I show that narrow banking can indeed alleviate the destabilizing force. Under narrow banking, a contraction of (non-narrow) bank balance sheets is inconsequential for the supply of transaction media because bank liabilities do not serve any transaction purpose. As a result, the money supply, which then solely consists of government-provided outside money, remains unaltered. There is no additional deflationary pressure, and no additional force that amplifies the portfolio reallocation from capital to money beyond what is triggered by the initial rise in risk premia.

Second, my model shows that stabilization under narrow banking goes further. In the absence of money-creating fractional reserve banks, the interaction between the two roles of money, a safe store of value and a medium of exchange, gives rise to a stabilizing force that has not been previously discussed in the literature. Specifically, portfolio choice between capital assets and money is determined by the total premium private agents require to hold capital assets instead of money. This premium is the sum of a capital risk premium and the money premium that originates from the transaction services provided by money. If the risk premium rises, then so does the demand for money as a safe asset. This tends to increase the real value of money for reasons unrelated to the need of making transactions and thereby lowers the money premium. As a consequence, the total increase in the required capital premium is smaller than the increase in the risk premium alone, which mitigates the desire to reallocate portfolios and the resulting asset price reaction. While in principle this same interaction between the two roles of money is also present under fractional reserve banking, the simultaneous endogenous contraction in the supply of bank deposits in response to disinflationary pressures overturns the
mitigation result and instead makes the money premium more positively correlated with the risk premium.

Third, I show that, besides improving financial stability, narrow banking also decreases long-run economic growth. Interestingly, these growth effects are not the result of reduced intermediation under narrow banking due to higher bank funding costs. While bank funding costs do increase when bank liabilities are no longer a medium of exchange, this does not lead to a lower scale of banking but to more equity funding. Instead, the growth effects of narrow banking are the result of increased crowding out of investment. Specifically, my model is at its core an AK-type endogenous growth model in which the growth rate is directly related to the investment rate. Outside money is a bubble that crowds out investment in a similar way as bubbles in OLG models (e.g. Tirole, 1985): it represents – unproductive – net wealth whose presence increases consumption through a wealth effect, which is paid for by a reduction in – productive – capital investments. In contrast, bank-created inside money under fractional reserve banking is ultimately backed by capital assets and does not crowd out investment in the same way as outside money. When most money is inside money, then most money demand is effectively met by capital assets, the real value of outside money is small, and there is little crowding out. Under narrow banking, however, such inside money no longer exists and all money demand must be met by outside money. The real value of outside money must then rise to satisfy this money demand. As a consequence, there is more crowding out and the economy grows at a lower rate. It is important to emphasize that this growth effect of narrow banking is about the real value of the money stock and cannot be undone by increasing the nominal money supply.

In addition to qualitatively describing the mechanisms through which narrow banking operates, I calibrate my model to provide a quantitative estimate of the growth and stability effects of adopting a narrow banking policy. The calibrated model suggests that narrow banking would have sizeable effects. It would lower the annual growth rate by almost 1%, but it would also reduce asset price volatility by four fifths and nearly eliminate the occurrence of banking crises.

Fourth, the welfare implications of implementing narrow banking are driven by several aspects. The improved stability under narrow banking is always welfare-improving whereas the welfare effect of a larger value of outside money is ambiguous. While the lower growth rate associated with investment crowding out tends to reduce welfare, a larger value of outside money in itself can also have positive implications for risk sharing. In my calibrated model, I find large beneficial effects from introducing narrow banking; the policy would be equivalent to increasing the consumption of all agents permanently by 6.68%.
The remainder of the paper is structured as follows. In Section 2, I discuss the related literature. Section 3 presents the model. In Section 4, I study a special case of the model without banks to gain insights for the inner workings of my model that hold also for the general version. Section 5 adds banks to the analysis and contrasts fractional reserve banking and narrow banking with regard to their positive implications. In Section 6, I study the welfare implications of narrow banking. Section 7 presents a number of model extensions I have abstracted from in my baseline analysis, including the existence of non-monetary government debt, different monetary policies and bank runs. Further details on the literature, additional results, the model solution procedure and all proofs are collected in Appendices A–E.4.

2 Related Literature

The earliest wave of narrow banking proposals originated from a short memorandum signed by economists at the University of Chicago (Knight et al., 1933), commonly referred to as the “Chicago Plan”, and the subsequent academic discussion. These works have been created under the impression of the Great Depression and are largely based on the presumption that the a major source of economic fluctuation is due to variations in the supply of the economy’s medium of exchange. For this reason, these proposals see in a 100% reserve requirement for demand deposits an adequate measure for economic stabilization by itself. To my knowledge, the present paper is the first to work out these arguments in a formal model.

Despite this agreement in the underlying economic argument, early narrow banking proposals differ substantially with respect to the funding restrictions they would impose on non-narrow financial institutions to prevent the emergence of “near monies”. With regard to this dimension, individual narrow banking proposals mostly fall in between two polar cases. The first is the “Fisher-style” version of narrow banking that I analyze in this paper and that is very permissive regarding short-term debt financing of non-narrow financial institutions. The second polar case, which one may call “Simons-style narrow banking” after Henry Simons (Simons, 1936, 1946), restricts banks to fund themselves fully or primarily with equity. Ultimately, Simons-
style narrow banking is Fisher-style narrow banking combined with very tight bank equity regulation. The latter is the subject of a large literature and its financial stability implications are well-understood. By studying Fisher-style narrow banking, I theoretically isolate the bank money creation component, which is the key aspect of narrow banking that sets it apart from conventional equity capital regulation.

After the first wave of narrow banking discussions in the 1930s, a variety of narrow banking policies have been proposed, partially under different headings such as “full reserve banking”, “sovereign money” or “limited-purpose banking”. These proposals have partially shifted emphasis to other mechanisms for stabilization such as reduced shock amplification by the financial system due to lower bank leverage or reduced systemic run risk due to a reduction of run-prone liabilities issued by non-narrow financial institutions. Most of these other arguments rest on liability restrictions for non-narrow financial institutions under narrow banking schemes that are closer to Simons-style narrow banking and are thus not directly related to restricted inside money creation under narrow banking. I therefore abstract from these additional considerations.

Almost all articles on narrow banking provide a verbal analysis of the effects of the policy, but do not study formal economic models as I do. An exception is Beneš and Kumhof (2013), who revisit a version of the original Chicago Plan and study its effects in a rich quantitative DSGE model. They conclude that an implementation of the plan would lead to output gains, government debt reductions, private debt reductions and a reduction of business cycle volatility. While they need to be credited for providing the first serious model of the macroeconomic effects of narrow banking, their analysis does not fully clarify through which channels the reform operates. In contrast, the present work does not attempt to provide a comprehensive quantitative analysis as Beneš and Kumhof (2013) do, but uses a more stylized environment to illuminate how the reform would operate. Brunnermeier and Niepelt (2019) state a set of sufficient conditions under which public and private money provision are equivalent. Their equivalence requirements are neither satisfied in my model nor in most narrow banking proposals.

---


9However, some advocates of Fisher-style narrow banking – most notably Fisher (1935) himself – claim that by removing the destabilizing force resulting from money creation under fractional reserve banking, run risk would be mitigated even in the absence of additional liability restrictions on non-narrow financial institutions. I discuss the validity of such claims in the context of a model extension in Section 7.3.
My model of fractional reserve banking builds on the money and banking framework developed by Brunnermeier and Sannikov (2016c). That paper shows how a deposit-funded financial sector in a monetary economy is simultaneously exposed to losses on its asset side and disinflationary revaluation of its liabilities, both caused by portfolio reallocation from capital assets to money, and how the resulting amplification spirals threaten financial stability in the absence of accommodative monetary policy or macroprudential regulation. The key difference in my model is that bank deposits also serve as a medium of exchange. This is important, as otherwise the quantity of deposits in itself would not matter and a disinflationary spiral as described by Fisher (1933, 1935) would no longer arise.

Methodologically, my paper is most closely related to the recent continuous-time macrofinance literature pioneered by Brunnermeier and Sannikov (2014) and He and Krishnamurthy (2013). In addition to Brunnermeier and Sannikov (2016c), several other papers in that literature study monetary economies. Brunnermeier and Sannikov (2016b) use a simplified version of the Brunnermeier and Sannikov (2016c) model without financial intermediaries to analyze the real effects of money and optimal money growth policies. Di Tella (2019) extends Brunnermeier and Sannikov (2016b) by adding a transaction motive for holding money and clarifies the interaction between precautionary demands and risk premia. In these papers, the key aspect of money is that it allows agents to self-insure against idiosyncratic risk, i.e., money is a safe store of value. In my model, the medium-of-exchange role of money is equally important. This allows me to discuss the interaction of the two motives for holding money, as a safe store of value and as a medium of exchange, and how this interaction shapes the stability properties of the financial system. Drechsler, Savov, and Schnabl (2018) also develop a theory of money and banking to study risk premium dynamics and monetary policy. In their framework, money is valued because banks hold it as buffers against withdrawal shocks to their deposit liabilities. Thus, their emphasis is on money as outside liquidity required by the financial sector and they abstracts from money in the form of inside liquidity created by the financial sector for other sectors. This makes their framework less suitable for analyzing narrow banking. None of the mentioned papers studies narrow banking or other bank regulation.

Brunnermeier and Sannikov (2016c) also use the term “disinflationary spiral” and cite Fisher (1933), but in their model this refers to a disinflation channel of the financial accelerator which is distinct from the money supply channel in my model, see Footnote 47 for more details. Without the medium of exchange nature of deposits, their model is unsuitable for an analysis of narrow banking. However, my results justify their approach for their question: under fractional reserve banking, my model implies risk premium dynamics that are qualitatively very similar to an equivalent framework without a medium-of-exchange role for money.

This literature itself is part of the larger literature on macroeconomics with financial frictions originating from Bernanke and Gertler (1989), Carlstrom and Fuerst (1997), Kiyotaki and Moore (1997), and Bernanke, Gertler, and Gilchrist (1999). For a survey of that literature, see Brunnermeier, Eisenbach, and Sannikov (2013).
Li (2017) shows how the role of financial intermediaries as money creators can make their leverage procyclical if firms need to carry liquid deposits to react flexibly to investment opportunities. The outside provision of liquidity in the form of government debt can deteriorate banks’ profit margins and destabilize the banking sector. This is different from narrow banking as the bank funding restrictions imposed by narrow banking would remove the source of procyclicality in the first place in Li’s (2017) framework.

Some papers in the micro banking literature also discuss narrow banking. Wallace (1996) studies narrow banking in the context of the Diamond and Dybvig (1983) model and concludes that it would eliminate the benefits of banking. As discussed by Freixas and Rochet (2008, Section 7.2.2) and Pennacchi (2012), there are other natural interpretations of narrow banking in the Diamond and Dybvig (1983) framework that do not lead to the same conclusion. This debate is somewhat tangential to my analysis as the Diamond and Dybvig (1983) framework is primarily about consumption risk sharing, not about money or the medium-of-exchange role of demand deposits.

Diamond and Rajan (2000, 2001) argue that the run-prone nature of bank deposits is a commitment device for bankers to contribute their future human capital in the collection of loans. They conclude that narrow banking would destroy this commitment ability, but presume a funding structure of non-narrow banks that is not run-prone as in Simons-style narrow banking proposals. Their concern does not apply to the Fisher-style narrow banking analyzed here because then liabilities of non-narrow banks remain in principle run-prone.

Kashyap, Rajan, and Stein (2002) argue that narrow banking would destroy a potential synergy between deposit-taking and lending through loan commitments: to the extent that commitment takedowns and deposit withdrawals are imperfectly correlated, an integrated fractional reserve bank can economize on liquidity buffer holdings. This argument is valid even under Fisher-style narrow banking as non-narrow banks would be prohibited from making loan commitments beyond their reserve holdings. This is a potential cost of narrow banking my analysis abstracts from.

---

12 I include here only papers that explicitly discuss narrow banking. For a more extensive review of micro banking theories and what they imply for narrow banking, see Pennacchi (2012).

13 Indeed, non-narrow banks could provide efficient consumption insurance through credit provision even under narrow banking. Existing studies on narrow banking in the Diamond and Dybvig (1983) framework implicitly preclude this possibility by assuming that all financial arrangements must take the form of a narrow bank.

14 The same argument applies analogously also to the Calomiris and Kahn (1991) theory of banking. That paper, however, does not explicitly discuss narrow banking.

15 However, Pennacchi (2012) points out that there is only correlation evidence in favor of significant synergy effects after the introduction of deposit insurance. It is thus unclear to what extent the empirical evidence in favor of Kashyap, Rajan, and Stein (2002) indicates fundamental synergies as opposed to distortions from deposit
3 Model

This section presents the model. Most of the discussion is concerned with the presentation of the model for the present fractional reserve banking regime without restrictions on the backing of bank deposits. The model counterpart of a policy shift to narrow banking consists in the mere exclusion of bank deposits from the set of transaction media. It is discussed in the penultimate paragraph of Section 3.2.

3.1 Overview

The model is based on the “I theory of money” by Brunnermeier and Sannikov (2016c). I modify their framework along three dimensions. First, I simplify the production structure. Second, I consider risk shocks instead of capital shocks as this leads to cleaner theoretical results. Third, and most important, I add a medium-of-exchange role for money, which is crucial to study narrow banking.

The Brunnermeier and Sannikov (2016c) framework is particularly suitable for an analysis of the general equilibrium effects of narrow banking, including its implications for growth and financial stability. For a sensible analysis of narrow banking, one would like to have the following elements present in the model. First, there should be a role for banks to exist. Second, there should be financial frictions and some form of financial crisis dynamics to make meaningful statements about financial stability. Third, as narrow banking proposals emphasize the monetary role of bank deposits, the model should be a monetary economy with a natural role for money. Furthermore, the monetary and banking sides should interact so that the unwinding of inside money creation by fractional reserve banks generates a disinflationary spiral in the spirit of Fisher (1933) as this was originally stressed in the Chicago plan proposals. Finally, to talk about the growth implications of narrow banking, the model economy should have, fourth, some endogenous growth feature. My modification of the Brunnermeier and Sannikov (2016c) model presented here leads to a tractable model that satisfies all four requirements.

An overview of the model structure is depicted in Figure 1, which shows the balance sheets of the agents in the model economy. The left panel illustrates the situation in the present situation of fractional reserve banking. There are two types of agents, bankers and households and two assets in positive net supply, capital and outside money. Each household operates a single firm that produces output with physical capital, but is subject to idiosyncratic risk. Markets are incomplete, so that households cannot write insurance contracts for idiosyncratic insurance.
risk. However, households can hold outside money and bank deposits to self-insure against idiosyncratic fluctuations. In addition, households also need to hold these monetary assets to make transactions in the production process. Under fractional reserve banking, bank deposits are monetary and thus from the perspective of households a perfect substitute for outside money. Bankers, in turn, do not produce anything. Instead, they hold equity in a bank that they manage and that has the distinctive advantage of being able to invest into many firms simultaneously and thereby eliminate some of the idiosyncratic risk through diversification. Apart from their owner’s equity, each bank funds itself by issuing monetary nominal deposits to households.

The right panel of Figure 1 depicts the balance sheets in the model economy under narrow banking. The only difference is that bank deposits are no longer monetary, that is they can no longer be used by households to make transactions (but otherwise remain nominally risk-free short-term debt claims). Households must then resort to government-provided outside money as the only available medium of exchange.

Note that my “narrow banking economy” does not explicitly contain narrow banks. Bankers in the model under narrow banking perform the function of other (non-narrow) financial institutions. The reason is that narrow banks in the model would be pure pass-through entities for outside money and can therefore be eliminated w.l.o.g. by putting the outside money directly on households’ balance sheets.

3.2 Model Environment

Preferences. The model economy is populated by two types of agents, a unit mass of households-entrepreneurs ($h$) and a unit mass of bankers ($b$). All agents have logarithmic preferences over

---

16In my model, there is no distinction between cash and central bank reserves and I assume that the full outside money stock can be held by households. This simplifies the exposition, but is otherwise without loss of generality (w.l.o.g.).
consumption streams with an identical time preference rate $\rho$,

$$E \left[ \int_0^\infty e^{-\rho t} \log c_i^t dt \right],$$

where $i \in \{b, h\}$ indexes the agent type.\textsuperscript{17}

**Production Technology.** Each household manages a single firm which produces the homogeneous output good using physical capital as its sole input. The firm’s flow production net of reinvestment in the capital stock is

$$(a - \iota_t - \Xi_t (\psi_t)) k_t dt,$$

where $k_t$ is the household’s capital stock, $\iota_t$ is physical investment per unit of capital, and $\Xi_t (\psi_t)$ is a proportional transaction cost specified below. Absent market transactions of capital, the household’s capital stock $k_t$ evolves according to

$$\frac{dk_t}{k_t} = (\Phi (\iota_t) - \delta) dt + \tilde{\sigma}_t d\tilde{Z}_t. \quad (1)$$

Here, $\Phi$ is a concave increasing function describing the transformation technology from output goods into capital goods. Following Brunnermeier and Sannikov (2016a), I make the functional form assumption\textsuperscript{18}

$$\Phi (\iota) = \frac{1}{\phi} \log (1 + \phi \iota), \quad (2)$$

which leads in equilibrium to an analytically tractable linear Tobin’s Q condition for the optimal investment rate, see equation (21) below. $\delta$ in equation (1) is the capital depreciation rate, and $\tilde{Z}$ is a household-specific idiosyncratic Browning motion, independent across households. The idiosyncratic risk loading $\tilde{\sigma}_t$ is common for all households and assumed to depend on an exogenous Markov state $s_t$, $\tilde{\sigma}_t = \tilde{\sigma} (s_t)$, where $\tilde{\sigma} (\cdot)$ is a given increasing function and

$$d s_t = \mu_s (s_t) dt + \sigma_s (s_t) dZ_t. \quad (3)$$

\textsuperscript{17}For notational convenience, indices referring to individual agents within the type groups are suppressed throughout. Agents within each type group are identical up to the scale of their wealth and consumption (which depends on their idiosyncratic shock history), so that this should not lead to any confusion.

\textsuperscript{18}$\phi \geq 0$ is a capital adjustment cost parameter. For $\phi = 0$ this function is to be interpreted as $\Phi (\iota) = \iota$, which is the limit for $\phi \to 0$. The higher $\phi$, the larger are capital adjustment costs.
with a Brownian motion $Z$ (aggregate shock) that is independent of all $\tilde{Z}$ processes.

**Intermediation Technology.** By assumption, households are not able to write contracts on idiosyncratic risk with each other, leaving each household exposed to the idiosyncratic risk of his firm. However, households can sell off risky claims to bankers and bankers in turn are able to invest into many firms simultaneously, thereby diversifying away a fraction $1 - \beta$ of the idiosyncratic risk. Risky claims take the form of equity claims collateralized by the household’s capital. Specifically, a household operating $k_t$ units of capital sells off the cash flows generated by $\kappa_t k_t$ units of capital, $\kappa_t \in [0, 1]$, as risky claims to bankers in exchange for a locally deterministic capital management fee.

For households to remain relevant for production, they must be restricted from selling off all capital risk to banks by choosing $\kappa = 1$. This is also realistic, as agency problems typically impose limits on outside funding. Rather than modeling agency conflicts explicitly, here I assume a parsimonious reduced-form skin-in-the-game constraint for households, $\kappa_t \leq \bar{\kappa}$ where $\bar{\kappa} \in (0, 1)$ is a model parameter. This implies that households must at all times be exposed to a fraction $1 - \bar{\kappa}$ of their capital risk.

**Assets.** There are two assets in positive net supply, capital and intrinsically worthless outside money. The latter may nevertheless have value in equilibrium, because it is a bubble or because it facilitates transactions. Capital is used in production and can be traded on Walrasian markets at a price $q_t K_t$. Outside money is assumed to be in constant supply normalized to $M \equiv 1$. It is formally convenient to describe the value of money not by the nominal price level $P_t$ of goods, but by the real value of the total money stock per unit of capital in the economy, $q_t^M := \frac{M}{P_t K_t}$, where $K_t$ is the aggregate capital stock. Furthermore, I denote by $\vartheta_t$ the share of total wealth in the economy that is nominal wealth,

$$\vartheta_t := \frac{M/P_t}{M/P_t + q_t^M K_t} = \frac{q_t^M}{q_t^M + q_t^K}. \tag{4}$$

In addition to the two physical – positive net supply – assets, bankers may hold risky claims issued by households and in turn issue bank deposits to households. Bank deposits are assumed to be nominal short-term debt claims.

---

19This assumption is not important for my results. A non-constant money supply and monetary policy are discussed in Section 7.2.
**Transaction Costs and Medium of Exchange Role of Money.** In equilibrium, money will be valued because it allows agents to self-insure against idiosyncratic risk, even if it is otherwise entirely useless.\(^{20}\) However, bank deposits are often viewed as special because of their role in the payment system. Since the main purpose of (Fisher-style) narrow banking is to restrict the supply of such special assets to the supply of outside money (reserves), a meaningful discussion of narrow banking must include such a special role for money. Here, this is done by assuming that firms have to pay a transaction cost \(\mathcal{T}_t(v_t) k_t\) in production. This can be interpreted as a transaction cost incurred in an unmodeled supply chain.\(^{21,22}\) The transaction cost is of the form

\[
\mathcal{T}_t(v) = \left( \frac{1}{\alpha v} \left( \frac{v}{\bar{v}} \right)^\alpha - \frac{1}{\alpha v^*} \left( \frac{v^*}{\bar{v}} \right)^\alpha \right) a.
\]

Here, \(\alpha, \bar{v} > 0\) are model parameters and

\[
v_t := \frac{P_t a k_t}{m_t}
\]

is the velocity of the household’s money holdings if the household holds nominal money balances \(m_t\). Similarly, \(v^*_t\) is the velocity chosen by all other households. Subtracting the average transaction costs incurred in equilibrium makes transaction costs a pure decision wedge in the household’s problem without real resource implications. A notable special case is obtained in the limit \(\alpha \to \infty\). Then, the transaction cost specification reduces to a cash-in-advance constraint \(v_t \leq \bar{v}\) on households. If this constraint happens to bind for all households simultaneously, the quantity theory of money holds in the aggregate, \(P_t a K_t = M_t \bar{v}\).

For the purpose of computing the velocity \(v_t\), nominal money balances \(m_t\) consist of both outside money holdings and deposit holdings of the household under fractional reserve banking. Under narrow banking, only outside money holdings contribute to money balances \(m_t\), bank deposits do not. This is the only difference between fractional reserve banking and narrow banking in the model. To treat both cases simultaneously, I assume in the following that a

\(^{20}\) Even in the absence of explicit contracts, households can partially share idiosyncratic risk by trading in money: a household hit by a negative idiosyncratic shock sells money on the market to smooth consumption, households hit by positive shocks are happy to buy the money in the expectation that they can pass it on to others in the future when they suffer negative idiosyncratic shocks. If idiosyncratic risk is sufficiently high, such expectations can be self-fulfilling.

\(^{21}\) I provide a modified model with a more explicit micro foundation in the spirit of New Monetarist models Williamson and Wright (2010a,b) in Appendix E.4. This modification has only minor effect on my results.

\(^{22}\) Putting transaction costs directly into the production function results in cleaner analytical characterizations. Qualitatively, the model behaves similarly if agents face a transaction cost or cash-in-advance constraint for all consumption and real investment expenditures.
fraction $\psi \geq 0$ of deposits enter money balances $m_t$. $\psi = 1$ then corresponds to fractional reserve banking and $\psi = 0$ to narrow banking.

**Aggregate Resource Constraint.** Let $C^b_t$ and $C^h_t$ be total consumption of all bankers and households, respectively. Under the assumption that all households choose the same real investment rate $\iota_t$ (which is indeed the case in equilibrium), the aggregate resource constraint for output goods is

$$C^b_t + C^h_t + \iota_t K_t = aK_t$$

and the aggregate law of motion of capital is

$$\frac{dK_t}{K_t} = (\Phi (\iota_t) - \delta) dt,$$

where the idiosyncratic risk present in equation (1) cancels out in the aggregate.

### 3.3 Decision Problems

**Household Problem.** Up to the real investment and velocity choice, households face a standard (constrained) consumption-portfolio choice problem. Each household chooses consumption $\{c^h_t\}_{t=0}^{\infty}$, real investment $\{\iota^h_t\}_{t=0}^{\infty}$, money velocity $\{v^h_t\}_{t=0}^{\infty}$, and portfolio weights for money $\{\theta^m_{t,h}\}_{t=0}^{\infty}$, deposits $\{\theta^{d,h}_{t}\}_{t=0}^{\infty}$, capital $\{\theta^{k,h}_{t}\}_{t=0}^{\infty}$, and risky claims $\{\theta^{x,h}_{t}\}_{t=0}^{\infty}$ to maximize utility

$$\mathbb{E}\left[ \int_0^{\infty} e^{-\rho t} \log c^h_t dt \right]$$

subject to the following constraints:

1. the household’s net worth evolution,

$$\frac{dn^h_t}{n^h_t} = -\frac{c^h_t}{n^h_t} dt + \theta^{m,h}_t dR^m_t + \theta^{d,h}_t dR^d_t + \theta^{k,h}_t dR^k_t (\iota_t, v_t) + \theta^{x,h}_t dR^x_t; \quad (8)$$

2. the constraint that portfolio weights sum to one, $\theta^{m,h}_t + \theta^{d,h}_t + \theta^{k,h}_t + \theta^{x,h}_t = 1$;

3. no-short sale constraints for money, deposits, and physical capital $\theta^{m,h}_t, \theta^{d,h}_t, \theta^{k,h}_t \geq 0$;

4. risky claims can only be issued, not purchased, $\theta^{x,h}_t \leq 0$;

5. the skin-in-the-game constraint for issuance of risky claims, $\kappa_t := -\theta^{x,h}_t / \theta^{k,h}_t \leq \bar{\kappa}$;
6. the definition of velocity,\textsuperscript{23}

\[ v_t = \frac{a}{q_t^K} \frac{\theta_t^{k,h}}{\psi \theta_t^{d,h}}, \tag{9} \]

where \( \psi = 1 \) under fractional reserve banking and \( \psi = 0 \) under narrow banking.

In this problem, the household takes the return processes \( dR_t^m, dR_t^d, dR_t^k (\iota_t, v_t), dR_t^x \) in equation (8) as given. These processes are stated explicitly in the end of this subsection.

**Banker Problem.** Also bankers face a standard consumption-portfolio choice problem. Each banker chooses consumption \( \{c_{bt}\}_{t=0}^\infty \), and portfolio weights for money \( \{\theta_{m,b,t}\}_{t=0}^\infty \), deposits \( \{\theta_{d,b,t}\}_{t=0}^\infty \), and risky claims \( \{\theta_{x,b,t}\}_{t=0}^\infty \) to maximize utility

\[ E \left[ \int_0^\infty e^{-\rho t} \log c^b_t dt \right] \]

subject to the following constraints:

1. the banker’s net worth evolution,

\[ \frac{dn_{bt}}{n_{bt}} = -\frac{c^b_t}{n_{bt}^h} dt + \theta_t^{m,b} dR_t^m + \theta_t^{d,b} dR_t^d + \theta_t^{x,b} dR_t^x; \tag{10} \]

2. the constraint that portfolio weights sum to one, \( \theta_t^{m,b} + \theta_t^{d,b} + \theta_t^{x,b} = 1 \);

3. no-short sale constraints for money and risky claims \( \theta_{m,b,t}, \theta_{x,b,t} \geq 0 \).

The problem of bankers differs from that of households in three respects. First, bankers cannot directly invest in capital, but only in risky claims issued by households. Second, they face a different risky claim return process \( dR_t^{x,b} \) instead of \( dR_t^x \), where the former reflects the partially diversified portfolio of risky claims the banker holds. Third, they can – and regularly do – hold a short position in bank deposits, that is issue deposits to households.

**Return Processes.** The problem description above is completed by an explicit statement of the return processes \( dR_t^m, dR_t^d, dR_t^k (\iota_t, v_t), dR_t^x, dR_t^{x,b} \) appearing in equations (8) and (10).

\textsuperscript{23}Nominal money balances are \( P_t \left( \theta_t^{m,h} + \theta_t^{d,h} \right) n_t^h \) under fractional reserve banking and \( P_t \theta_t^{m,h} n_t^h \) under narrow banking. Substituting this and \( k_t = \theta_t^{k,h} n_t^h / q_t^K \) into equation (5) yields equation (9).
To derive these return processes, assume that prices \( q^K \) and \( q^M \) follow a geometric Ito process in equilibrium,\(^{24}\)

\[
\begin{align*}
\frac{dq^K_t}{q^K_t} &= \mu_t^{q,K} dt + \sigma_t^{q,K} dZ_t, \\
\frac{dq^M_t}{q^M_t} &= \mu_t^{q,M} dt + \sigma_t^{q,M} dZ_t
\end{align*}
\]  

(11)

with potentially time- and state-dependent drifts \( \mu_t^{q,K}, \mu_t^{q,M} \) and volatilities \( \sigma_t^{q,K}, \sigma_t^{q,M} \) that are determined in equilibrium.\(^{25}\) The return on capital as a function of the managing household’s investment \( \iota_t \) and velocity \( v_t \) choice is then

\[
\begin{align*}
\text{dividend yield} & \quad \frac{dR^{K}_t (\iota_t, v_t)}{q^K_t k_t} = \left( \frac{a - \iota_t - \Phi (\iota_t)}{q^K_t k_t} + \Phi (\iota_t) - \delta + \mu_t^{q,K} \right) dt + \sigma_t^{q,K} dZ_t + \tilde{\sigma}_t d\tilde{Z}_t. \\
\text{capital gains} & \quad \frac{d\left( q^K_t k_t \right)}{q^K_t k_t} = \left( \frac{a - \iota_t - \Phi (\iota_t)}{q^K_t k_t} + \Phi (\iota_t) - \delta + \mu_t^{q,K} \right) dt + \sigma_t^{q,K} dZ_t + \tilde{\sigma}_t d\tilde{Z}_t.
\end{align*}
\]  

(12)

The return on a risky claim collateralized by this capital must have the same risk terms, but may differ in its expected return due to the management premium households earn, that is

\[
\begin{align*}
\text{dividend yield} & \quad \frac{dR^x_t}{q^K_t} = \frac{d}{q^K_t} \left( \frac{\Phi (\iota_t)}{q^K_t} - \delta + \mu_t^{q,K} \right) dt + \sigma_t^{q,K} dZ_t + \tilde{\sigma}_t d\tilde{Z}_t. \\
\text{capital gains} & \quad \frac{d\left( q^K_t k_t \right)}{q^K_t k_t} = \left( \frac{a - \iota_t - \Phi (\iota_t)}{q^K_t k_t} + \Phi (\iota_t) - \delta + \mu_t^{q,K} \right) dt + \sigma_t^{q,K} dZ_t + \tilde{\sigma}_t d\tilde{Z}_t.
\end{align*}
\]  

(13)

The return on a risky claim collateralized by this capital must have the same risk terms, but may differ in its expected return due to the management premium households earn, that is

\[
\begin{align*}
\text{dividend yield} & \quad \frac{dR^{x,b}_t}{q^K_t} = \frac{d}{q^K_t} \left( \frac{\Phi (\iota_t)}{q^K_t} - \delta + \mu_t^{q,K} \right) dt + \sigma_t^{q,K} dZ_t + \tilde{\sigma}_t d\tilde{Z}_t. \\
\text{capital gains} & \quad \frac{d\left( q^K_t k_t \right)}{q^K_t k_t} = \left( \frac{a - \iota_t - \Phi (\iota_t)}{q^K_t k_t} + \Phi (\iota_t) - \delta + \mu_t^{q,K} \right) dt + \sigma_t^{q,K} dZ_t + \tilde{\sigma}_t d\tilde{Z}_t.
\end{align*}
\]  

(14)

Bankers invest into many of these risky claims and thereby eliminate a fraction \( 1 - \beta \) of the idiosyncratic risk through diversification. The return process of the risky claim portfolio of a banker is therefore

\[
\begin{align*}
\text{dividend yield} & \quad \frac{dR^{x,b}_t}{q^K_t} = \frac{d}{q^K_t} \left( \frac{\Phi (\iota_t)}{q^K_t} - \delta + \mu_t^{q,K} \right) dt + \sigma_t^{q,K} dZ_t + \tilde{\sigma}_t d\tilde{Z}_t. \\
\text{capital gains} & \quad \frac{d\left( q^K_t k_t \right)}{q^K_t k_t} = \left( \frac{a - \iota_t - \Phi (\iota_t)}{q^K_t k_t} + \Phi (\iota_t) - \delta + \mu_t^{q,K} \right) dt + \sigma_t^{q,K} dZ_t + \tilde{\sigma}_t d\tilde{Z}_t.
\end{align*}
\]  

(15)

Outside money does not pay a dividend, so its return is simply equal to its capital gains rate,

\[
\begin{align*}
\text{dividend yield} & \quad \frac{dR^m_t}{q^K_t} = \frac{d}{q^K_t} \left( \frac{\Phi (\iota_t)}{q^K_t} - \delta + \mu_t^{q,M} \right) dt + \sigma_t^{q,M} dZ_t + \tilde{\sigma}_t d\tilde{Z}_t.
\end{align*}
\]  

(16)

\(^{24}\)The Ito process assumption is common in the literature and rules out (i) nonabsolutely continuous time drifts and (ii) price jumps. The former is economically of little relevance, the assumption of no price jumps eliminates the possibility of discontinuous sunspot fluctuations that typically exist in these types of models (see Brunnermeier and Sannikov (2015), Mendo (2018) and the model extension presented in Section 7.3).

\(^{25}\)These assumptions immediately imply that also \( \vartheta \) must be an Ito process. In line with the notation for \( q^K \) and \( q^M \), its drift and volatility are denoted by \( \mu_t^\vartheta \) and \( \sigma_t^\vartheta \).
where it has been used that $1/P_t = q_t^M K_t$ by the definition of $q^M$ and that $M = 1$ by assumption. Deposits are nominally risk-free debt claims, that is debt claims that promise payments in outside money. The return on deposits therefore equals the return on outside money up to an additional nominal deposit rate $i^d_t$,

$$dR_t^d = i^d_t dt + dR_t^m = \left(i^d_t + \Phi(t) - \delta + \mu t^q,q^M\right) dt + \sigma t^q,q^M dZ_t. \quad (16)$$

### 3.4 Equilibrium

In the following, let $N^i$ denote the aggregate net worth of agents of type $i \in \{b, h\}$ and let $\eta := \frac{N^b}{N^b + N^h}$ be the share of aggregate wealth owned by bankers. Note that total wealth in the economy satisfies the equation $N^b_t + N^h_t = (q_t^K + q_t^M) K_t$ because net wealth only exists in the form of capital and outside money.\(^{26}\)

The model economy features five markets, for goods, capital, risky claims, money, and deposits. The goods market clears if the aggregate resource constraint (6) is satisfied. The capital market clears if the total value of the capital stock, $q_t^K K_t$, equals total capital demand by households, $\theta_t^{k,h} N^h_t$. After rearranging and using the definition of $\vartheta$ (equation (4)), capital market clearing can be equivalently written as

$$\theta_t^{k,h} = 1 - \vartheta_t \frac{1}{1 - \eta_t}. \quad (17)$$

The market for risky claims clears if the value of claims issued by households, $-\theta_t^{x,h} N^h_t$, equals the value of claims demanded by bankers, $\theta_t^{x,b} N^b_t$. Risky claims market clearing and the definition of $\kappa_t (= -\theta_t^{x,h}/\theta_t^{k,h})$ are then equivalent to the two equations

$$\theta_t^{x,b} = \frac{\kappa_t}{\eta_t} (1 - \vartheta_t), \quad \theta_t^{x,h} = -\frac{\kappa_t}{1 - \eta_t} (1 - \vartheta_t). \quad (18)$$

The money market clears if the value of outside money, $q_t^M K_t$, equals the value of outside money demanded by bankers and households, $\theta_t^{m,b} N^b_t + \theta_t^{m,h} N^h_t$, or equivalently

$$\theta_t^{m,b} \eta_t + \theta_t^{m,h} (1 - \eta_t) = \vartheta_t. \quad (19)$$

Finally, the deposit market clears by Walras’ law.

As is standard in the literature, I limit attention to Markov equilibria. There are three

---

\(^{26}\)In this model, outside money is net wealth because it is a bubble.
state variables, the exogenous state \( s_t \) that affects the level of idiosyncratic risk \( \sigma_t \) in the economy, the wealth distribution across types summarized by the wealth share of bankers \( \eta_t \), and the aggregate capital stock \( K_t \). Total output and wealth are linear in \( K_t \) and individual decision problems are scale-invariant in individual net worth. Therefore, movements in \( K_t \) simply scale the overall economy up or down, but leave allocations otherwise unaffected. Up to scale, everything can be expressed in terms of the two remaining state variables \( s_t \) and \( \eta_t \).

The formal equilibrium definition for the descaled economy is then as follows. Here, I use the notation \( \check{c}^i := c^i/n^i \), \( i \in \{b,h\} \) for the consumption-wealth ratios chosen by agents of type \( i \in \{b,h\} \).

**Definition 1.** A competitive Markov equilibrium in the state variable \((s, \eta)\) consists of functions \( q^K, q^M, \iota, \mu^n, \sigma^n, \kappa, \check{c}^b, \check{c}^h, \theta^{m,b}, \theta^{m,h}, r^x, i^d : (0, \infty) \times [0, 1) \rightarrow \mathbb{R} \) such that with \( \vartheta \) defined as in (4)

1. The exogenous state \( s_t \) evolves according to (3) and is expected by all agents to do so;
2. Bankers’ wealth share \( \eta_t \) evolves according to
   \[
   \frac{d\eta_t}{\eta_t} = \mu^n(s_t, \eta_t) \, dt + \sigma^n(s_t, \eta_t) \, dZ_t
   \]
   and is expected by all agents to do so;
3. \( \check{c}^h_t = \check{c}^h(s_t, \eta_t) n^h_t, \, \iota_t = \iota(s_t, \eta_t), \, \nu_t \) as in (9), \( \theta_t^{m,h} = \theta^{m,h}(s_t, \eta_t), \, \theta_t^{\kappa,b} \) as in (17), \( \theta_t^{r,h} \) as in (18) and \( \theta_t^{d,h} = 1 - \theta_t^{m,h} - \theta_t^{\kappa,h} - \theta_t^{r,h} \) solve the household’s problem, where individual household net worth \( n^h_t \) evolves according to (8) and returns are given by (12), (13), (15), and (16);
4. \( \check{c}^b_t = \check{c}^b(s_t, \eta_t) n^b_t, \, \theta_t^{m,b} = \theta^{m,b}(s_t, \eta_t), \, \theta_t^{x,b} \) as in (18) and \( \theta_t^{d,b} = 1 - \theta_t^{m,b} - \theta_t^{\kappa,b} - \theta_t^{x,b} \) solve the banker’s problem, where individual bank net worth \( n^b_t \) evolves according to (10) and returns are given by (14), (15), and (16);
5. \( r^x_t = r^x(s_t, \eta_t) \) in return equations (13) and (14) and \( i^d_t = i^d(s_t, \eta_t) \) in return equation (16);

The wealth distribution of individual agents within each type group does not matter for macro aggregates, as individual decision problems are scale-invariant in individual net worth \( n' \) and therefore the aggregate choices of the household and banking sector are independent of the wealth distribution within these sectors.
6. The goods market clears, \(^{28}\)

\[
\left( e^b (s_t, \eta_t) \eta_t + e^b (s_t, \eta_t) (1 - \eta_t) \right) (q^K (s_t, \eta_t) + q^M (s_t, \eta_t)) = a - \iota (s_t, \eta_t);
\]

7. The money market clears, that is equation (19) holds;

If, in addition, \( q^M (s, \eta) > 0 \) for \( s \) in the support of the ergodic distribution of \( s_t \) and for all \( \eta \in [0, 1) \), then the equilibrium is called a competitive monetary Markov equilibrium.

Note that capital and risky claims market clearing is already implicit in the representation of portfolio shares \( \theta^{ji} \) in 3. and 4. This equilibrium definition is the same in both the fractional reserve banking and the narrow banking regime. The only difference is that the velocity equation (9) in 3. differs across regimes because \( \psi \) takes a different value.

In the following, the term “equilibrium” always refers to a competitive monetary Markov equilibrium. This convention involves an implicit equilibrium selection, as I rule out the no bubble equilibrium and the continuum of nonstationary (non-Markovian in \((s, \eta)\)) inflationary equilibria that typically exist in rational bubble models of money. \(^{29}\)

### 3.5 Parameter Assumptions

For the remainder of the paper, I also make two additional parameter assumptions. Both assumptions are not very restrictive, but required in some places. The first assumption excludes situations in which the monetary friction is so severe that agents may prefer to dispose their capital instead of coming up with the necessary money balances to operate it. The second assumption ensures that in steady state bankers always issue deposits. It merely excludes the practically irrelevant case in which banks issue deposits only occasionally.

**Assumption 1.** Model parameters satisfy

\[
\tilde{v} > \left( \frac{\alpha - 1}{\alpha \rho} \right)^{\frac{1}{\alpha}} \frac{\phi a}{1 + \phi a} \rho.
\]

**Assumption 2.** The function \( \mu_s (s) \) in the exogenous state process (3) is nonincreasing and has a unique zero \( \bar{s} \geq 0 \) (“steady state”). At \( \bar{s} \), model parameters satisfy the inequality \( \beta \tilde{\sigma} (\bar{s}) < \sqrt{\rho} \).

\(^{28}\)This is a reformulation of (6) that uses \( C^i_t = c^i_t N^i_t \ (i \in \{b, h\}) \) and \( N^b_t + N^h_t = (q^K_t + q^M_t) K_t \).

\(^{29}\)One can make the equilibrium selection explicit by adding out-of-equilibrium fiscal backing of money to the monetary Markov equilibrium as in Brunnermeier, Merkel, and Sannikov (2020).
4 Special Case without Bankers

In this section, I discuss a version of the model without bankers, since it allows me to illustrate much of the inner workings of my model in a simplified setting. In particular, this section clarifies the determinants of household money demand and explains how money crowds out investment, an important component of the mechanism behind the growth implications of narrow banking.

Formally, the economy without bankers corresponds to the full model at the absorbing state $\eta = 0$. Then, bankers have no net worth and thus they do not consume and optimally choose to intermediate no capital, $\kappa = 0$. The economy is then isomorphic to a model without bankers. Versions of this model have been studied previously by Brunnermeier and Sannikov (2016b), Di Tella (2019) and Brunnermeier, Merkel, and Sannikov (2020).\footnote{The version in Di Tella (2019) is closest to the present model with the only difference being the money demand specification.}

The solution of the model proceeds in two steps.\footnote{I only sketch the solution procedure here. Additional technical details and proofs of the propositions can be found in Appendix B.1.1.} The first step takes the portfolio choice between money and capital and thus the money wealth share $\vartheta$ as given and derives the resulting equilibrium consumption-savings decision of agents and implied values for $q^K$ and $q^M$. Specifically, this step combines three equilibrium conditions. The first is the optimal consumption choice of households. Here, households have logarithmic preferences, which have the well-known property that the optimal consumption rule involves consuming a constant fraction $\rho$ of net worth $n^h_t$,

$$c^h_t = \rho n^h_t. \quad (20)$$

This condition immediately implies that all households choose the same consumption-wealth ratio $z^h_t = \rho$. The second condition is the optimal investment choice. The investment rate $\iota_t$ enters the household problem only through the drift term in the capital return expression (12) and thus it is optimal to choose $\iota_t$ to maximize the expected capital return, which implies the Tobin’s Q condition

$$q^K_t = \frac{1}{\Phi'(\iota_t)} = 1 + \phi \iota_t. \quad (21)$$

A consequence of this equation is that all households choose the same investment rate $\iota_t$ because
they face the same capital price \( q^K_t \). The third condition is goods market clearing:

\[
\tilde{c}_t^h (p^K_t + q^M_t) = a - \iota_t. \tag{22}
\]

Combining the three equations and using the definition of \( \vartheta_t \) (equation (4)) to substitute \( 1 - \vartheta_t \) for \( q^K_t / (p^K_t + q^M_t) \) results in a single equation that relates \( \iota_t \) to \( \vartheta_t \). After solving for \( \iota_t \), equation (21) determines \( q^K_t \) and equation (4) determines \( q^M_t \). The result is summarized in the following proposition.

**Proposition 1.** In equilibrium, asset prices and the investment rate are given by

\[
q^K_t = (1 - \vartheta_t) \frac{1 + \phi a}{1 - \vartheta_t + \phi \rho}, \quad q^M_t = \vartheta_t \frac{1 + \phi a}{1 - \vartheta_t + \phi \rho}, \quad \iota_t = (1 - \vartheta_t) a - \rho \frac{1 - \vartheta_t + \phi \rho}{1 - \vartheta_t + \phi \rho}.
\]

Besides reducing the number of equilibrium variables to solve for by expressing the three unknowns \( q^K_t, q^M_t \) and \( \iota_t \) in terms of the single unknown \( \vartheta_t \), this representation of asset prices and the investment rate has two important economic insights. First, the equation implies that asset price fluctuations are the result of fluctuations in \( \vartheta_t \). By the money market clearing condition (19), \( \vartheta_t \) is directly related to the aggregate demand for money, so the ultimate cause of asset price fluctuations in this model is portfolio reallocation between capital and money. The only source of aggregate risk in this model is asset price risk \( \sigma_{q,K}^t \) and \( \sigma_{q,M}^t \) in the return expressions for capital, money and risky claims. Thus, any force that amplifies such reallocation amplifies aggregate risk, any force that mitigates portfolio reallocation mitigates aggregate risk.

Second, an increase in \( \vartheta_t \) does not simply increase \( q^M_t \) and decrease \( q^K_t \) by identical amounts, but it also increases the sum \( q^K_t + q^M_t \), that is total wealth per unit of capital in the economy. In other words, for a given quantity of capital \( K_t \), a larger fraction of money wealth in the economy is associated with larger total wealth. As agents feel overall more wealthy, a wealth effect increases their consumption demand. Yet, total output \( a K_t \) is independent of \( \vartheta_t \) as the additional consumption demand derives from unproductive money wealth. Thus, market clearing requires that larger money wealth is met by a reduction in investment activity. Therefore, \( \iota_t \) is a decreasing function of \( \vartheta_t \). In this sense, **money crowds out investment.**

---

32This conclusion is only true for \( \phi > 0 \). However, in the special case \( \phi = 0 \), individual households are indifferent between any \( \iota_t \) value and it is thus w.l.o.g. to assume also then that they all choose the same investment rate.

33Compare Definition 1.

34The same mechanism is also present in other models of rational bubbles, e.g. Tirole (1985).
investment.

Proposition 1 summarizes the results of the first step in the model solution procedure. The second step consists in characterizing the remaining unknown, the relative money demand $\vartheta_t$ by households’ portfolio choice between capital and money. The following proposition summarizes the resulting equilibrium conditions.

**Proposition 2.** The equilibrium money wealth share $\vartheta_t$ and velocity $v_t$ are jointly determined by two money demand conditions:

1. The money valuation equation,
   \[
   \mu_t^d = \rho - \left( (1 - \vartheta_t)^2 \tilde{\sigma}_t^2 + \lambda_t v_t \right) \tag{23}
   \]

2. The quantity equation,
   \[
   v_t = \frac{a}{1 + \phi a} \left( 1 - \vartheta_t + \phi \rho \right) \vartheta_t \tag{24}
   \]

Here, $\lambda_t$ is a measure of the value of transaction services\(^{35}\) and satisfies the condition

\[
\begin{cases}
\lambda_t = \frac{1}{a} \mathcal{F}'(v_t) v_t, & \alpha < \infty \\
\min\{\lambda_t, \bar{v} - v_t\} = 0, & \alpha = \infty
\end{cases}
\]

The two money demand conditions in the proposition emphasize different roles of money. The first, which I call the “money valuation equation” following the terminology of Brunnermeier and Sannikov (2016c), is about money as a store of value. It captures the portfolio choice between money and capital assets. Formally, this equation is a backward stochastic differential equation (BSDE) for the the money wealth share $\vartheta$. It describes the expected evolution of $\vartheta$ on any (rational expectations) equilibrium path. For interpretation, it is more instructive to write the equation in integral form,

\[
\vartheta_t = \mathbb{E}_t \left[ \int_t^\infty e^{-\rho(s-t)} \left( (1 - \vartheta_s)^2 \tilde{\sigma}_s^2 + \lambda_s v_s \right) \vartheta_s ds \right]. \tag{25}
\]

This equation states that $\vartheta_t$ is the discounted integral of the sum of two premia that relate to two motives for holding money in this economy. The first, $\lambda_t$, is the risk premium required by households to bear their idiosyncratic net worth risk. Because trading money

\(^{35}\)In the cash-in-advance case ($\alpha = \infty$), $\lambda_t$ is the Lagrange multiplier on the cash-in-advance constraint.
allows agents to self-insure against idiosyncratic wealth fluctuations, money is more valuable if this idiosyncratic risk premium is large. The second term, $\lambda_s v_s$, is the money premium and measures the transaction services that money provides to households.

The value of the money premium links the money valuation equation to the second condition in Proposition 2, the quantity equation, which emphasizes the role of money as a medium of exchange. The equation as stated in the proposition is simply the definition of velocity combined with market clearing. Its economic content originates from the relationship between velocity and the money premium $\lambda_s v_s$ expressed in condition (24). The interplay between the two money demand conditions determines equilibrium money demand.

Propositions 1 and 2 together with the law of motion (3) of the exogenous state fully determine the equilibrium of the model without bankers. It is instructive to consider first the case $\sigma_s \equiv 0$ with a mean-reverting drift function $\mu_s(s) = b(s - \bar{s})$ and discuss comparative statics of the model’s steady state with respect to the idiosyncratic risk parameter $\tilde{\sigma} := \tilde{\sigma}(\bar{s})$. For sharp theoretical results, I limit attention to the case of a cash-in-advance constraint, $\alpha = \infty$. Then the steady-state solution is available in closed form:

**Proposition 3.** Assume $\sigma_s(s) = 0$, $\mu_s(s) = b(s - \bar{s})$ (deterministically mean-reverting state process) and $\alpha = \infty$ (cash-in-advance constraint). Then there is a unique (monetary) steady state equilibrium. In this steady state, $p^K$, $p^M$ and $\iota$ are as in Proposition 1 and $\vartheta = \max\{\vartheta^I, \vartheta^M\}$ with two mutually exclusive regimes:

1. “I theory regime”:
   If $\tilde{\sigma} \geq \tilde{\sigma}^I := \frac{\bar{\delta} + \alpha(1 + \varphi)}{\bar{\delta} + \alpha(\bar{\delta} - \rho)} \sqrt{\rho}$, the cash-in-advance constraint is slack, $v < \bar{v}$, $\lambda = 0$, and money demand is dominated by the self-insurance motive. The steady-state value of $\vartheta$ is determined by the steady-state money valuation equation and given by
   $$\vartheta = \vartheta^I := \frac{\tilde{\sigma} - \sqrt{\rho}}{\tilde{\sigma}},$$
   in particular, it is (locally) strictly increasing in idiosyncratic risk $\tilde{\sigma}$, strictly decreasing in the time preference rate $\rho$ and independent of all other model parameters.

2. “monetarist regime”:
   If $\tilde{\sigma} < \tilde{\sigma}^I$, the cash-in-advance constraint binds, $v = \bar{v}$, $\lambda > 0$, and money demand is dominated by the transaction motive. The steady-state value of $\vartheta$ is determined by the

---

36How money trading allows agents to self-insure and thereby makes money valuable to agents is discussed in more detail by Di Tella (2019) and Brunnermeier, Merkel, and Sannikov (2020)
quantity equation (with $v = \bar{v}$) and given by

$$\vartheta = \vartheta^M := \frac{a (1 + \phi \rho)}{\bar{v} + a (1 + \phi \bar{v})},$$

in particular, it is (locally) strictly increasing in $a$, (weakly) increasing in $\rho$, strictly decreasing in $\bar{v}$, and independent of idiosyncratic risk $\tilde{\sigma}$. The money premium $\lambda \bar{v}$ adjusts to make the steady-state money valuation equation hold for $\vartheta = \vartheta^M$,

$$\lambda \bar{v} = \rho - (1 - \vartheta^M)^2 \tilde{\sigma}^2.$$

This proposition says that in steady state one of the two motives for holding money dominates and just one of the two money demand conditions of Proposition 2 matters. When idiosyncratic risk is high, the economy is in what one may call the “I theory regime”, because the transaction motive does not matter and the model reduces to a version of the I theory of money (Brunnermeier and Sannikov, 2016c) without banks. In this regime, the money valuation equation determines the value of $\vartheta$, whereas the quantity equation is not of independent economic significance. $\vartheta$ must then be strictly increasing in idiosyncratic risk $\tilde{\sigma}$: the money valuation equation relates $\vartheta$ to the risk premium on idiosyncratic risk and for any given $\vartheta$ a larger $\tilde{\sigma}$ is associated with a larger risk premium $(1 - \vartheta)^2 \tilde{\sigma}^2$. In particular, an (unexpected) increase in risk in this regime leads to a portfolio reallocation from capital to money.

In contrast, when idiosyncratic risk is low, the transaction motive dominates money demand and the quantity equation with $v = \bar{v}$ determines the value of money. Now instead the money valuation equation loses independent economic significance. It simply determines passively the money premium required to make $\vartheta = \vartheta^M$ consistent with agents’ portfolio choice. In this regime, $\vartheta$ depends on the maximum velocity and on capital productivity, but no longer on idiosyncratic risk $\tilde{\sigma}$. The reason is that any risk premium increase caused by a larger $\tilde{\sigma}$ is precisely offset by a fall in the money premium such that the total premium $(1 - \vartheta^M)^2 \tilde{\sigma}^2 + \lambda \bar{v}$ in the money valuation equation remains unaffected. In particular, an (unexpected) rise in risk does not lead to a portfolio reallocation from capital to money in this regime. Because of the prominent role of the quantity equation, I call this regime the “monetarist regime”.

The $\tilde{\sigma}$-comparative statics in the monetarist regime hint at a self-stabilizing force of the economic system in the presence of a transaction demand for money: portfolio reallocation in response to changes in risk premia that are unrelated to the need to make transactions is mitigated by offsetting movements in the money premium. The reason is that an increase in

23
risk premia raises demand for money as a safe store of value, which in isolation would cause agents to demand more money and drive up the relative money value \( \vartheta \). But as money becomes more valuable relative to capital, transaction media become less scarce, velocity falls and this drives down the money premium. The latter effect limits the additional money demand to begin with and the overall effect on \( \vartheta \). In the case of a binding cash-in-advance constraint, only the money premium adjusts and the equilibrium money value and velocity remain unaltered. More generally (for \( \alpha < \infty \)), one should expect some velocity adjustment and portfolio reallocation to take place.

Figure 2 shows that this intuition derived from the steady state equations carries over to the dynamic setting. The figure has been created under the assumption that idiosyncratic risk follows the Heston model of stochastic volatility (Heston, 1993) which requires choosing \( \tilde{\sigma}(s) = \sqrt{s} \) and

\[
\frac{ds_t}{s_t} = b(\bar{s} - s)dt - \sigma_s \sqrt{s_t} dZ_t
\]

for the exogenous risk process. The left panel depicts the equilibrium value of \( \vartheta \) as a function of \( \tilde{\sigma} \). The blue line simply plots the steady-state comparative statics with respect to \( \bar{s} \) in line with the equations in Proposition 3. The remaining two lines plot the equilibrium solution for a fixed \( \bar{s} \) in the dynamic model with risk shocks, with a cash-in-advance constraint (red solid}
line) and a transaction cost technology featuring $\alpha < \infty$ (purple dashed line). The right panel depicts the equilibrium money premium $\lambda v$ as a function of $\tilde{\sigma}$ for the same specifications using the same color coding as the left panel. The figure shows that the steady-state comparative statics and the dynamic equilibrium solution for the cash-in-advance model are qualitatively very similar. In both cases, $\vartheta$ is constant for low $\tilde{\sigma}$ and increases in risk are met by decreases in the money premium. When the money premium reaches zero, $\vartheta$ is strictly increasing in $\tilde{\sigma}$. The only difference in the dynamic model is that $\vartheta$ increases less steeply in $\tilde{\sigma}$ because agents expect the economy to revert back to the “monetarist” region with a lower equilibrium value of $\vartheta$.

The dashed lines in Figure 2 confirm that the economic logic of the cash-in-advance model carries over to the version with a transaction cost technology. The strict decomposition of the state space into a monetarist and an I theory region is no longer valid. $\vartheta$ is everywhere increasing in $\tilde{\sigma}$ – as in the I theory regime – and the money premium is everywhere positive – as in the monetarist regime. Yet, the solution functions still retain a similar qualitative shape as in the cash-in-advance model. In the left part of the state space, the $\vartheta$ function is relatively insensitive to changes in $\tilde{\sigma}$ and the money premium decreases strongly in response to rising $\tilde{\sigma}$. In the right part of the state space, the opposite pattern holds. In this sense, the solution still suggests the presence of a “monetarist region” for low $\tilde{\sigma}$ and an “I theory region” for large $\tilde{\sigma}$, although the boundary between the two is somewhat blurred relative to the cash-in-advance case.

I close this section with a preview of the effects of adding banks to the analysis in the next section. Regardless of banking policy, banks’ ability to diversify risk improves agents’ ability to manage idiosyncratic risk so that less idiosyncratic risk remains on households’ balance sheets. If banks are fractional reserve banks, then they also provide additional transaction media to households, thereby decreasing velocity. If banks are operating under a narrow banking policy with non-monetary deposits liabilities, their presence does not directly affect the supply of transaction media. These considerations alone suggest (1) that under fractional reserve banking the money premium is likely lower and thus $\vartheta$ tends to be lower and (2) that under fractional reserve banking the I theory region covers a larger fraction of the state space than under narrow banking because transaction media are more abundant. The first conclusion together with the results in Proposition 1 points to more investment crowding out and lower growth under narrow banking. The second conclusion together with the relative insensitivity of the $\vartheta$ function

A similar approximate separation into two regions can also be observed with the money in utility specification of Di Tella (2019), compare his Figure 6. The high curvature region in the middle of the state space around $\sigma \approx 0.25$ in the top right and bottom left panel separates a regime akin to my monetarist regime to the left – in which the risk premium reacts strongly and investment reacts weakly to changes in risk – and a regime akin to my I theory regime to the right – in which the risk premium is almost constant and investment reacts strongly to changes in risk.
with respect to fluctuations in \( \tilde{\sigma} \) in the monetarist region point to more stability under narrow banking, provided the \( \vartheta \) function in this section is any guide to the equilibrium behavior of \( \vartheta \) in the full model. As I show in the next section, these conjectures are correct, but an incomplete description of the effects of introducing banking. While the narrow banking economy is indeed quite close to the economy without banks studied in this section, dynamics under fractional reserve banking are associated with an additional destabilizing force that undermines the self-stabilizing force in the monetarist regime discussed above.

5 Full Model Solution

In this section I add banks to the analysis and use the model to study the effects of narrow banking. Formally, I focus on an initial value of \( \eta \) that starts in the interior of the state space, \( \eta \in (0,1) \). One can then show that \( \eta \) always remains inside this interval and thus bankers are always relevant. I start by discussing the key equilibrium conditions in Section 5.1. That discussion highlights through which channels narrow banking affects the equilibrium. I then proceed by characterizing the equilibrium in the absence of aggregate shocks in Section 5.2. There, I show how narrow banking reduces both bank leverage and economic growth in steady state. I conclude the analysis of the model with banks by considering the stochastic economy in Section 5.3. That discussion reveals the presence of a dynamic amplification mechanism that is absent in the model without banks. I show there how the introduction of narrow banking mitigates amplification and thereby stabilizes the financial system.

5.1 Key Equilibrium Conditions

The solution of the full model proceeds along the same lines as the solution of the model without bankers discussed in Section 4. Here, I report and discuss the key equilibrium equations. Detailed derivations are relegated to Appendix B.1. To avoid the need for case distinctions, I make throughout the assumption that in equilibrium always \( \kappa_t (1 - \vartheta_t) \geq \eta_t \), which is equivalent to assuming that bankers always issue deposits.\(^{38}\)

As in Section 4, asset prices and the investment rate can be expressed as functions of the money wealth share \( \vartheta_t \). The derivation is precisely the same as in Section 4.

**Proposition 4.** In equilibrium, asset prices \( q_t^K, q_t^M \) and the investment rate \( \iota_t \) are as in Proposition 1.

\(^{38}\)Assumption 1 guarantees that this is indeed the case in steady state. I also verify this property for my numerical model solution.
As a consequence of this proposition, the discussion following Proposition 1 in the context of the model without banks applies equally to the model with banks. In particular, asset price movements in this model are a consequence of portfolio reallocation between capital assets and money and a larger value of outside money crowds out investment and thereby lowers growth.

Proposition 4 implies that asset prices, the investment rate and through market clearing also (aggregate) consumption are functions of the money wealth share \( \vartheta \), so that \( \vartheta \) is the key variable for model dynamics – precisely as in the model without banks. As in Section 4, \( \vartheta \) itself is determined by a money valuation equation that relates the current value of \( \vartheta \) to the expected future risk premia for aggregate and idiosyncratic risk and the expected future money premium.

**Proposition 5 (Money Valuation Equation).** The equilibrium money wealth share \( \vartheta \) satisfies the equation

\[
\mu_t \vartheta = \rho - \left( \frac{\eta_t \sigma_t^2}{\eta_t(1-\eta_t)} \right) + \left( \frac{(1-\kappa_t)^2}{1-\eta_t} + \frac{\kappa_t^2 \beta^2}{\eta_t} \right) \left( 1 - \vartheta_t \right)^2 \tilde{\sigma}_t^2 + (1 - \psi \eta_t) \lambda_t v_t ,
\]

(27)

Here, \( \eta_t \sigma_t^2 \) is the \( dZ_t \)-loading of the endogenous state variable \( \eta_t \).

What has been said in Section 4 regarding the interpretation of equation (23) applies similarly to equation (27). As there, the money valuation equation should be interpreted in integral form (compare (25)) as a discounted integral of future risk and money premia. Relative to the equation in the previous section, a third term appears here, namely \( \frac{(\eta_t \sigma_t^2)^2}{\eta_t(1-\eta_t)} \). It is the aggregate risk premium that total wealth has to earn in excess of money and measures how valuable money is as a safe asset in the sense of hedging aggregate risk. In addition, there are again an idiosyncratic risk premium, \( \left( \frac{(1-\kappa_t)^2}{1-\eta_t} + \frac{\kappa_t^2 \beta^2}{\eta_t} \right) \left( 1 - \vartheta_t \right)^2 \tilde{\sigma}_t^2 \), that measures how valuable money is to agents for self-insurance against idiosyncratic risk, and a liquidity premium proportional to the money premium \( \lambda_t v_t \), that measures the value of transaction services provided to agents by money. The risk premia and liquidity premia are weighted averages, weighted by the wealth shares \( \eta \) and \( 1 - \eta \), of the excess expected returns that bankers and households require to earn on their net worth portfolio in excess of the expected money return. The banking policy, fractional reserve banking (\( \psi = 1 \)) versus narrow banking (\( \psi = 0 \)), affects equation (27) only marginally through a difference in the liquidity premium term.\(^{39}\) This difference is of secondary importance

\(^{39}\)Under fractional reserve banking, capital assets held by bankers do not need to earn the money premium because bankers can refinance themselves at the money rate. In contrast, under narrow banking, bankers must pay the money premium on their non-monetary deposit liabilities. Then, for bankers to be willing to fund assets with non-monetary deposits, bank assets need to earn the money premium on top of the risk premium required by bankers.
to understand the effects of narrow banking.

In order to fully characterize the equilibrium, three sets of variables related to the three premium terms in equation (27) must be determined: the variables $\lambda_t$ and $\nu_t$, which determine the transaction value of money, the capital risk allocation $\kappa_t$, which is relevant for the weighted-average idiosyncratic risk premium, and the law of motion of the endogenous state $\eta_t$. These three sets of variables are discussed in the remainder of this subsection.

The first variables, $\lambda_t$ and $\nu_t$, are determined by a version of the quantity equation in analogy to Proposition 2 for the model without bankers. However, the precise form of this equation now depends crucially on the banking regime in place:

**Proposition 6 (Quantity Equation).** The equilibrium velocity $\nu_t$ satisfies the equation

$$
\nu_t = \frac{a}{1 + \phi \alpha \nu_t + \psi (\kappa_t (1 - \vartheta_t) - \eta_t)},
\tag{28}
$$

and $\lambda_t$ is determined by

$$
\begin{cases}
\lambda_t = \frac{1}{\alpha} \mathcal{J}(\nu_t) \nu_t, & \alpha < \infty \\
\min\{\lambda_t, \bar{\nu} - \nu_t\} = 0, & \alpha = \infty
\end{cases}
\tag{29}
$$

Condition (29) for the price of transaction services $\lambda_t$ is identical to condition (24) in the model without bankers. The important difference to the model without bankers is in the quantity equation (28) itself, provided a fractional reserve banking policy is in place ($\psi = 1$). Specifically, an additional term $\kappa_t (1 - \vartheta_t) - \eta_t$ appears in the denominator of $\nu_t$ under fractional reserve banking. $\kappa_t$ is the share of capital assets intermediated by bankers and $1 - \vartheta_t$ is the fraction of total wealth that is capital wealth, so $\kappa_t (1 - \vartheta_t)$ represents the value of bank assets as a fraction of total wealth. The difference between this quantity and bankers’ wealth share $\eta_t$ is the value of aggregate deposits as a fraction of total wealth. The presence of this deposit term has two important implications. First, the deposit term is always positive in equilibrium, meaning that the presence of monetary deposits under fractional reserve banking lowers velocity $\nu_t$ relative to narrow banking for any given $\vartheta_t$. In equilibrium, this means that $\vartheta_t$ is typically smaller under fractional reserve banking. Second, the deposit term $\kappa_t (1 - \vartheta_t) - \eta_t$ also matters for dynamics. As the bank balance sheet variables $\kappa_t$ and $\eta_t$ fluctuate, so does velocity and via condition (29) the money premium. The money valuation equation (27) in turn translates such money premium fluctuations into fluctuations in the relative money value $\vartheta_t$. This dynamic mechanism is absent under narrow banking.\footnote{This remark refers specifically to the channel that causes fluctuations in $\vartheta$ through fluctuations in the money} These two conclusions from equation (29) make the quantity
The share of intermediated capital $\kappa_t$ is determined jointly by banks’ risky claim demand $\theta_t^{x,b}$ and households’ supply of risky claims $\theta_t^{x,h}$ and demand for physical capital $\theta_t^{k,h}$. The resulting condition can be written as a single “capital allocation equation”:

**Proposition 7 (Capital Allocation Equation).** The equilibrium share of intermediated capital $\kappa_t$ satisfies the following condition:

$$
\min \left\{ \frac{\sigma^2_t}{1 - \eta_t} \frac{\sigma^2_t}{1 - \vartheta_t} + \left( \frac{1 - \kappa_t}{1 - \vartheta_t} - \frac{\kappa_t}{\eta_t} \beta^2 \right) (1 - \vartheta_t) \sigma^2_t + \psi \lambda_t, \bar{\kappa} - \kappa_t \right\} = 0.
$$

In both policy regimes, the capital allocation equation is a complementary slackness condition that requires both arguments in the minimum expression to be nonnegative and at least one of the two to be precisely zero. If the second argument is zero, then $\kappa_t = \bar{\kappa}$ and thus banks intermediate the maximum quantity of capital that is consistent with households’ skin-in-the-game constraints. Instead, whenever banks do not intermediate the maximum possible quantity of capital, $\kappa_t < \bar{\kappa}$, the first argument in the capital allocation equation must vanish. This first argument has an intuitive interpretation as the sum of risk and liquidity premium differentials for banks and households:

- The first term, $\frac{\sigma^2_t}{1 - \eta_t} \frac{\sigma^2_t}{1 - \vartheta_t}$, is the aggregate risk premium that households require in excess of the one banks require for holding capital assets instead of money. It represents the marginal aggregate risk sharing benefit of shifting risk from household balance sheets to bank balance sheets by increasing $\kappa_t$. In equilibrium, this term is typically negative. With regard to the aggregate risk allocation, increasing $\kappa_t$ is then costly because it shifts aggregate risk from households to banks even though banks require a higher compensation for being exposed to this risk.

- The second, $\left( \frac{1 - \kappa_t}{1 - \eta_t} - \frac{\kappa_t}{\eta_t} \beta^2 \right) (1 - \vartheta_t) \sigma^2_t$, similarly captures the idiosyncratic risk premium that households require in excess of the one banks require to hold more capital assets. Because banks can eliminate idiosyncratic risk through diversification – represented by the factor $\beta^2$ – this term is typically positive. With regard to the idiosyncratic risk allocation, increasing $\kappa_t$ is then beneficial as it shifts idiosyncratic risk away from households to bankers who have a higher risk-bearing capacity for this type of risk.

Under narrow banking, these two terms are the only terms present in the capital allocation equation (because $\psi = 0$). The scale of banks’ operation $\kappa_t$ therefore balances the costs of higher aggregate risk exposure of bankers with the benefits of a better idiosyncratic risk allocation premium term $\lambda \bar{\kappa}$. Even under narrow banking, fluctuations in $\eta$ and $\kappa$ may affect $\vartheta$ through the idiosyncratic risk premium term in equation (27).
whenever $\kappa_t < \bar{\kappa}$. The determination of $\kappa_t$ is thus solely about the risk allocation and not directly affected by the transaction services provided by money. This is different under fractional reserve banking. Then there is an additional third term in the capital allocation condition, the money premium $\lambda_t v_t$. The reason is that bank deposits serve as inside money for households. Banks can create more of this inside money by taking more assets on their balance sheet. In the aggregate, this requires an increase in $\kappa_t$. The capital allocation equation under fractional reserve banking therefore says that banks increase their excess aggregate risk exposure by more than under narrow banking, not because holding more risky claims helps households managing risk but because these claims can serve as backing for additional money-like deposits.

Combining the observations from Propositions 6 and 7 leads to the important corollary that narrow banking separates banks’ asset choice and households’ transaction demand. The absence of the money premium $\lambda_t v_t$ in equation (30) means that banks’ asset choice is only about the risk allocation, not about households’ transaction demand. This is because banks cannot create transaction media in the form of inside money for households. Transaction demand, in turn, does not directly depend on bank assets ($\kappa_t$) and bank capitalization ($\eta_t$) as the variables $\kappa_t$ and $\eta_t$ do not appear in equation (28). This is the case because changes in the quantity of bank deposits do not affect the total money supply under narrow banking. This separation is in stark contrast to the situation under fractional reserve banking where banks’ asset choice and households’ transaction demand are interdependent: a high money premium induces banks to take on more leverage and thus more aggregate risk (equation (30)), and more leverage in turn reduces velocity (equation (28)) and thereby the money premium.

The model is closed by a set of equations that determine the dynamics of the endogenous state variable $\eta_t$. The risk of this variable, $\sigma^\eta_t$, also appears directly in the aggregate risk premium terms in the money valuation and capital allocation equations. The following proposition describes the evolution of $\eta$.

**Proposition 8 (Equilibrium State Dynamics).** The endogenous state variable $\eta$ follows an Ito process,

$$
\frac{d\eta_t}{\eta_t} = \mu^\eta_t dt + \sigma^\eta_t dZ_t.
$$
and the equilibrium drift and volatility of $\eta$ are given by

$$\mu_t^\eta \eta_t = \eta_t \left( \frac{1 - 2\eta_t}{1 - \eta_t} \sigma_t^\eta - \sigma_t^\theta \right) + \eta_t (1 - \eta_t) \left( \left( \frac{\kappa_t \beta_t}{\eta_t} \right)^2 - \left( \frac{1 - \kappa_t}{1 - \eta_t} \right)^2 \right) (1 - \vartheta_t)^2 \sigma_t^2$$

$$\sigma_t^\eta \eta_t = -\psi \eta_t (1 - \eta_t) \lambda_t v_t,$$

$$\sigma_t^\eta \eta_t = - (\kappa_t - \eta_t) \sigma_t^\theta.$$

The drift equation is here of secondary importance – both to understand model dynamics and to understand the effects of narrow banking – and simply stated for the sake of completeness. The volatility equation says that the volatility of the endogenous state $\eta$ depends on two factors. The first, $\kappa_t - \eta_t$, is the mismatch between the fraction of capital risk held by bankers and their wealth share. When this term is positive, the proportion of total capital risk held by bankers exceeds the proportion of total outside assets they own. As a consequence, bankers are then negatively exposed to increases in $\vartheta_t$ as such increases lower the value of capital assets ($q_t^K$) and increase the value of monetary assets ($q_t^M$). This is the typical situation in equilibrium. The second factor, $\sigma_t^\theta$, is the volatility of the money wealth share $\vartheta_t$. This volatility is caused by portfolio reallocation between capital and money in response to aggregate shocks $dZ_t$ and measures the magnitude of relative price adjustments between the two assets. I present a more explicit expression for this volatility in Section 5.3 when I discuss amplification dynamics.

5.2 Steady State Effects of Narrow Banking: Crowding Out

In this section, I assume that the exogenous risk state $s_t$ is constant and thus idiosyncratic risk $\tilde{\sigma}$ is a time-invariant parameter. Dynamics are then deterministic and converge over time to a steady state. The following proposition states the formal result. In the special case of a cash-in-advance constraint, the steady state can be expressed in closed form which allows for tighter theoretical statements.

**Proposition 9 (Steady State Equilibrium).** Assume that $s_t = \bar{s}$ is constant. Then for each banking policy, there is at most one (monetary) steady state equilibrium with $\eta = \eta^{ss} \in (0, 1)$. Such a steady state always exists under narrow banking or if $\tilde{\sigma}$ is sufficiently large or $\bar{v}$ is sufficiently small. At $\eta^{ss}$, banks intermediate the maximum feasible fraction of capital, $\kappa^{ss} = \ldots$

---

41This can be made consistent with the law of motion (3) by assuming $\sigma_v \equiv 0$, an arbitrary mean-reverting drift function $\mu_s$ (e.g. $\mu_s(s) = b(s - \bar{s})$) and that $s_0$ starts at a steady-state value ($s_0 = \bar{s}$ in the example).

42Under fractional reserve banking with a weak demand for money, both for transaction (large $\bar{v}$) and self-insurance (small $\tilde{\sigma}$) purposes, it is possible that no monetary equilibrium exists because bank-created inside money is sufficient to satisfy all money demand. This situation is not of interest for the present paper.
\( \kappa \).

In the special case \( \alpha = \infty \), the steady state values of \( \eta \) and \( \vartheta \) are given by

\[
\eta^{ss} = \min \{ \eta^I, \eta^M \}, \quad \vartheta^{ss} = \max \{ \vartheta^I, \vartheta^M \},
\]

where

1. under fractional reserve banking,

\[
\begin{align*}
\eta^I &= \frac{\bar{\kappa} \beta}{1 - \bar{\kappa} + \bar{\kappa} \beta}, \\
\vartheta^I &= \left(1 - \bar{\kappa} + \bar{\kappa} \beta\right) \tilde{\sigma} - \sqrt{\rho} \left(1 - \bar{\kappa} + \bar{\kappa} \beta\right) \tilde{\sigma}, \\
\eta^M &= \frac{\bar{\kappa} \beta \tilde{\sigma}}{a \sqrt{\rho} + (1 + \phi a) \left(\bar{\kappa} \beta \tilde{\sigma} + (1 - \bar{\kappa}) \sqrt{\rho}\right) \tilde{b}}, \\
\vartheta^M &= \frac{a \left(1 + \phi \rho\right) - \bar{\kappa} (1 + \phi a) \left(1 - \frac{\beta \tilde{\sigma}}{\sqrt{\rho}}\right) \tilde{b} - \bar{\kappa} \left(1 + \phi a\right) \left(1 - \frac{\beta \tilde{\sigma}}{\sqrt{\rho}}\right) \tilde{b}}{\tilde{b} + a (1 + \phi \tilde{b}) - \bar{\kappa} (1 + \phi a) \left(1 - \frac{\beta \tilde{\sigma}}{\sqrt{\rho}}\right) \tilde{b}}.
\end{align*}
\]

2. under narrow banking,

\[
\begin{align*}
\eta^I &= \frac{\bar{\kappa} \beta}{1 - \bar{\kappa} + \bar{\kappa} \beta}, \\
\vartheta^I &= \left(1 - \bar{\kappa} + \bar{\kappa} \beta\right) \tilde{\sigma} - \sqrt{\rho} \left(1 - \bar{\kappa} + \bar{\kappa} \beta\right) \tilde{\sigma} - \beta \tilde{\sigma} \sqrt{\rho} \bar{\kappa}, \\
\eta^M &= \frac{\bar{\kappa} \beta}{1 - \bar{\kappa} + \bar{\kappa} \beta}, \\
\vartheta^M &= \frac{a \left(1 + \phi \rho\right) - \bar{\kappa} (1 + \phi a) \left(1 - \frac{\beta \tilde{\sigma}}{\sqrt{\rho}}\right) \tilde{b} - \bar{\kappa} \left(1 + \phi a\right) \left(1 - \frac{\beta \tilde{\sigma}}{\sqrt{\rho}}\right) \tilde{b}}{\tilde{b} + a (1 + \phi \tilde{b}) - \bar{\kappa} (1 + \phi a) \left(1 - \frac{\beta \tilde{\sigma}}{\sqrt{\rho}}\right) \tilde{b}}.
\end{align*}
\]

The first part of the proposition implies that \( \kappa^{ss} \) is independent of the banking policy in place. Consequently, narrow banking does not affect the long-run scale of banking activities. This is important because it highlights that narrow banking does not deprive firms of access to bank funding, at least in the long run. The result therefore clarifies that any potential adverse effect of a narrow banking policy on output or growth cannot be caused by an inefficiently low scale of intermediation activities. The argument for \( \kappa^{ss} = \bar{\kappa} \) is as follows. For any given value of \( \eta \) and \( \vartheta \), the capital allocation equation in Proposition 7 implies that for \( \kappa < \bar{\kappa} \), the expected return on capital assets in excess of the money return are equalized across bankers and households. This is a simple consequence of perfect competition and the fact that both agent types are marginal in the risky claims market. But bankers enjoy this (positive) excess return on a larger fraction of their wealth because they hold a more concentrated position in capital assets than households. As a consequence, their net worth grows faster than households’, resulting in an increase in \( \eta \) until bankers’ risk-bearing capacity is sufficiently large to intermediate the maximum amount of capital, \( \kappa = \bar{\kappa} \). This mechanism is at work regardless of banking policy.

The second part of the proposition reports closed-form expressions for the steady-state values
of $\eta$ and $\vartheta$ when transaction demand is derived from a cash-in-advance specification. The result mirrors Proposition 3 for the model without banks. Again, there are two parameter regions, an “I theory” region and a “monetarist” one. In the former, the cash-in-advance constraint is non-binding and thus there is an abundance of transaction media. The resulting steady state values are unaffected by the banking policy – although banking policy does matter for the parameter region for which they are valid – and these expressions coincide with the ones that can be obtained from the Brunnermeier and Sannikov (2016c) model. $\vartheta^I$ is as in the model without banks if idiosyncratic risk $\tilde{\sigma}$ there is replaced with the effective residual risk after partial diversification, $(1 - \bar{\kappa} + \bar{\kappa}\beta)\tilde{\sigma}$. The equilibrium value of bankers’ wealth share $\eta^I$ equals the fraction of idiosyncratic risk that they hold.

The interesting and novel results in this paper concern the latter, monetarist, region in which the cash-in-advance constraint binds and thus the quantity of transaction media matters. Under narrow banking, the equilibrium value of $\vartheta$ is identical to the one obtained in the model without banking, simply because nothing that enters the quantity equation (28) is affected by the presence of banks. In addition, bankers’ equilibrium wealth share is the same in both regions under narrow banking, $\eta^M = \eta^I$.

The situation is more complicated under fractional reserve banking. In the monetarist region, banks’ wealth share must be below the narrow banking counterpart (because $\eta^{ss} = \eta^M < \eta^I$). As banks intermediate the same amount of capital in both banking regimes, the lower bank capitalization under fractional reserve banking is associated with larger leverage. In addition, the value of $\vartheta^M$ is strictly smaller under fractional reserve banking than under narrow banking due to the additional negative term $-\bar{\kappa}(1 + \phi\alpha)(1 - \beta\tilde{\sigma}/\sqrt{\rho})\bar{v}$ in both the numerator and the denominator. In line with the discussion following Proposition 1 in Section 4 and Proposition 4 in Section 5.1, this implies that there must be more investment crowding out and less growth under narrow banking.

While there is no closed-form solution for the steady state in the more general transaction cost formulation ($\alpha < \infty$), the effects of narrow banking on the steady state just discussed remain valid more generally. The following result summarizes the steady-state effects of narrow banking. For completeness, it restates the result that $\kappa^{ss}$ is unaffected by the banking policy.

**Proposition 10** (Steady-state Effects of Narrow Banking). *Assume that $s_t = \bar{s}$ is constant and $\alpha > 2$.*\(^{43}\) Then implementing a narrow banking policy has the following steady state effects:

\(^{43}\)The assumption $\alpha > 2$ is a crude, but simple sufficient condition that is used in the proof that narrow banking increases $\vartheta$. A closer inspection of the argument reveals that the results of this proposition may remain valid for values of $\alpha$ well below 2.
1. The scale of banking activity $\kappa$ is identical under narrow banking and fractional reserve banking.

2. If $\alpha = \infty$ and $\tilde{\sigma} \geq \tilde{\sigma}^I := \frac{\tilde{v} + a(1 + \phi \tilde{v})}{\tilde{v} + a(\tilde{v} - \rho) 1 - \kappa + \kappa^2}$, transaction media are abundant and narrow banking has no effect in equilibrium.

3. If $\alpha < \infty$ or $\tilde{\sigma} < \tilde{\sigma}^I$, $\eta^{ss}$ and $\vartheta^{ss}$ are strictly larger under narrow banking than under fractional reserve banking. As a consequence,

   (a) narrow banking reduces bank leverage and
   (b) narrow banking reduces economic growth.

The first two results stated in the proposition have already been discussed above. The economic intuition behind the third result is as follows.

The larger $\eta^{ss}$ is a result of the higher financing rate banks face under narrow banking. Under fractional reserve banking, banks can fund themselves at the (low) money rate and are thus willing to expand their balance sheets to a point where their required risk premium for holding capital assets is larger than households' (compare the capital allocation equation (30) with $\psi = 0$). In contrast, under narrow banking, banks must fund themselves at the higher illiquid deposit rate and thus only expand their balance sheets to the point where the capital risk premium required by households and bankers is equalized (equation (30) with $\psi = 0$). For banks to hold the same fraction $\tilde{\kappa}$ of capital risk in steady state, their net worth share $\eta$ must thus be larger under narrow banking.

The larger $\vartheta^{ss}$ is a result of the absence of inside money under narrow banking. Specifically, by the quantity equation (28), steady-state velocity under fractional reserve banking is

$$\psi^{ss} = \frac{\alpha}{1 + \phi a} \left( 1 - \vartheta^{ss} + \phi \rho \right) \left\{ \frac{\vartheta^{ss}}{\text{outside money}} + \frac{\tilde{\kappa} (1 - \vartheta^{ss}) - \eta^{ss}}{\text{inside money}} \right\}$$

where the denominator is the sum of the value of outside money $\vartheta^{ss}$ and the value of bank deposits $\tilde{\kappa} (1 - \vartheta^{ss}) - \eta^{ss}$, which represent inside money, both as a fraction of total wealth. Under narrow banking, deposits can no longer be used for transactions and thus the (positive) inside money term disappears, which increases velocity for any given value of outside money $\vartheta^{ss}$. This raises the equilibrium money premium (condition (29)) and – through the money valuation equation (27) – the steady-state value of $\vartheta^{ss}$. Intuitively, the outside money bubble with value $\vartheta^{ss}$ is the only form of money supply under narrow banking, whereas under fractional
reserve banking a substantial fraction of the money supply may take the form of inside money that ultimately derives its value from intermediated capital (whose value is $\bar{\kappa} (1 - \vartheta^{ss})$). The value of the outside money bubble $\vartheta^{ss}$ must therefore be larger under narrow banking to meet any given demand for money as a medium of exchange. This is the most important steady-state effect of narrow banking.

5.3 Dynamic Effects of Narrow Banking: Stabilization

In this section, I analyze the stochastic economy. In the process, I discuss how the financial system under fractional reserve banking amplifies portfolio reallocation between capital assets and monetary assets in response to changes in risk premia. The idiosyncratic risk shocks in my model should here be thought of as a simple device to generate such risk premium variations. Other devices such as time-varying risk aversion would lead to similar conclusions. I then proceed by showing theoretically how narrow banking can fully eliminate amplification and thereby stabilize the economy in the special case $\alpha = \infty$. I conclude with the presentation of a numerically solved calibrated example for the more realistic case $\alpha < \infty$.

Theoretical Results. In the model, shocks affect allocations only if they affect the wealth share of bankers $\eta$ or if they lead to a revaluation of assets by affecting the share of money wealth $\vartheta$. By Proposition 8, the effects of a shock on both variables are linked, $\sigma_t^\eta \eta_t = - (\kappa_t - \eta_t) \sigma_t^\vartheta$. As discussed following that proposition, bank balance sheets are composed of capital assets and monetary liabilities when $\kappa_t > \eta_t$, so that an increase in $\vartheta$ hurts bankers as it simultaneously decreases the value of their assets and increases the value of their liabilities. But as a change in $\vartheta$ affects $\eta$ and thereby aggregate bank capitalization, banks react and this may amplify or mitigate the initial effects on $\vartheta$. The following proposition restates the risk of the endogenous state variable $\eta$ in a more explicit form that highlights the forces behind shock amplification and mitigation.

Proposition 11 (Risk Generation Equation). In equilibrium, the risk of the endogenous state variable $\eta$ is given by

$$\sigma_t^\eta \eta_t = \frac{(\kappa_t - \eta_t) \frac{\partial \vartheta(s_t, \eta_t)}{\partial(s_t, \eta_t)}}{1 + (\kappa_t - \eta_t) \frac{\partial \vartheta(s_t, \eta_t)}{\partial(s_t, \eta_t)}} \sigma_s(s_t, \eta_t).$$

44 More precisely, this is only true when $\kappa (1 - \vartheta) > \eta$. If $\kappa > \eta \geq \kappa (1 - \vartheta)$, banks do not have monetary liabilities, but hold outside money. They merely hold a larger proportion of their wealth in capital assets than households. I emphasize in my discussion the case of monetary liabilities because this is the typical case under Assumption 2.
The terms in this risk generation formula have simple interpretations. The factor \( \sigma_s(\eta_t, s_t) \) is the marginal increase in the risk state \( s_t \) for a negative shock (\( dZ_t < 0 \)). The direct effect of this increase in \( s_t \) (and thus in risk \( \bar{\sigma}_t = \bar{\sigma}(s_t) \)) is a rise in the idiosyncratic risk premium that leads agents in the economy to reallocate their portfolios from capital to monetary assets, even for fixed \( \eta \) – compare the money valuation equation (27). As a result, the nominal wealth share \( \vartheta_t \) changes by \( \partial_s \vartheta/\vartheta \). The associated asset price adjustment affects \( \eta \) because of the mismatch on banks’ balance sheets that is increasing in the difference \( \kappa - \eta \). The numerator of the risk generation equation therefore captures the marginal direct effect of a shock on \( \eta_t \) in the absence of amplification caused by banks’ reaction to the shock.

The denominator captures equilibrium amplification. After a negative shocks, \( \eta_t \) falls and bankers find themselves with a higher than desired ex post leverage and thus higher net worth risk exposure. If leverage is sufficiently high, i.e., for low \( \eta_t \), bankers respond to this situation by deleveraging. In the process, banks shrink their balance sheet which reduces both their short position in (nominal) deposits and their long position in (real) risky claims, thereby triggering a further round of aggregate portfolio reallocation from capital assets to money. The overall equilibrium effect of this response on \( \vartheta \) is \( \partial_\eta \vartheta/\vartheta \), which is a negative number because more bank losses are associated with a stronger reallocation to money assets.\footnote{This argument presumes that either \( \eta \) is actually low enough that banks respond by shrinking their balance sheet or the prospects of moving \( \eta \) closer to such a region lead to a sufficiently large appreciation in \( \vartheta \). Under fractional reserve banking, there is also a second, offsetting, effect if this is not the case: a reduction in \( \eta \) with no simultaneous reduction in \( \kappa(1 - \vartheta) \) increases the real value of inside money and this tends to lower \( \vartheta \), thereby mitigating the effect of the shock. I do not focus on this channel in my discussion because it is quantitatively weak in the ergodic distribution of my calibrated example below.} As for the direct effect, the price adjustments associated with the change in \( \vartheta \) translate into further losses for banks, whose magnitude depend on the magnitude of the balance sheet mismatch, \( \kappa - \eta \). In total, the denominator is typically smaller than 1 and therefore scales up the total risk of \( \eta \). In this sense, it captures amplification.

The risk generation equation is useful in that it clarifies that the derivatives \( \partial_s \vartheta/\vartheta \) and \( \partial_\eta \vartheta/\vartheta \) are the right measures for endogenous risk and amplification. But it is not directly helpful in understanding how narrow banking affects endogenous risk. Indeed, the same equation holds in identical form for both banking regimes and it also holds in an “I theory version” of the model in which there is no demand for money as a medium of exchange at all (formally either \( \alpha \to 1 \) or \( \bar{v} \to \infty \)) so that the distinction between fractional reserve banking and narrow banking becomes irrelevant. What matters for understanding the difference between these cases is therefore the shape of the \( \vartheta \) functions, which in turn depends on how the sum of the three premia in the money valuation equation (27) changes over the state space. The aggregate risk premium term in that
equation depends itself on $\sigma^\eta \eta$, which simply means that amplification can be self-reinforcing. For understanding the original sources of endogenous risk, one must therefore understand the two other premia in the equation, the idiosyncratic risk premium and the money premium.

Let me first assume an “I theory” world as in Brunnermeier and Sannikov (2016c) in which the money premium is absent. Then $\partial_s \theta/\theta > 0$ because an increase in $s_t$ increases the idiosyncratic risk premium term

$$\left( \frac{(1 - \kappa_t)^2}{1 - \eta} + \frac{\kappa_t^2 \beta^2}{\eta} \right) (1 - \theta_t)^2 \tilde{\sigma}_t^2$$

(31)

for any given values of $\kappa_t$ and $\eta_t$. When the initial impact of the shock generates losses for banks, a liquidity spiral amplifies these initial losses. Specifically, bankers fire-sell capital assets (risky claims) and simultaneously shrink deposits. These fire sales increase the fraction of capital risk that households have to hold directly, $1 - \kappa_t$, and thereby increase the total undiversified idiosyncratic capital risk, $(\kappa_t \beta + 1 - \kappa_t) s_t$, in the economy. This has two implications. First, households are only willing to hold the additional capital risk if the capital price $q_t^K$ falls. Thus, fire sales depress the capital price which generates additional losses on the asset side of banks’ balance sheets. Second, households end up holding more idiosyncratic risk and therefore demand more money for self-insurance purposes, which raises the value of outside money $q_t^M$. Thus, fire sales also increase the real value of bank deposits which generates even more losses on the liability side of banks’ balance sheets. The combined effect of both implications is reflected formally in the first factor of the idiosyncratic risk premium expression (31). After the shock, $\kappa_t$ and $\eta_t$ have simultaneously decreased in a way that makes this first factor larger.

With a transaction demand for money, an additional money premium term $\lambda_t \theta_t$ appears in the money valuation equation and it is the combined premium that matters for portfolio reallocation and thus the derivatives $\partial_s \theta/\theta$ and $\partial_\eta \theta/\theta$. The presence of this money premium term in itself is stabilizing as has been explained in the context of the model without banks. Both the direct effect of the initial shock and the indirect effect from amplification increase money demand for reasons other than the need to make transactions – as a store of value to hedge against aggregate risk or to self-insure against idiosyncratic wealth fluctuations. As agents want to hold more money for these other reasons, the opportunity cost of holding transaction media decreases. This means that $\lambda_t \theta_t$ must fall, which affects the total premium in the money valuation equation in the opposite direction as the effect of the shock and of amplification on the idiosyncratic risk premium. As a consequence, the presence of a transaction demand for

---

46 Alternatively, they may leave the quantity of deposits unaltered and replace sold capital assets with outside money reserves. This is economically equivalent in my model.
money in itself tends to mitigate both the initial effect of the shock – resulting in a lower $\partial_s \vartheta / \vartheta$ – and the amplification from banks’ reaction to the shock – resulting in a larger $\partial \eta \vartheta / \vartheta$. Both effects lower endogenous risk $\sigma^\rho \eta$ relative to an “I theory” world.

Under fractional reserve banking, however, the discussion in the last paragraph is incomplete, as it ignores how banks’ reaction to shocks affects the supply of transaction media. Specifically, as banks shrink their balance sheet, they reduce the total supply of inside money, which makes transaction media scarcer. The stabilizing reduction of the money premium $\lambda_t \vartheta_t$ may then not materialize. Instead, movements in the money premium may even become destabilizing: even though disinflationary pressure (an increase in $\vartheta_t$) tends to lower the money premium, the simultaneous losses it generates for banks may cause a Fisherian disinflationary spiral: a higher real value of money leads to larger bank losses and banks contract their balance sheets by so much that the real supply of transaction media shrinks despite the disinflation, thereby increasing the demand for outside money for transaction purposes, which increases the real value of money even more.\(^\text{47,48}\) This is the source of instability caused by fractional reserve banking that narrow banking advocates such as Fisher are concerned with (e.g. Fisher, 1935, Chapter VII).

This discussion highlights how narrow banking attempts to stabilize the financial system. By breaking the link between bank deleveraging and the money supply, narrow banking eliminates the disinflationary spiral and allows the automatic stabilization mechanism present in the model without banks to operate. While a rise in risk premia and bank deleveraging may still cause disinflationary pressure that hurts banks, this disinflationary pressure simultaneously lowers the money premium, so that the total premium in the money valuation equation (27) is smaller. Thus, portfolio reallocation between capital assets and money remains limited. As stated previously, the reduced desire to adjust portfolios decreases not only amplification (less negative $\partial \eta \vartheta / \vartheta$), but also the initial direct effect of the shock (lower $\partial_s \vartheta / \vartheta$).

\(^{47}\)Note that Brunnermeier and Sannikov (2016c) use a different terminology and use the label “disinflationary spiral” in reference to the disinflation resulting from capital fire sales and the induced reallocation of idiosyncratic risk, whereas they limit the term “liquidity spiral” to the effects of the fire sales on the capital price $q^K_t$. I find that usage misleading, at least in reference to Fisher (1933), because in their framework it is not the reduction of the inside money supply in itself that causes the disinflation, but the increase in undiversified idiosyncratic capital risk due to a reduction in the share of capital intermediated by banks ($\kappa_t$).

\(^{48}\)Nothing definite can be said in general with regard to the question whether amplification under fractional reserve banking with a transaction demand for money is larger or smaller than in an otherwise identical model without a transaction demand (with only the mechanism emphasized by Brunnermeier and Sannikov (2016c) present). This depends on parameters and the question whether endogenous reactions of the inside money supply make the money premium positively correlated with the two risk premia or merely less negatively correlated than under narrow banking. In the context of my calibrated model, amplification is indeed larger than without a transaction demand.
The ability of narrow banking to limit even the direct effect of the shock is starkest in the case of a cash-in-advance specification for the transaction demand. Then, the quantity equation (28) implies a one-to-one mapping between velocity and \( \vartheta \) that is not affected by bank balance sheet variables \( \eta \) and \( \kappa \), so that whenever velocity is at its upper bound, \( \vartheta \) remains insulated from shocks and instead movements in the money premium fully absorb changes in risk premia.

This reasoning implies the following theoretical result.

**Proposition 12 (Stability under Narrow Banking).** Suppose \( \alpha = \infty \) and a narrow banking policy is in place. Then any equilibrium has the following properties:

1. **Local stability:**
   
   For any state \((s_0, \eta_0)\) such that \( \lambda(s, \eta) > 0 \) in a neighborhood of \((s_0, \eta_0)\) (i.e., the cash-in-advance constraint binds locally), \( \vartheta(s_0, \eta_0) = \vartheta^M \), \( \vartheta \) is locally constant around \((s_0, \eta_0)\) (i.e., \( \mu^\vartheta(s_0, \eta_0) = \sigma^\vartheta(s_0, \eta_0) = 0 \)) and the endogenous state \( \eta \) is locally deterministic (\( \sigma^\eta(s_0, \eta_0) = 0 \)) and drifts towards \( \eta^M \).

2. **Global stability:**
   
   If the exogenous process for \( s_t \) is such that \( \mathbb{P}(\bar{\sigma}(s_t) \leq \bar{\sigma}^I) = 1 \), then \( \vartheta_t = \vartheta^M \) and \( \eta_t = \eta^{ss} \) are constant in equilibrium.

Here, \( \vartheta^M \) and \( \eta^M \) are the “monetarist” steady state values as stated in Proposition 9 and \( \bar{\sigma}^I \) is the threshold defined in Proposition 10.

The local stability result exactly formalizes the idea that whenever the cash-in-advance constraint is binding, the quantity equation determines \( \vartheta \) and the money premium adjusts to make the money valuation equation hold, i.e., to stabilize the value of \( \vartheta \) at the level consistent with the quantity equation. This immediately implies that \( \vartheta \) must be locally constant around such states. By the risk generation equation, fluctuations in the endogenous state \( \eta \) are caused by relative price fluctuations between capital assets and monetary assets. A constant \( \vartheta \) means a stable relative price and thus \( \eta \) must evolve (locally) deterministically.

The global stability result is more restrictive in nature. The additional condition \( \mathbb{P}(\bar{\sigma}(s_t) \leq \bar{\sigma}^I) = 1 \) ensures that the cash-in-advance constraint is always binding and thus the local result applies globally. If this is the case, then the narrow banking economy remains in steady state, even though idiosyncratic risk and its associated risk premium fluctuate over time. The additional condition is certainly restrictive, but there is no discontinuity between exogenous state processes with \( \mathbb{P}(\bar{\sigma}(s_t) \leq \bar{\sigma}^I) = 1 \) and those with \( \mathbb{P}(\bar{\sigma}(s_t) \leq \bar{\sigma}^I) < 1 \). If idiosyncratic risk is sometimes
larger and sometimes smaller than $\tilde{\sigma}'$, then a picture similar to the red solid line in Figure 2 in the model without banks emerges: for low $\tilde{\sigma}$ states, the local stability result is valid, $\vartheta$ is constant at $\vartheta^M$ and $\eta$ mean-reverts to $\eta^M$, for large $\tilde{\sigma}$ states, $\vartheta$ is slightly increasing in $s$ and thus endogenous risk can be positive.

**Calibration and Numerical Solution Procedure.** I next turn to a quantitative example for the case that the conditions of the global stability result in Proposition 12 are violated. In this case, the model does not have a closed-form solution anymore. However, one can use standard procedures to transform the forward-backward equation system defined by the equations in Propositions 5 and 8 into a second-order partial differential equation for the function $\vartheta(s, \eta)$. I solve this equation with a finite-difference time stepping procedure. Monotonicity of the scheme is ensured by the use of upwind differences for the first-order derivatives and a customized algorithm based on coordinate rotation and interpolation for the second-derivatives. Details are presented in Appendix B.2.

I summarize my calibration strategy for the model parameters only briefly and relegate a detailed discussion to Appendix C. For the calibration, I use a slightly more general model version than presented so far to allow for the existence of non-monetary government debt. I show in Section 7.1 that this generalized model can be mapped into the model discussed so far by adjusting the parameter $\bar{v}$ and using $\psi > 1$ instead of $\psi = 1$ under fractional reserve banking. Consequently, all the intuition and the model equations discussed previously remain valid.

A time period in the calibrated model represents one year. For the exogenous volatility process, I assume again a Heston model, compare equation (26). I calibrate the parameters of the volatility process $b$, $\sigma_s$ and $\bar{s}$ to match the evidence on establishment-level TFP shocks presented in Bloom et al. (2018). I calibrate the remaining model parameters so that the model is largely consistent with empirical observations in U.S. data that inform the quantitative strength of my mechanisms. Because the computation time of a model solution is substantial (several hours), a calibration to target moments of the stochastic economy is not feasible and I resort to a calibration of the steady state instead. I calibrate $\alpha$, so that the interest-elasticity of money demand in the model is consistent with the empirical estimates. I calibrate $\rho$, $\beta$ and $\bar{\kappa}$ jointly to be consistent with the leverage of depository institutions, the money multiplier and the average share of government-provided safe assets in the portfolio of private agents. I choose $\psi$ in line with the transformation stated in Section 7.1 to match the share of government-provided safe assets that are non-monetary. I calibrate $\bar{\vartheta}$, so that the steady-state money premium is in line with estimates in the literature. Finally, I calibrate $a$ to match the empirical investment-output
ratio, while I follow Di Tella and Hall (2020) in their calibration of the capital adjustment cost \( \phi \). The resulting parameter values are \( a = 0.0673, \rho = 0.047, \phi = 5, b = 0.15, \sigma_s = 0.037, \bar{s} = 0.085, \beta = 0.062, \bar{\kappa} = 0.244, \alpha = 10, \bar{v} = 0.235, \psi = 1.539 \).

**Numerical Results.** All equilibrium quantities depend on the two-dimensional state vector \((s, \eta)\), so that one could in principle illustrate global equilibrium dynamics by plotting the full equilibrium functions. As three-dimensional plots are hard to read, I plot instead sections of equilibrium functions as a function of the endogenous state \( \eta \) for fixed values of \( s \). It turns out that most equilibrium functions are monotonic either directly in \( s \) or along curves in the state space that are very close to being lines parallel to the \( s \) axis. Showing for each variable an equilibrium function for high and for low \( s \) gives therefore a sufficiently complete picture of the full equilibrium dynamics.

Figures 3 depicts the equilibrium functions for the money value \( q^M \), the capital price \( q^K \), the share of capital intermediated by banks \( \kappa \), and the money premium \( \lambda v \) as a function of \( \eta \) and Figure 4 depicts the dynamics of the state variable \( \eta \) itself by plotting its (arithmetic) drift \( \mu^\eta \eta \) and volatility \( \sigma^\eta \eta \). The blue lines show the equilibrium under fractional reserve banking whereas the red lines show the equilibrium under narrow banking. In both cases, the darker color represents the low risk state (low \( s \)) and the lighter color the high risk state (high \( s \)). I first discuss equilibrium dynamics under fractional reserve banking.

The top panels of Figure 3 show the equilibrium price functions \( q^M \) and \( q^K \). In line with the economic logic discussed above, \( q^M \) is increasing and \( q^K \) is decreasing in the risk state \( s \). This implies that an increase in risk reduces bank capitalization \( \eta \). As a function of \( \eta \), \( q^M \) is decreasing for low \( \eta \) and increasing for large \( \eta \), whereas the opposite pattern is true for \( q^K \). The reason for the difference between low and large \( \eta \) is visible in the bottom left panel of Figure 3. For low \( \eta \), the fraction of capital intermediated by banks is increasing in \( \eta \). In this region, a reduction in bank capitalization leads to deleveraging in line with the discussion above. Due to \( \partial_\eta q^M < 0, \partial_\eta q^K > 0 \), price movements then amplify endogenous risk. For large \( \eta \), banks intermediate the maximum possible quantity of capital, \( \kappa = \bar{\kappa} \). In this region, asset price movements mitigate instead of amplify shocks. I have not focused on this possibility in

---

49 Note that the calibration of \( a \) and \( \phi \) does not affect the dynamics of the state variable \( \eta \) and the solution function \( \vartheta(s, \eta) \). In this sense, the choice of these parameters is less important than the choice of the remaining parameters. The two parameters do affect my welfare results presented in the next section, however.

50 The model also features an additional parameter, the depreciation rate \( \delta \) which I set to be consistent with the average GDP growth rate. However, \( \delta \) does not enter any of the equilibrium conditions presented in Section 5.1 and therefore its value is irrelevant for the following results.

51 The reason is that for \( \partial_\eta \kappa = 0 \), a decrease in \( \eta \) implies an increase in bank leverage and thus an increase in
Figure 3: Asset prices $q^M$, $q^K$, capital allocation $\kappa$, and money premium $\lambda v$ under a fractional reserve banking (blue) and narrow banking (red) policy. The dark blue and dark red lines depict the respective equilibrium function at a low risk state (low $s_t$), the light blue and light red lines depict the respective equilibrium function at a high risk state (high $s_t$).

Figure 4: Volatility of $\eta$.

my discussion above, because, as will become clear below, the system spends most of its time at $\eta$ values in or close to the region where $\kappa < \bar{\kappa}$.

Equilibrium amplification is also visible in the right panel of Figure 4, which plots the volatility of $\eta$. In the low $\eta$ region, volatility of the state variable is large, both relative to high $\eta$ states and relative to the drift of $\eta$, which is depicted in the left panel of the same figure. The latter observation is important. Whenever the absolute value of the drift is large relative to the volatility, drift dynamics dominate and the system almost deterministically moves in the quantity of deposits, see footnote 45. This lowers the money premium as visible in the bottom right panel of Figure 3 and thereby mitigates the impact on portfolio reallocation.
direction of the drift. This is the case for large $\eta$ above a value of around 0.09, when $\mu^\eta \eta$ is strongly negative and $\sigma^\eta \eta$ is small. Then the system quickly drifts back to lower $\eta$ values so that these states are rarely visited in equilibrium. Instead, in the low $\eta$ region where $\kappa < \bar{\kappa}$, the drift is positive (consistent with the steady state results in Section 5.2), yet the volatility is quantitatively large so that occasionally a sequence of bad shocks pushes $\eta$ very close to zero.

The bottom right panel of Figure 3, which depicts the money premium $\lambda v$, confirms the economic logic behind amplification under fractional reserve banking. First, when the risk state $s$ rises, then for any given $\eta$ the money premium falls. This itself is a stabilizing force. But second, as a function of $\eta$, the money premium steeply decreases in the region where $\kappa < \bar{\kappa}$. Because a risk shock in this region of the state space reduces $\eta$, the money premium falls by less than for fixed $\eta$ or even rises after a shock that increases risk.

The red lines in Figures 3 and 4 depict the equilibrium under a narrow banking policy. As is visible in the top panels of Figure 3, $q^M$ is uniformly larger and $q^K$ is uniformly smaller under narrow banking. This is consistent with the steady-state results from Section 5.2, that narrow banking raises the value of outside money and crowds out investment. In addition, $q^M$ and $q^K$
are almost constant as a function of $\eta$. As a consequence, there is almost no amplification and almost no endogenous risk under narrow banking. This is clearly visible in the right panel of Figure 4. Volatility $\sigma^\eta \eta$ is uniformly lower than under fractional reserve banking and no longer rises sharply when $\eta$ becomes small.

A consequence of the increased stability under narrow banking is that banks almost never shrink their balance sheet in response to negative shocks as can be seen in the bottom left panel of Figure 3. The $\kappa$ function is nearly insensitive to the risk state $s$ itself and remains constant in $\eta$ except for very low levels of $\eta$. This result may look surprising in light of Proposition 10 that in steady state bank leverage is lower under narrow banking. The difference between that result and the situation here is that in steady state there is no endogenous risk, $\sigma^\eta \eta = 0$, in either banking regime whereas here there is substantial endogenous risk under fractional reserve banking, which through the capital allocation equation (30) translates into a lower equilibrium value for $\kappa$.

Again, the bottom right panel of Figure 3 confirms that also the increased stability under narrow banking is consistent with the intuition given above. Now – as under fractional reserve banking – the money premium is strictly decreasing in the risk state $s_t$, which is stabilizing. But it is no longer the case that the money premium is decreasing in $\eta$, the destabilizing force under fractional reserve banking. Instead, the money premium is largely constant in $\eta$.

The differences in the dynamics of the endogenous state variable $\eta$ under fractional reserve and narrow banking visible in Figure 4 translate into differences in the ergodic state distribution for the two banking policies. This is depicted in Figure 5, which plots the conditional ergodic densities of $\eta$, conditional on a fixed risk state $s_t$. As before, I plot for each banking policies two functions, one at a low risk state (dark lines) and one at a high risk state (light lines). Under both banking policies, the $\eta$ distribution conditional on the high risk state is shifted to the left. This is mechanical: shocks that increase $s_t$ hurt bankers ($\sigma^\eta \eta$ is everywhere positive), thus conditional on a high level of $s_t$ there have been many of these shocks in the recent history. The difference is, however, by how much the $\eta$ distribution is shifted to the left for large $s_t$. Under fractional reserve banking, a substantial fraction of the (conditional) probability mass is to the left of the $\eta$ level below which bankers no longer intermediate the maximum fraction of capital ($\kappa < \bar{\kappa}$), compare the bottom left panel of Figure 3. Under narrow banking, in contrast, the distribution at the high level of $s_t$ is only slightly shifted to the left and only a negligible fraction of the probability mass is in the region where $\kappa < \bar{\kappa}$. Narrow banking therefore not

\footnote{It is even mildly increasing, reflecting the fact that for low $\eta$, money demand for self-insurance purposes slightly increases and the resulting (small) increase in the real value of money makes transaction media less scarce.}
only stabilizes the system in the sense that it reduces endogenous volatility (lower $\sigma_\eta^2$ and $\sigma_\theta^2$), but also in the sense that it reduces the amount of time spent in a “crisis state” where $\kappa < \bar{\kappa}$.

Overall, the observations from this numerical example confirm the theoretical results and economic intuition discussed earlier. Fractional reserve banking is associated with a destabilizing disinflationary spiral that amplifies the effects of changes in risk premia on banks’ balance sheets and asset prices. A narrow banking policy stabilizes the financial system by eliminating this disinflationary spiral. Then, the money premium acts as a stabilizing force that offsets changes in risk premia and thereby limits the portfolio reallocation caused by them. However, a narrow banking policy also restricts the supply of transaction media which increases the total value of outside money and leads to more investment crowding out and lower economic growth.

How large are these effects quantitatively? My calibrated model simulation suggests that average growth rate, $\Phi(t) - \delta$ averaged over the ergodic distribution, falls by 0.95% under narrow banking. I compute two measures of the stabilizing effect of narrow banking. First, I compare
the average volatility of the relative capital value $q^K/q^M$ across the policies. Under fractional reserve banking, this average volatility is 4.7%, while under narrow banking it falls to 1.1%, almost a fifth of the volatility under fractional reserve banking. As a second measure, I compute the fraction of time the economic system spend in the region where banks do not intermediate the maximum quantity of capital, that is $\kappa < \bar{\kappa}$. Under fractional reserve banking, this fraction is 6.6%, whereas under narrow banking it drops to below 0.1%.

6 Welfare

In this section, I complement the positive analysis in the previous section with a normative assessment of narrow banking. The results from Section 5 seem to imply a policy trade-off between financial stability and economic growth. I show that improved financial stability under narrow banking is indeed always welfare-improving. A policy trade-off emerges, if the negative growth effects of narrow banking are welfare-reducing. The latter may not always be the case because a larger value of money does not only reduce growth but also improves idiosyncratic risk sharing.

I start with a representation of expected utility of an individual agent in the model. In the previous sections, I have suppressed explicit references to individual agents within each type group to avoid cluttering notation with uninformative indices or function arguments. In the following, it serves clarity to introduce such explicit references. To do so, I assume that the interval $\mathbb{I} := [0, 1]$ represents the continuum of all agents. I refer to an individual agent by $\tilde{i} \in \mathbb{I}$ and to the type this agent belongs to by $t(\tilde{i}) \in \{h, b\}$. I attach the individual index $\tilde{i}$ to variables not as an additional superscript but as a function argument, e.g., $c(\tilde{i})$ and $c^{t(\tilde{i})}(\tilde{i})$ both refers to the consumption process of agent $\tilde{i}$.

**Proposition 13.** The expected continuation utility of individual $\tilde{i} \in \mathbb{I}$ at the state $(s_0, \eta_0, K_0)$ and the initial wealth distribution $(n_0(\tilde{i}))_{\tilde{i} \in \mathbb{I}}$ is given by

$$V_0(\tilde{i}) := \mathbb{E}\left[ \int_0^\infty e^{-\rho t} \log c_t(\tilde{i}) \mid s_0, \eta_0, K_0, (n_0(\tilde{i}))_{\tilde{i} \in \mathbb{I}} \right]$$

$$= \frac{\log \tilde{\eta}_0(\tilde{i}) + \log K_0}{\rho} + v^{t(\tilde{i})}(s_0, \eta_0)$$
where for $i \in \{h, b\}$,

$$v^i(s_0, \eta_0) := \mathbb{E} \left[ \int_0^\infty e^{-\rho t} \left( \log (a - \iota_t) + \frac{\Phi(t) - \delta}{\rho} \right) dt \mid s_0, \eta_0 \right]$$

$$+ \mathbb{E} \left[ \int_0^\infty e^{-\rho t} \left( \frac{\mu^\eta,i^i}{\rho} - \frac{(\sigma^\eta,i^i)^2}{2\rho} \right) dt \mid s_0, \eta_0 \right]$$

$$- \mathbb{E} \left[ \int_0^\infty e^{-\rho t} \frac{(\tilde{\sigma}^\eta,i^i)^2}{2\rho} dt \mid s_0, \eta_0 \right]$$

(32)

and $\tilde{\eta}_t(\tilde{i}) := \frac{n_t(i)}{N^h_t + N^b_t}$ is the wealth share of individual $\tilde{i}$ as a fraction of total wealth, $\eta_t^i := \frac{N^i_t}{N^h_t + N^b_t}$ is the wealth share of type $i$ across types, $\mu^\eta,i^i$ and $\sigma^\eta,i^i$ are the geometric drift and volatility of $\eta^i_t$ and $\tilde{\sigma}^\eta,i^i$ is the common geometric idiosyncratic volatility of the individual’s net worth $n_t(\tilde{i})$ for all individuals $\tilde{i}$ of type $t(\tilde{i}) = i$. The latter is explicitly given by

$$\tilde{\sigma}^\eta,i^i = \begin{cases} 
\frac{1 - \kappa_t}{1 - \phi_t} (1 - \vartheta_t) \tilde{\sigma}_t, & i = h \\
\frac{\kappa_t}{\eta_t} (1 - \vartheta_t) \beta \tilde{\sigma}_t, & i = b
\end{cases}.$$ 

The proposition decomposes expected utility into a term that depends only on the initial conditions for the wealth distribution across individuals and for aggregate capital and into a term that depends only on the type of the agent and the component $(s_0, \eta_0)$ of the aggregate state vector. To assess the welfare effects of narrow banking (or of any policy that does not affect initial conditions), only the second term is relevant. The proposition says further that this second term $v^i(s_0, \eta_0)$ can itself be decomposed into three terms. The first captures the aggregate consumption-savings trade-off of a hypothetical representative agent. Under complete markets, this would be the only term and thus the welfare effect of any policy would be fully described by the deviation of the investment rate $\iota_t$ from its first-best level $\iota^{FB}_t = \frac{a - \rho}{1 + \phi}$. In my incomplete markets model, it is always true that $\iota_t < \iota^{FB}_t$ in any monetary equilibrium, yet this may not be inefficient, because $\iota_t$ is directly related to $\vartheta_t$ which in turn affects the two other terms in the decomposition.

The second term in $v^i(s_0, \eta_0)$ captures future redistribution across types through changes in the endogenous state variable $\eta$. Expected changes are summarized by the drift term $\mu^\eta,i^i$. These can both increase and decrease utility relative to the first term in the decomposition.
depending on whether agents of type $i$ expect on average to become poorer or richer relative to the other agent type. The risk in $\eta$ is summarized by the volatility term $\sigma^{\eta,i}_t$ and always affects utility negatively. Finally, the third term in the decomposition of $v^i(s_0, \eta_0)$ captures idiosyncratic risk exposure. Its impact on utility is always negative.

In my heterogeneous agent economy, assessing normative implications of a policy based on a fixed social welfare function with given welfare weights evaluated at a given initial state faces the issue that the quantitatively largest effects of the policy can easily be redistributive ones: if narrow banking happens to decrease the average level of $\eta$ and the proportion of bankers and households are kept fixed across banking regimes, then bankers tend to be worse off and households tend to be better off. If this redistributive effect is strong, then a welfare function that puts a lot of weight on bankers will favor fractional reserve banking and a welfare function that puts a lot of weight on households will favor narrow banking, irrespective of any growth or risk implications of narrow banking. To abstract from such redistributive effects, I consider the experiment of a hypothetical planner who can intervene at the initial time ($t = 0$) by distributing wealth arbitrarily across agents and by assigning each agent $\tilde{i} \in \mathbb{I}$ a type $(\tilde{i}) \in \{h, b\}$ that the agent then keeps forever.\footnote{This experiment is similar to the one Di Tella (2017) considers. However, in that paper, the planner is only allowed to redistribute wealth while the type assignments remain fixed whereas I also allow types to be adjusted. My reasoning for this is that types are fixed in the model for simplicity, but in reality they would likely adjust to the long-run redistributive effects of the policy. If a policy increased the need for bank capital in the ergodic distribution (higher average $\eta$), then likely more people would start providing such capital and thereby become “bankers” so that eventually the share of bankers in the population would increase in line with the wealth share of bankers. Allowing the planner only to redistribute initial wealth could have the undesirable side effect that the planner may instead want to starve bankers of wealth initially (and potentially move to a “crisis state”) in order to balance the fact that they become richer over time relative to households.}

The planner has to make these initial assignments in a way that respects a given distribution (density) $(u(\tilde{i}))_{\tilde{i} \in \mathbb{I}}$ of relative utilities with the property $\int_{\mathbb{I}} u(\tilde{i}) d\tilde{i} = 1$. Specifically, after the planner has made initial assignments, the resulting initial expected utilities must have the property\footnote{Instead of the value functions $V^i_0(\tilde{i})$, I use equivalent permanent consumption $\exp (\rho V^i_0(\tilde{i}))$ to measure utility, because this alternative measure is always positive.}

$$\forall \tilde{i} \in \mathbb{I} : \frac{\exp (\rho V^i_0(\tilde{i}))}{\int_{\mathbb{I}} \exp (\rho V^i_0(j)) d\tilde{j}} = u(\tilde{i}).$$

(33)

Under this constraint, any welfare objective that puts nonnegative weights on all $V^i_0(\tilde{i})$ leads to the same wealth allocation choice and type assignments by the planner. Thus, expected utilities after the planner assignment under the two banking policies can always be Pareto ranked. This Pareto ranking could still depend on the chosen distribution of relative utilities $(u(\tilde{i}))_{\tilde{i} \in [0,1]}$. The
next proposition shows that this is not the case.

**Proposition 14.** For any initial capital stock $K_0$, risk state $s_0$ and distribution of relative utilities $(u(\tilde{i}))_{\tilde{i} \in I}$, any planner that chooses the type assignment $t : I \to \{h, b\}$ and the initial wealth distribution $(\tilde{\eta}_0(\tilde{i}))_{\tilde{i} \in I}$ in order to maximize any welfare function that is nondecreasing in all individual utilities $V_0(\tilde{i})$ such that (33) is satisfied, chooses the initial banker wealth share $\eta_0$ so that it solves the problem

$$v_0^p(s_0) := \max_{\eta_0} \left[ (1 - \eta_0) e^{\rho v_h(s_0, \eta_0)} + \eta_0 e^{\rho v_b(s_0, \eta_0)} \right],$$

where $v^h$ and $v^b$ are defined as in Proposition 13.

Furthermore, the Pareto ranking between the banking policies only depends on the value of $v_0^p(s_0)$.

The previous result is powerful because it says that the planner’s choice of the initial state $\eta_0$ is always the same and $v_0^p(s_0)$ is a sufficient statistic for the Pareto ranking across different policies, no matter which distribution of relative utilities the planner is forced to respect. For this reason, I use $v_0^p(s_0)$, either at a fixed initial $s_0$ or integrated over the ergodic distribution of $s_0$, as a welfare index to assess the normative effects of narrow banking.

I start by considering the steady state model discussed in Section 5.2. Assuming for now that the planner chooses $\eta_0 = \eta^{ss}$, the terms related to the evolution of $\eta$ in equation (32) disappear. This assumption turns out to be correct only under narrow banking, yet starting from the steady-state objective is informative also for fractional reserve banking. After substituting in the steady-state values and computing the integrals in $v^h$ and $v^b$, the steady-state welfare index can be written as

$$\log v^p = \log(a - \iota^{ss}) + \Phi(\iota^{ss}) - \frac{\delta}{\rho} + \log \left( \eta^{ss} \exp \left( -\frac{1}{\rho} \left( \frac{\kappa}{\eta^{ss}} (1 - \vartheta^{ss}) \beta \tilde{\sigma} \right)^2 \right) + (1 - \eta^{ss}) \exp \left( -\frac{1}{\rho} \left( 1 - \frac{1}{\eta^{ss}} (1 - \vartheta^{ss}) \frac{1}{\tilde{\sigma}} \right)^2 \right) \right).$$

(34)

where $\iota^{ss}$ is a function of $\vartheta^{ss}$, compare Proposition 4. The values of $\eta^{ss}$ and $\vartheta^{ss}$ are explicitly given by Proposition 9 for a cash-in-advance constraint and in general we know that $\eta^{ss}$ and $\vartheta^{ss}$
are both larger under narrow banking than under fractional reserve banking, compare Proposition 10. The first term in the expression for $\log v^p$ captures the trade-off between consumption and savings. For $\iota^{ss} < \iota^{FB}$, this term is strictly increasing in $\iota^{ss}$ and thus strictly decreasing in $\vartheta^{ss}$. The second term captures idiosyncratic risk sharing in the economy. It is strictly increasing in $\vartheta^{ss}$ because a larger value of money allows for more self-insurance against idiosyncratic risk.

For any given $\eta^{ss}$, this implies a static trade-off between growth and risk sharing. This is the same trade-off as in the papers by Brunnermeier and Sannikov (2016b) and Di Tella (2019) who consider similar models without bankers. These authors show that in the absence of a monetary friction, the competitive equilibrium leads to a too low value of $\vartheta$ and thus over-investment for low idiosyncratic risk $\tilde{\sigma}$, and the competitive equilibrium leads to a too large value of $\vartheta$ and thus under-investment for high idiosyncratic risk $\tilde{\sigma}$. Di Tella (2019) further shows that in the presence of a monetary friction modeled as money in the utility function, the parameter region that features under-investment expands. Both results are also true in my setting. As a consequence, the lower steady-state growth under narrow banking in my framework may actually increase welfare if idiosyncratic risk is low and the monetary friction is not severe.

In addition to the larger $\vartheta^{ss}$, narrow banking also leads to an increase in $\eta^{ss}$. In the appendix, I show that this increase is always welfare-improving. The reason is that under narrow banking, $\eta^{ss}$ adjusts such that both agent types earn an identical idiosyncratic risk premium on their net worth portfolio. As a result, $\eta^{ss}$ under narrow banking minimizes the idiosyncratic risk sharing term in equation (34) for any given $\vartheta^{ss}$. Under fractional reserve banking, $\eta^{ss}$ is below this minimizing wealth share because bankers choose a higher leverage in order to create additional media of exchange for households, which over-exposes them to idiosyncratic risk relative to households, compare the discussion following Proposition 7 in Section 5.1 and following Proposition 10 in Section 5.2. While the planner will for this reason generally not choose $\eta_0 = \eta^{ss}$ so that equation (34) represents only a lower bound for welfare under fractional reserve banking, it remains true that this additional portfolio distortion under fractional reserve banking tends to be welfare-reducing.

In total, the steady-state welfare effects of narrow banking are ambiguous: the negative growth effect due to crowding out can both improve or reduce welfare, but narrow banking also removes a portfolio choice distortion for bankers and this tends to improve welfare. A

---

55 How the monetary friction increases the under-investment region is particularly easy to see in my framework: because the monetary friction is a pure decision wedge without real resource implications, a stronger monetary frictions (lower $\tilde{v}$) raises $\vartheta$ and thereby lowers investment, yet does not affect the planner’s objective, so that $\vartheta$ may end up being too large.
sufficient condition for narrow banking to be welfare-reducing without aggregate shocks is that the steady-state welfare index stated in equation (34) is larger under fractional reserve banking than under narrow banking.

I now move to the full model with a stochastic $s$ process. Then, the additional term in (32) relating to the drift and volatility of $\eta$ in the welfare decomposition matters. The effects of the drift-dependent term are hard to characterize theoretically, but they are arguably of secondary importance for the planner’s objective, because the drift always enters household and banker utility with opposite sign. With regard to the volatility term, I have shown in Section 5.3 that narrow banking reduces endogenous volatility, so that this term is larger for both households and bankers under narrow banking. In addition to the direct effect of lower endogenous risk on the $\sigma^\eta$-dependent term, lower risk in $\eta$ also reduces the time variation of $\eta_t$, $\sigma_t^{h,h}$ and $\sigma_t^{h,b}$ along the equilibrium path. Because the first and third term in equation (32) are concave in these variables, this reduction in time variation tends to improve welfare. Overall, these considerations imply that the stability under narrow banking is welfare-improving. The total welfare effect of narrow banking results from the combination of this positive stability effect with the – negative or positive – steady state effect.

I assess welfare quantitatively in the context of my calibrated model from Section 5.3. Figure 6 plots the maximized planner objective $v^p(s_0)$ as a function of the idiosyncratic risk $\tilde{\sigma}(s_0)$ implied by the initial exogenous state $s_0$. It shows that for this example narrow banking is always welfare-improving. When integrating $v^p(s_0)$ over the ergodic distribution of $s_0$, the value under fractional reserve banking is 0.0494 whereas the value under narrow banking is 0.0527, so that also by this criterion the welfare effects of narrow banking are positive. These numbers mean that one would have to scale up each agent’s consumption by 6.68% in the planner’s allocation under fractional reserve banking to make them indifferent to the planner’s allocation under narrow banking.

Figure 7 provides further insights into the sources of the welfare effects in the numerical example. The black dashed lines show the differences $v^{i,nb}(s_0, \eta_0^{nb}) - v^{i,frb}(s_0, \eta_0^{frb})$ between the $v^i$-components of agents’ value functions (compare Proposition 13) under narrow banking (“nb”) and fractional reserve banking (“frb”), each evaluated at the respective initial $\eta$-value chosen by the planner. The four solid colored lines depict the differences of the individual components in equation (13), where the second term in that equation is split in the figure further into a drift-dependent and a volatility-dependent component.

For both agents, the aggregate consumption-savings term is lower under narrow banking which is represented by the negative yellow lines. This is due to additional investment crowding.
Figure 6: Welfare under fractional reserve banking (blue line) and narrow banking (red line) as a function of the exogenous risk state. $v^p$ is the planner’s welfare objective as defined in Proposition 14.

Figure 7: Decomposition of the welfare effects of narrow banking. Each line shows the difference between the respective term of $v^i$ in decomposition (32) under narrow banking and under fractional reserve banking. The term related to the $\eta$ evolution in (32) is here further split into a drift-dependent and a volatility-dependent term. The left panel depicts the decomposition for households, the right for bankers.
out under narrow banking. In contrast, the term related to the volatility of the state variable $\eta$ is larger under narrow banking, which is represented by the positive green lines. The larger $\eta$-volatility term is a direct effect of the reduced endogenous risk under narrow banking. This term is primarily relevant for the welfare of bankers as they bear most of the aggregate risk under fractional reserve banking.

The light blue lines depict the term related to the idiosyncratic risk exposure of agents. This term is strictly larger for households, both due to the steady-state effect discussed above that comes from a larger level of the money wealth share $\vartheta$ and to the lower steady-state $\eta$, which implies reduced time spent in the “crisis region” where $\kappa < \overline{\kappa}$ and thus idiosyncratic risk diversification by bankers is inhibited. For bankers, the term is negative instead because due to the lower steady-state $\eta$, bankers operate with higher leverage under narrow banking and thus higher idiosyncratic risk exposure per unit of wealth.

The purple lines represent the differences of the drift terms. These turn out to be small for households and for bankers as long as $\tilde{\sigma}$ is not too large. For larger $\tilde{\sigma}$ the term becomes quantitatively relevant because the planner under fractional reserve banking chooses the initial $\eta_0$ to the left of the ergodic mean so that bankers expect to get richer over time, while this does not happen under fractional reserve banking. However, this planner behavior does not appear to be a robust feature across parameter specifications.\textsuperscript{56}

\section{Extensions and Model Variations}

In this section, I discuss three topics I have abstracted from in my baseline analysis, the existence of non-monetary government debt, monetary policy, and bank runs.

\subsection{Government Debt}

So far I have assumed that money is the only government liability. Taken literally, $\vartheta$ should be the ratio of the value of the monetary base to the value of all assets held by the private sector. This ratio is tiny, smaller than typical values of $\vartheta$ in my calibration. Yet, government debt is also a nominally safe asset – at least if the country issues nominal debt in its own currency – with properties very similar to the properties of money in the model, except that it may not be

\textsuperscript{56}Note also that the impact of the drift term makes the total difference $v^{b,\text{nb}}(s_0, \eta^{\text{nh}}_0) - v^{b,\text{frb}}(s_0, \eta^{\text{frb}}_0)$ negative for bankers if $s_0$ is large. This should not be confused to imply that bankers may lose from a narrow banking policy whereas households always gain. $v^b$ only measures welfare per unit of initial wealth. The planner engineers a Pareto improvement by also giving each banker more wealth under narrow banking than under fractional reserve banking.
usable as a medium of exchange. Here, I show that it is justified to restrict attention to a model with only money. Specifically, I extend the model to incorporate both money and government bonds that do not provide any transaction services. I show that without further restrictions, the equilibrium under fractional reserve banking is unaltered, as banks optimally hold all the debt and issue deposits against it, so that all government debt is effectively outside money. If narrow banks are also permitted to back deposits by government debt, then the presence of non-monetary government debt does not affect my analysis of narrow banking.\footnote{For the sake of simplicity, my analysis abstracts from long-maturity debt and (nominal) interest rate risk. In the presence of long-term government debt, these conclusions would have to be somewhat qualified.}

Because in the data some government debt is held outside of the banking sector and thus not “enhanced” in the same way by banks to provide transaction services, I also consider a slightly generalized version in which a reduced-form friction prevents bankers from absorbing all government bonds.\footnote{This is important for the calibration of my model presented in Section 5.3.} Specifically, I add two types of short-term government bonds to the model. Let $B_t^1$ be the nominal face value of type-1 bonds and $B_t^2$ be the nominal face value of type-2 bonds outstanding at time $t$. Denote by $i^{b,1}_t$, $i^{b,2}_t$ the nominal interest rate promised on either type of bonds. These rates adjust to clear the market for government bonds. I abstract from taxation and assume that the government uses a combination of money printing and issuance of additional bonds to make interest payments. The nominal government budget constraint is

$$dM_t + dB_t^1 + dB_t^2 = \left( i^{b,1}_t B_t^1 + i^{b,2}_t B_t^2 \right) dt.$$  

The only difference between the two types of bonds is that type-1 bonds are eligible to be held by both bankers and households, whereas type-2 bonds can only be held by households for exogenous reasons. I limit attention to the case that the government keeps the ratios $b^1 := \frac{B_t^1}{M + B_t^1 + B_t^2}$ and $b^2 := \frac{B_t^2}{B_t^1 + B_t^2 + M}$ of each variety of bonds to total liabilities constant over time.\footnote{Nothing substantial changes for more general $b^1$, whereas the dynamics of the model become more involved for more general $b^2$.} Money and bonds differ by how they affect the velocity of households’ portfolio. Specifically, I continue to assume that velocity under fractional reserve banking is determined by equation (9), so that the portfolio weight invested in bonds, $\theta_t^{b,h}$, does not enter the definition of velocity, whereas the combined portfolio weight invested in money and deposits, $\theta_t^{m,h} + \theta_t^{d,h}$, does. For the narrow banking specification, I explicitly add a competitive sector of narrow banks to the model. The representative narrow bank operates with zero equity, invests in outside money and type-1 bonds and issues monetary narrow banking deposits to households. Both outside money holdings and narrow banking deposits enter household velocity in the denominator.
Under these assumptions, the equilibrium is precisely the same as in my baseline analysis with only money if \( b^2 = 0 \), and is equivalent to a variant of my model with only money in which bank deposits are the superior medium of exchange if \( b^2 > 0 \).

**Proposition 15.** If \( b^1 > 0 \), then bankers (under fractional reserve banking) or narrow banks (under narrow banking) hold all bonds and issue monetary deposits against them. The equilibrium nominal bond rate is \( i^{b,1} = 0 \) and the dynamics of \( \vartheta, \eta, \kappa \) and \( v \) are identical to the model with \( b^1 = 0 \).\(^{60}\) Furthermore:

1. If in addition \( b^2 = 0 \), then the equilibrium is identical to the one of the model with only outside money presented in Section 3.

2. If \( b^2 > 0 \) and \( \kappa_t (1 - \vartheta_t) \geq \eta_t \) in equilibrium at all times, then the model is isomorphic to model with only outside money with a non-integer value for \( \psi \) and an additional restriction that bankers cannot hold outside money.\(^ {61}\) The parameter transformation between the original and isomorphic model is given by\(^ {62}\)

\[
\begin{align*}
\bar{v}' &= (1 - b^2)^{1 - \frac{1}{\alpha}} \bar{v}, \\
\psi' &= \begin{cases} 
0, & \text{narrow banking} \\
\frac{1}{1 - b^2}, & \text{fractional reserve banking}
\end{cases}
\end{align*}
\]

### 7.2 Monetary Policy

In my baseline analysis, I have assumed a constant money supply and ignored any interaction between monetary policy and the banking regime. The first purpose of this section is to show that this assumption is w.l.o.g., provided monetary policy is conducted by paying interest on reserves or by open market operations with government bonds. The second purpose is to discuss how narrow banking can enlarge or restrict the policy space for other types of central bank policies.

Suppose first that there are no bonds and the monetary authority pays a nominal interest \( i^m_t \) on reserves (outside money)\(^ {63}\) and finances the interest payments by money printing, i.e.

\[^{60}\text{For the purpose of this extended model, } \vartheta \text{ is defined as the ratio of all nominal net wealth relative to total net wealth, } \vartheta_t := \frac{(M_t + B^1_t + B^2_t)}{p_t}.
\]

\[^{61}\text{This additional restriction does not change any of the insights discussed in Section 5. In fact, all equations presented there have been derived under the assumption } \kappa_t (1 - \vartheta_t) \geq \eta_t \text{ and then it is w.l.o.g. to assume that households hold all outside money. Therefore, these equations remain valid in the isomorphic transformed model.}
\]

\[^{62}\text{One has to re-interpret the money premium } \lambda' v' \text{ in the transformed model. It maps to the money premium } \lambda v \text{ in the original model according to the identity } \lambda' v' = (1 - b^2) \lambda v. \text{ All other equilibrium quantities are identical in both models.}
\]

\[^{63}\text{I abstract from cash holdings and assume that all outside money is interest-paying.}
\]
\[ dM_t = i^m_t M_t dt. \] Monetary policy is then super-neutral, because inflationary effects of money printing are exactly offset by interest payments on money.

**Proposition 16** (Super-neutrality of interest rate policies). *Monetary policy conducted by altering the interest rate paid on outside money is super-neutral.*

Formally, for any nominal interest rate process \( \{i^m_t\}_{t=0}^\infty \), if the monetary authority makes interest payments \( i^m_t M_t dt \) on the outstanding outside money stock and finances them by money printing, \( dM_t = i^p_t M_t dt \), then all equilibrium solution functions (compare Definition 1) are on the support of the ergodic state distribution identical to the ones of the baseline equilibrium with a constant money supply and no interest on reserves – with the exception that \( i^d_t \) must be replaced with \( i^d_t - i^m_t \).

A second common monetary policy is the use of open market operations to change the fraction of government debt that is outside money. It appears that this should be a powerful policy that can potentially stabilize the financial system even under fractional reserve banking by using discretionary policy to create stabilizing money premium movements in a similar way as narrow banking does automatically. The next proposition, which assumes the generalized setting from Section 7.1, shows that this may not be the case.

**Proposition 17** (Super-neutrality of open market operations). *Monetary policy conducted by open market operations that swap outside money for government bonds is super-neutral, provided the central bank cannot affect the fraction of bonds held outside the banking sector.*

Formally, in the model extension presented in Section 7.1, if the central bank can vary \( b^1 \) by choosing a process \( \{b^1_t\}_{t=0}^\infty \), then all equilibrium solution functions (compare Definition 1) are on the support of the ergodic state distribution identical regardless of the choice of \( \{b^1_t\}_{t=0}^\infty \).

The proposition assumes that the central bank cannot impact the fraction \( b^2 \) of government debt held outside the banking sector. This is most likely a knife-edge case. While doing so is beyond the scope of the present paper, modeling explicitly the friction that leads to \( b^2 > 0 \) in the first place would likely lead to the conclusion that open market operations have some effect on the fraction \( b^2 \).\(^{64}\) Nevertheless, Proposition 17 is an important benchmark result that shows

\(^{64}\)E.g., one conceivable “friction” is limited competition in the banking sector (in a model with just one type of government bonds). Then banks optimally choose to absorb not all bonds to earn a spread between the deposit rate and the bonds in their portfolio. When the central bank exchanges non-interest-bearing money for bonds, holding on to the additional money would reduce the average spread that banks earn and induce them to demand more bonds from households instead.
that using monetary policy to replicate the automatic stabilizing force under narrow banking must rely on some additional frictions and is certainly not trivial.\footnote{In addition, even if monetary policy was able to control the fraction $b^2$ in some way, a similar stability-growth trade-off would emerge as for narrow banking: in order to reduce the money premium in response to adverse shocks by reducing $b^7$, the government would need to keep some “ammunition” in good times in the form of non-monetary government bonds held outside the banking sector. But this additional non-monetary government debt would again generate more crowding out and thus lower growth relative to a situation where all government debt is money.}

The previous two results show why monetary policy considerations are not necessarily strongly interrelated with the choice of the banking regime. There are nevertheless other central bank policies that are directly affected by introducing narrow banking. These are central bank lending to the banking sector\footnote{Lending here takes the form of either uncollateralized lending or lending against private assets on banks’ balance sheets (which are ultimately backed by capital). Lending against collateral in the form of government debt could not achieve an increase in the quantity of transaction media under narrow banking.} and reserve requirements. Formally, I model central bank lending by the central bank offering bankers funds in the form of nominally safe short-term debt at a nominal rate $i^c_{t} \geq i^m_{t}$.\footnote{The lending rate cannot be lower than the reserve rate as otherwise their would be arbitrage opportunities for bankers.} I model reserve requirements as an additional portfolio restriction $\theta^{m,b}_t + \varrho_t \theta^{d,b}_t \geq 0$ for monetary deposits (but not for non-monetary deposits), where $\varrho_t$ is the required reserve ratio.

**Proposition 18** (Additional central bank policies). Suppose the central bank can choose state-contingent paths for $\{i^c_t\}_{t=0}^\infty$ and $\{\varrho_t\}_{t=0}^\infty$.

1. With respect to the instrument $\{i^c_t\}_{t=0}^\infty$, the policy space under narrow banking is strictly larger than under fractional reserve banking in the sense that any allocation implementable under fractional reserve banking can also be implemented under narrow banking.

2. With respect to the instrument $\{\varrho_t\}_{t=0}^\infty$, the policy space under fractional reserve banking is strictly larger than under narrow banking.

The first result is effectively a restatement of the equivalence between private and public money provision discussed by Brunnermeier and Niepelt (2019). Narrow banking prevents non-narrow banks from issuing monetary deposits themselves, but if the central bank provides loans to non-narrow banks at the money rate, it effectively intermediates funds between narrow banks and non-narrow banks, thereby replicating fractional reserve banking. It can implement additional allocations by charging a spread between the loan rate $i^c_t$ and the money rate $i^m_t$. The second result is a simple consequence of the fact that under narrow banking, reserve requirements...
have no effect at all because non-narrow banks are prohibited from issuing monetary deposits at all and narrow banks face already a 100% reserve requirement.

The conclusion from the results in this section is that monetary policy and narrow banking are largely independent (Propositions 16 and 17) and to the extent that they are not, no clear ranking can be given regarding which banking regime – fractional reserve or narrow banking – is more conductive for effective monetary policy (Proposition 18).

7.3 Bank Runs

In my baseline analysis, I have disregarded the possibility of bank runs. Preventing bank runs is an often-claimed benefit of narrow banking, but the narrow banking literature has not reached an agreement whether narrow banking can only prevent runs on money-issuing banks – which is obviously the case – or also runs on financial institutions at large. Authors who emphasize run risk therefore often advocate a form of Simons-style narrow banking (e.g. Chamley, Kotlikoff, and Polemarchakis, 2012; Cochrane, 2014; Kotlikoff, 2010). Here I revisit a position taken by Fisher (1935) that even Fisher-style narrow banking would reduce run risk on non-narrow banks despite a funding structure of these banks that remains in principle run-prone.68

Specifically, I discuss vulnerability to the risk of unanticipated systemic sunspot runs.69 A systemic bank run in the model is an event in which depositors withdraw their funds from banks in the expectation that banks have to default, provided all other depositors run. In the attempt to service deposit withdrawals, banks simultaneously sell assets, thereby shifting capital assets to households and reducing the quantity of deposits. This results in a reduction of the capital price $q^K$ and an increase in the money value $q^M$. If such price movements are large enough, they can lead to banks’ ex post default and thus make the initial expectations of depositors self-fulfilling. Such a systemic run is costly, even though there are no liquidation costs in the model, because it leads to a reduction of aggregate bank capitalization $\eta$. Modeling systemic runs in this fashion is a common approach in the literature (e.g. Gertler and Kiyotaki, 2015; Mendo, 2018).

While a simultaneous default of all bankers would lead to $\eta = 0$ after a run, I allow for the possibility that bankers can rescue a fraction of their pre-run wealth in the event of a run and

68 Specifically, he refers to savings deposits that are not subject to a reserve requirement (Fisher, 1935, Chapter IX).

69 I focus on unanticipated runs for simplicity. A full analysis of runs that are expected in equilibrium is possible, but beyond the scope of the present paper. I have verified that for small positive sunspot probabilities, equilibrium functions change only moderately when taking anticipation effects into account. In particular, there is no discontinuity when moving from unanticipated to anticipated runs and the following discussion remains relevant if runs are expected.

58
Figure 8: Vulnerability to unanticipated systemic runs under fractional reserve banking (left) and narrow banking (right). The color maps correspond to the smallest proportion of bank losses in a run that are necessary for a systemic run to be self-fulfilling. The white area in the right panel means no run vulnerability.

use these funds to start a new bank. I view this as a simple device that models in reduced form the impacts of a bailout, of heterogeneous exposures of banks to runs (so that some banks survive) or entry of new banks after the systemic run.\footnote{Without such a device in the model, a run would lead to the default of the entire banking system and drive the economy permanently to the \( \eta = 0 \) state that resembles the model without bankers discussed in Section 4. The relative asset price movements associated with such an event would be unrealistically severe and thus systemic run vulnerability would be overstated.} Formal details on this model extensions are provided in Appendix E.3.

Figure 8 plots a measure of run vulnerability as a function of the model’s state variables \( (\tilde{\sigma}(s), \eta) \) under fractional reserve and narrow banking. Specifically, the figure depicts for each state the minimum proportion of wealth that the banking sector as a whole must lose in a run such that an immediate run at that state can be self-fulfilling. If the loss is larger than that minimum, then a run is always also self-fulfilling because the function \( \vartheta(s, \eta) \) is strictly decreasing in \( \eta \) for low \( \eta \) values, so that bankers face more adverse price movements for larger downward jumps of \( \eta \).

The figure shows that there is substantial run vulnerability under fractional reserve banking.
Run vulnerability is largest when $\eta$ is close to its stochastic steady state and the aggregate risk state $s$ is small, a region where endogenous risk $\sigma^\eta \eta$ is not necessarily elevated (compare Section 5.3). As in Mendo (2018), there can be substantial “hidden risk” in the form of run risk.

The situation is very different under narrow banking as the right panel in Figure 8 shows. In most of the state space, the system is never vulnerable to runs, even if those runs would drive $\eta$ permanently to zero. Where the system is vulnerable at all, reductions in $\eta$ would need to be larger than 80% to trigger price movements that make a run self-fulfilling. The reason for this reduced run vulnerability is the same as for reduced endogenous risk under narrow banking: for the reasons discussed in Section 5.3, asset prices $q^K$ and $q^M$ are almost insensitive to changes in bank capitalization $\eta$, so that prices do not react sufficiently to a run to make banks ex post insolvent.

In total, my model confirms the claim of Fisher (1935) that narrow banking would substantially reduce run risk even for non-narrow banks who retain a run-prone liability structure under narrow banking.

8 Conclusion

This paper analyzes the macroeconomic implications of narrow banking in a macro model with financial frictions and money. My analysis confirms the claim of narrow banking advocates that such a policy would substantially improve financial stability. In my model, both movements in the relative prices of monetary and real assets and in the capitalization of the banking sector are primarily caused by portfolio reallocation between capital assets and money. The determinant of agents’ desire to undertake such portfolio reallocation is the total required capital premium over money, which consists of the sum of a risk premium that captures the demand for safety and the money premium that captures the demand for transaction services. Under fractional reserve banking, the risk premium and the money premium tend to covary positively, because banks shrink the supply of inside money precisely when the risk premium rises. Narrow banking enhances stability by insulating the money supply from banks’ balance sheet adjustments. A rise in the risk premium is then met by a reduction in the money premium that automatically stabilizes the total capital premium and thereby mitigates portfolio reallocation.

My analysis further implies that the restrictions on the financial sector imposed by narrow banking do not affect the steady-state production allocation substantially. Claims that narrow banking would be costly because it limits the ability of banks to fund real economic activity appear unfounded from the perspective of the framework developed in this paper. A narrow
banking policy can nevertheless have output and growth effects, although through a more subtle channel, i.e., investment crowding out resulting from a wealth effect on unproductive outside money that represents positive net wealth for the private sector. I find is that narrow banking decreases investment and growth through this crowding out channel.

In this paper, I have abstracted from a number of issues. Most importantly, I have assumed a dichotomy between transaction media and other types of nominal debt and taken the position that effective legal barriers can be imposed that prevent the usage of non-narrow bank deposits in transactions under narrow banking. I have thereby sidestepped the issue of how to treat “near-monies” under narrow banking, an issue that is widely debated in the narrow banking literature, particularly by advocates of Simons-style narrow banking. While I believe that my modeling choice serves theoretical clarity of the present paper, future research should clarify what additional restrictions on non-narrow banks are required to make narrow banking effective and to what extent these additional restrictions make narrow banking more costly.

References

Ball, Laurence (2001). “Another look at long-run money demand”. In: Journal of Monetary Economics 47.1, pp. 31–44.


Friedman, Milton (1959). *Debt Management and Banking Reform.*
Kotlikoff, Laurence J. (2010). *Jimmy Stewart is dead: ending the world’s ongoing financial plague with limited purpose banking*. John Wiley & Sons.


Appendix

A A Review of the Narrow Banking Literature

In this appendix, I provide a brief review of the literature on narrow banking policies. More extensive surveys of the literature are provided by Beneš and Kumhof (2013) and Lainà (2015). Phillips (1992) gives a historical account of the Chicago Plan and how it influenced New Deal banking legislation.

The first version of the original Chicago Plan was communicated in the form of a short memorandum sent to the secretary of agriculture, Knight et al. (1933). The document makes a number of concise policy recommendations. While the majority constitutes emergency measures to fight the ongoing Great Depression, the authors also recommend long-term reforms regarding the Federal Reserve System and banking. The banking reform portion recommends the gradual resolution of existing financial institutions and establishment of two new types of institutions, narrow banks that “shall serve exclusively as institutions for deposits and transfers of funds” and are “required to maintain reserves of 100%” (p. 3) and “investment trusts” or “Lending Companies” that should perform remaining banking functions. The memorandum is somewhat ambiguous with regard to the means of financing such investment trusts should have access to.\(^{71}\) A second unpublished memorandum containing a more detailed outline of the reforms was disseminated in November 1933 (see Phillips, 1992). No specific summary of its contents can be given here, as I have been unable to obtain a copy of this second memorandum. Versions of the reform plan were subsequently outlined in Fisher (1935), Simons (1936), Douglas et al. (1939) and discussed by Hart (1935), Angell (1935), Fisher (1936) and Lehmann (1936). To these authors, the 100% reserve requirement serves the primary purpose of reducing cyclical volatility of the money supply in the fractional reserve system that in their view is the primary driver of business cycle fluctuations.\(^{72}\) Accordingly, Fisher (1935) and Douglas et al. (1939) maintain

\(^{71}\) On p. 4 the memorandum says that these institutions, “like other corporations, would be in position to lend and invest only the funds invested by its stockholders (and, perhaps, bondholders)”, suggesting that the authors have in mind a financial system that is primarily equity-funded, but would potentially also allow for bond-financing. In addition, “[f]or investors preferring the savings-deposit form, the facilities of the Postal Savings System would still be available” (p. 4). Depending on how the authors thought the Postal Savings System to operate after the reform, this aspect could have introduced some backdoor short-term funding of financial institutions into the new financial architecture.

\(^{72}\) An exception is Lehmann (1936), who is generally less enthusiastic of the reform proposals. With regard to the relationship between banking and business cycles, he claims to express a “view [that] is sharply opposed to the doctrine maintained by Fisher and others that the destruction of bank money is responsible for falling prices and growing unemployment”(p. 46). He favors an explanation that is more in line with a financial accelerator
that banking could largely continue as before the reform, but demand deposit funding would have to be replaced with savings or time deposits and appropriate regulation to ensure that such claims to banks cannot circulate as means of exchange. Simons (1936) is more skeptical and stresses that one should “conceive the problem broadly as that of achieving a financial structure in which the volume of short-term borrowing would be minimized” (p. 17). He discusses a number of financial arrangements with regard to the question which would be most conductive to real activity. Among those discussed, he identifies as “approximately ideal” a system where there are “no fixed money contracts at all – ... all property ... held in a residual-equity or common-stock form”. This ideal system would abandon private debt contracts at large and thus go much further than merely restricting bank money creation. An intermediate position is taken by Hart (1935), who does not explicitly endorse Simon’s proposal, but sees non-reserve backed time deposits as a fragile arrangement, and Angell (1935), who proposes to convert time deposits into longer-maturity time obligations.

In addition to a stable money supply and a reduction in business cycle volatility, claims were made that the Chicago Plan could eliminate bank runs (Angell, 1935; Douglas et al., 1939; Fisher, 1935, 1936), it would allow to eliminate deposit insurance and much of existing banking regulation (Douglas et al., 1939; Fisher, 1935, 1936; Graham, 1936; Hart, 1935), it would enable the government to appropriate seigniorage rents from banks (Fisher, 1935, 1936; Graham, 1936) and its introduction would lead to a massive reduction of government debt (Fisher, 1935, 1936). With respect to these claims, the academic debate of the late 30s and early 40s did not reach full agreement. The elimination of bank runs for demand deposits is generally acknowledged, but Hart (1935), Lehmann (1936) and – to a lesser extent – Higgins (1941) worry that certain models of implementation of the Chicago Plan would make other funding sources of banks more run-prone. That regulation could be reduced or eliminated appears to be wide agreement,

mechanism and maintains that the “volume of bank money is a consequence of calling credits and selling securities rather than a factor of independent significance” (p. 46).

Simons (1936), footnote 17 (p. 19): “The so-called 100 per cent scheme of banking reform can easily be defended only as a proper first step toward reconstruction of our whole financial organization. Standing by itself, as an isolated measure, it would promise little but evasion – small effects at the price of serious disturbance – and would deserve classification as merely another crank scheme.”

This view is reinforced in a later article (Simons, 1946, p. 85): “In its more important, converse aspect, 100 per cent reserve banking is simply 100 per cent equity financing of all incorporated enterprises”.

Like Fisher (1935) and Douglas et al. (1939), Simons (1936) also sees fluctuations in the quantity of the medium of exchange as the primary concern, but unlike the aforementioned authors considers “near monies” like short-term debt to contribute to that quantity. While some of his arguments describe a financial accelerator mechanism or run risk, this is not his main point of emphasis. The mechanism he primarily emphasizes are deflationary pressures induced by fluctuations in the effective money supply.

Lehmann (1936), for example, discusses a “run” in which a large number of time depositors do not extend
but is only stressed by some. The appropriation of seigniorage rents is particularly stressed by Graham (1936), whereas Brown (1940) argues that in the fractional reserve system depository institutions distribute seigniorage gains to the public by providing free banking services. The possibility of government debt reduction in the introduction phase appears to be an isolated position of Fisher (1935) that for most other commentators of the time is – in Hart’s (1935) words – “illusory” (p. 105). Despite this substantial disagreement regarding the precise benefits of the plan and even larger disagreement with respect to implementation details, most authors at the time endorse the policy proposal at least partially (Angell, 1935; Fisher, 1935; Graham, 1936, 1941; Hart, 1935; Higgins, 1941; Simons, 1936, 1946). Lehmann (1936), Brown (1940) and Thomas (1940) take a critical position.

B Model Solution

B.1 Analytical Solution Steps

In this appendix, I provide additional details and proofs to the model solution, which has only been sketched in Sections 4 and 5 of the main text. While the model without banks discussed in Section 4 follows from the general version as a special case at the initial state \( \eta = 0 \), many of the solution steps are simpler in this case, so that I discuss them separately first.

B.1.1 The Model of Section 4 without Banks

Consumption-savings Choice and Proposition 1. As in the main text, I start with households’ consumption-savings and optimal investment choice. Because households have logarithmic preferences with time preference \( \rho \), they optimally consume a constant fraction \( \rho \) of their net worth, which implies equation (20). This is a well-known property of logarithmic utility. For a derivation, see e.g. Brunnermeier and Sannikov (2016a) or Merkel (2019).

Households always hold a long position in physical capital, compare the description of the household problem in Section 3.3, and the investment rate choice \( \tau_t \) only affects the expected return on capital in equation (12). Consequently, households find it optimal to choose \( \tau_t \) in

---

\( \eta = 0 \) indicates the initial state of the model. The time loan, but transfer the funds to a demand deposit account with the same bank to obtain more flexibility in the future. Such a “run” would trigger only little or no loan liquidations in a fractional reserve system. In a full reserve system, in contrast, large loan or asset liquidations would be necessary to come up with the required reserves for demand deposits.

\(^{77}\) Angell (1935) and Higgins (1941) appear to share this position, other cited articles in this paragraph do not comment explicitly on the issue.
order to maximize that expected return on capital. Taking the first order condition of the drift in (12) with respect to $\iota_t$ leads directly to equation (21).

Combining the goods market clearing equation (22) with household’s optimal consumption choice (20) and dividing the resulting equation by $q_t^K$ yields

$$
\rho \frac{q_t^K + q_t^M}{q_t^K} = a - \iota_t.
$$

Next, $\frac{q_t^K + q_t^M}{q_t^K} = \frac{1}{1 - \vartheta_t}$ by equation (4) (definition of $\vartheta$) and $q_t^K = 1 + \phi_t$ by equation (21). Substituting these expressions into the previous equation and solving for $\iota_t$ yields

$$
\iota_t = \frac{(1 - \vartheta_t)a - \rho}{1 - \vartheta_t + \phi \rho},
$$

which is the expression for $\iota_t$ stated in Proposition 1. Substituting this back into $q_t^K = 1 + \phi_t$ and plugging the result into $q_t^M = \frac{\vartheta_t}{1 - \vartheta_t}q_t^M$ implies the expressions for $q_t^K$ and $q_t^M$ stated in Proposition 1. This concludes the proof of that proposition.

**Portfolio Choice and Proposition 2.** I next turn to the portfolio choice problem that characterizes the dynamics of the remaining unknown $\vartheta_t$. I express the optimal portfolio choice in terms of martingale pricing conditions following the methodology advocated by Brunnermeier and Sannikov (2016a). Because my household problem is nonstandard due to the dependence of the capital return on velocity, I rederive these conditions for my model using the stochastic maximum principle. The resulting expressions turn out to become more transparent and simpler if a change of numeraire is performed first before setting up the Hamiltonian. Specifically, instead of working with the net worth evolution (8) where net worth is measured in units of consumption, I express net worth in the following as a fraction of aggregate wealth $\bar{q}_t := (q_t^K + q_t^M)K_t$.\(^{78}\) Restating (8) in terms of this new numeraire requires transforming asset returns. The following lemma provides a generic transformation formula for returns as the numeraire is changed in an environment with Brownian shocks.

**Lemma 1.** Let $R$ be a cumulative return process of an asset following an Ito evolution

$$
dR_t = r_t dt + \sigma_t^r dZ_t + \tilde{\sigma}_t^r d\tilde{Z}_t
$$

\(^{78}\)The model has been introduced using the consumption good as a numeraire instead of starting with this new numeraire to begin with because this is standard in most of macroeconomics.
and let $X_t$ be a positive \textit{Ito process}
\[
dX_t \over X_t = \mu_t^X dt + \sigma_t^X dZ_t
\]
expressed in the same units as the numeraire of $R$. Then the return of the asset in the $X$ numeraire, $\hat{R}$, satisfies
\[
d\hat{R}_t = dR_t - (\mu_t^X + \sigma_t^X (\sigma_t^r - \sigma_t^X)) dt - \sigma_t^X dZ_t.
\]

\textbf{Proof.} Let $x_t$ be the value process of a self-financing strategy investing into the asset (and reinvesting all dividends), i.e.
\[
dx_t = dR_t.
\]
$x_t$ is measured in the original numeraire. In the new numeraire, the value of the same self-financing strategy is $\hat{x}_t = x_t X_t$. Thus, the asset return $\hat{R}$ in the new numeraire must satisfy
\[
d\hat{x}_t \over \hat{x}_t = (r_t - \mu_t^X - \sigma_t^X (\sigma_t^r - \sigma_t^X)) dt + (\sigma_t^r - \sigma_t^X) dZ_t + \sigma_t^r d\tilde{Z}_t
\]
which proves the assertion.

In the special case considered here, $X_t = q_t K_t$ is aggregate wealth. A straightforward application of the previous lemma to the return processes (12) and (15) yields:

\textbf{Lemma 2.} In terms of the aggregate wealth numeraire $q_t K_t$, the return processes on capital $\hat{R}_t^k(v_t)$ and money $\hat{R}_t^m$ are given by
\[
d\hat{R}_t^k(v_t) = \left(\rho \over 1 - \vartheta_t - \bar{\tau}_t(\vartheta_t) - \vartheta_t \mu_t^q \over 1 - \vartheta_t\right) dt - \vartheta_t \sigma_t^q \over 1 - \vartheta_t dZ_t + \bar{\sigma}_t d\tilde{Z}_t,
\]
\[
d\hat{R}_t^m = \mu_t^q dt + \sigma_t^q dZ_t.
\]

\textbf{Proof.} The drift and volatility of the aggregate wealth numeraire are $\mu_t^X = \mu_t^q + \mu_t^K = \mu_t^q +$
\[ \Phi(t) - \delta \text{ and } \sigma_t^X = \sigma_t^q. \] Applying Lemma 1 to the returns (12) and (15)

\[ d\hat{R}_k^k(v_t) = \left( a - \mu_t^q - \frac{\xi_t(v_t)}{q_t^K} \right) dt + \left( \sigma_t^{q,K} - \sigma_t^q \right) dZ_t + \tilde{\sigma}_t d\tilde{Z}_t, \]

\[ d\hat{R}_m^m = \mu_t^{q,M} - \sigma_t^q \left( \sigma_t^{q,M} - \sigma_t^q \right) dt + \left( \sigma_t^{q,M} - \sigma_t^q \right) dZ_t. \]

To obtain the asserted expressions, use \( \vartheta = q^M / \bar{q} \) and \( 1 - \vartheta = q^K / \bar{q} \), which imply \( \sigma_t^{q,M} - \sigma_t^q = \sigma_t^\vartheta \), \( \mu_t^{q,M} - \sigma_t^q = \sigma_t^\vartheta = \sigma_t^{1-\vartheta} = -\vartheta \sigma_t^q / (1 - \vartheta) \) and \( \mu_t^{q,K} - \sigma_t^q = \sigma_t^{q,K} - \sigma_t^q = \mu_t^{1-\vartheta} = -\vartheta \sigma_t^q / (1 - \vartheta) \). In addition, one can replace \( a - \mu_t \) in the capital return with \( \rho(q_t^K + q_t^M) \) using goods market clearing, which results in the expression \( \rho / 1 - \vartheta \) for the dividend yield.

The net worth evolution in the new numeraire is then structurally identical to equation (8), one simply has to replace the return processes \( dR_m^m \) and \( dR_k^k(v_t) \) with the processes \( d\hat{R}_m^m \) and \( d\hat{R}_k^k(v_t) \) from Lemma 2 and express consumption and net worth in the new units (if \( x \) is expressed in consumption units, define \( \hat{x} := x / (\bar{q} K) \) here and in the following)

\[ \frac{d\hat{n}_t^h}{\hat{n}_t^h} = -\frac{\hat{c}_t^h}{\hat{n}_t^h} dt + \theta_t^{m,h} d\hat{R}_m^m + \theta_t^{k,h} d\hat{R}_k^k(v_t). \] (35)

I now set up the Hamiltonian of the household. Let \( \xi_t \) denote the costate associated with the net worth evolution in the form (35). Economically, \( \xi_t^h \) acts as a Lagrange multiplier on the agent’s budget constraint, so it measures the marginal (time 0) value of an additional unit of wealth (where units are expressed as a fraction of aggregate wealth) and is thus the stochastic discount factor (SDF) of the agent. When it is written in the form

\[ \frac{d\xi_t^h}{\xi_t^h} = -r_t^h dt - \xi_t^h dZ_t - \xi_t^h d\tilde{Z}_t, \] (36)

\( r_t^h \) is the shadow risk-free rate for the household, \( \xi_t^h \) is the price of aggregate risk \( (dZ_t \text{ risk}) \) that the household requires to be willing to assume exposure to aggregate shocks (relative to the numeraire) and similarly, \( \xi_t^h \) is the household’s price of idiosyncratic risk \( (d\tilde{Z}_t \text{ risk}) \). The
Hamiltonian of the household’s problem is (using $\theta^m_t + \theta^k_t = 1$ to eliminate $\theta^m_t$)

$$H^h_t = e^{-\rho t} \log c^h_t + \xi_t \left(-\frac{c^h_t}{q^K_t} + \hat{n}^h_t \left(\mu^\vartheta_t + \theta^k_t \left(\frac{\rho}{1-\vartheta_t} - \frac{\Sigma_t(v_t)}{q^K_t} - \frac{\mu^\vartheta_t}{1-\vartheta_t}\right)\right)\right)$$

$$- \varsigma_t \xi_t \hat{n}^h_t \sigma^\vartheta_t \left(1 - \frac{\theta^k_t}{1-\vartheta_t}\right) - \varsigma_t \xi_t \hat{n}^h_t \theta^k_t \tilde{\sigma}_t$$

where velocity $v_t$ must satisfy

$$\frac{a}{q^K_t} \theta^k_t = v_t \left(1 - \theta^k_t\right)$$

by equation (9). By the stochastic maximum principle, the optimal choice $(c^h_t, \theta^k_t, v_t)$ must maximize the Hamiltonian subject to equation (37). For the derivation of first-order conditions, let $\xi_t \hat{n}^h_t \lambda_t$ be the Lagrange multiplier on equation (37). $\lambda_t$ can be interpreted as the price (expressed in the same units as wealth) of an additional unit of velocity in the same way as $\varsigma_t$ and $\tilde{\varsigma}_t$ are the wealth prices of additional units of risk.

The first-order condition for the maximization of $H^h_t$ with respect to $\theta^k_t$ is

$$\xi_t \hat{n}^h_t \left(\frac{\rho}{1-\vartheta_t} - \frac{\Sigma_t(v_t)}{q^K_t} - \frac{\mu^\vartheta_t}{1-\vartheta_t}\right) = \varsigma_t \xi_t \hat{n}^h_t \left(\frac{-\sigma^\vartheta_t}{1-\vartheta_t}\right) + \varsigma_t \xi_t \hat{n}^h_t \tilde{\sigma}_t + \xi_t \hat{n}^h_t \lambda_t \left(v_t + \frac{a}{q^K_t}\right).$$

After dividing by $\xi_t \hat{n}^h_t$, multiplying by $1 - \vartheta_t$, using equation (37) and that in equilibrium $\theta^k_t = 1 - \vartheta_t$ and $\Sigma_t(v_t) = 0$, and solving for $\mu^\vartheta_t$, this equation becomes

$$\mu^\vartheta_t = \rho - \varsigma_t \left(-\sigma^\vartheta_t\right) + \xi_t \left(1 - \vartheta_t\right) \tilde{\sigma}_t + \lambda_t v_t.$$

Equation (38) is essentially the money valuation equation stated in Proposition 2 in the main text, although additional steps are required to bring it into the form stated there. The form (38) justifies the interpretation given in the main text that the second term consists of a sum of risk premia and the money premium. Indeed, $-\sigma^\vartheta_t$ is the incremental aggregate risk in households’ net worth portfolio in excess of the risk of an all-money portfolio. Consequently, $\varsigma_t \left(-\sigma^\vartheta_t\right)$ as the product of the price of risk with the risk is the risk premium that households require on top of the money return to hold the aggregate risk in their portfolio. Similarly, $(1 - \vartheta_t) \tilde{\sigma}_t$ is the idiosyncratic risk in households’ net worth portfolio in excess of an all-money portfolio (which would be free of idiosyncratic risk) and thus $\xi_t \left(1 - \vartheta_t\right) \tilde{\sigma}_t$ is the risk premium households require for bearing this idiosyncratic risk. $\lambda_t v_t$ is the product of households’ price of velocity $\lambda_t$ with the velocity of money holdings $v_t$, so it is the money premium.
To bring the risk premium terms in equation (38) into the form stated in Proposition 2, one has to determine $\varsigma_t$ and $\tilde{\varsigma}_t$. To do so, take first-order conditions for the maximization of $H^h_t$ with respect to $c_t$ and solve for $\xi^h_t$, which implies

$$\xi^h_t = e^{-\rho t} \tilde{q}^h_t K_t = e^{-\rho t} \tilde{q}^h_t K_t \rho_n_t = e^{-\rho t} \frac{1}{\rho_n t}.$$

Here, the second equation uses the optimal consumption rule $c_t = \rho_n t$ and the third equation applies the definition of $\hat{n}_t$. One can now recover the prices of risk $\varsigma_t$ and $\tilde{\varsigma}_t$ by applying Ito’s formula to this equation for $\xi_t$ and comparing the volatility loadings with the ones in equation (36):

$$\varsigma_t = \left(1 - \theta_{k,h}^t \right) \sigma^h_t = 0, \quad \tilde{\varsigma}_t = \theta_{k,h}^t \sigma^h_t = (1 - \vartheta_t) \tilde{\sigma}_t.$$

Substituting these prices of risk into equation (38) implies equation (23) as stated in Proposition 2.

For the remaining equations asserted in Proposition 2, take first-order conditions for the maximization of $H_t$ with respect to $v_t$,

$$\xi^h_t \hat{n}^h_t \lambda_t \left(1 - \theta_{k,h}^t \right) = \xi^h_t \theta_{k,h}^t \frac{\chi'_t(v_t)}{q^K_t}.$$

Using $\theta_{k,h}^t \frac{\chi'_t}{q^K_t} = v_t \left(1 - \theta_{k,h}^t \right)$ (equation (37)) and dividing by all factors that appear on both sides, this equation simplifies to

$$\lambda_t = \frac{1}{a} \chi'_t(v_t) v_t,$$

which is precisely equation (24) in the case $\alpha < \infty$. For the case $\alpha = \infty$, note that the first-order conditions with respect to $v_t$ do not make sense, but instead one has the cash-in-advance constraint $v_t \leq \bar{v}$. It is easy to see that the multiplier $\lambda_t$ is then positive if and only if the cash-in-advance constraint is binding (and zero otherwise). This leads to a complementary slackness condition that can be expressed equivalently as equation (24) for the case $\alpha = \infty$.

The final equation asserted in Proposition 2 is simply a consequence of the constraint (37) combined with market clearing, $\theta_{k,h}^t = 1 - \vartheta_t$, and the representation of $q^K_t$ stated in Proposition 4. This completes the proof of Proposition 2.

Furthermore, note that the integral form of the money valuation equation discussed in the
main text following Proposition 2 can be easily derived by using
\[ d(e^{-\rho t} \vartheta_t) = -\rho e^{-\rho t} \vartheta_t dt + e^{-\rho t} \vartheta_t \, dt \]
integrating both sides over the interval \([t, T]\), taking expectations and then the limit \(T \to \infty\).

**Steady State and Proposition 3.** Due to Proposition 1, there is a monetary steady state if and only if there is a constant solution \(\vartheta \in (0, 1)\) to the money demand conditions stated in Proposition 2 and conversely any such constant solution corresponds to a steady state.\(^8\) If \(\vartheta > 0\) is constant, then in particular \(\mu_t^\vartheta = 0\). Substituting this into the money valuation equation in Proposition 2 yields the equation
\[ \rho = (1 - \vartheta)^2 \sigma^2 + \lambda_t \nu_t. \]
It follows then immediately that also the money premium \(\lambda_t \nu_t\) must be constant and be given by
\[ \lambda \bar{v} = \rho - (1 - \vartheta)^2 \sigma^2. \quad (39) \]
By condition (24) in Proposition 2, this expression must be nonnegative and whenever it is positive, then \(v = \bar{v}\). There are therefore two possible cases:

1. If \(\lambda = 0\), then the cash-in-advance constraint is not binding and solving (39) for \(\vartheta\) implies \(\vartheta = \frac{\sigma - \sqrt{\sigma^2}}{\sigma} = \vartheta^I\) as the only positive (and thus valid) solution. The steady state candidate must therefore be of the “I theory type” in this case. It is a valid steady state if it is also consistent with condition (24), which is for \(\lambda = 0\) equivalent to \(v \leq \bar{v}\). Substituting the quantity equation from Proposition 2 and \(\vartheta = \vartheta^I\) into this condition yields the inequality
\[ \frac{a}{1 + \phi a} \sqrt{\frac{\vartheta + \phi \rho \sigma}{\vartheta - \rho}} \leq \bar{v}. \]
Rearranging implies \(\bar{\sigma} \geq \frac{\vartheta + a(1+\phi \rho)}{\bar{\sigma} - a \vartheta - \rho} \sqrt{\bar{\sigma}} = \vartheta^I\), which is the condition stated in Proposition 3. Therefore, this “I theory” steady state exists if and only if \(\bar{\sigma} \geq \vartheta^I\).

\(^8\)One easily verifies that the \(q^K, q^M, \iota\) as in Proposition 1, \(\mu^n = \sigma^n = \kappa = 0\), \(\vartheta^b = \vartheta^h = \rho\), \(\theta^{m,b} = 0\) and \(\theta^{m,h} = \vartheta\) satisfy all the conditions of a monetary equilibrium in Definition 1 under the additional initial condition \(\eta_0 = 0\). One similarly verifies that any monetary equilibrium according to that definition with \(\eta_0 = 0\) and a constant \(s\) state must be of this form.
2. If \( \lambda > 0 \), then the cash-in-advance constraint is binding and thus \( \bar{v} = \bar{v} \). Combining this with the quantity equation from Proposition 2 yields the equation

\[
\bar{v} = \frac{a}{1 + \phi a} \frac{1 - \vartheta + \phi \rho}{\vartheta}
\]

and solving for \( \vartheta \) implies \( \vartheta = \frac{a(1 + \phi \rho)}{a + a(1 + \phi \rho)} = \vartheta^M \). This candidate steady state therefore satisfies the properties of the “monetarist” steady state in Proposition 3. Note that by Assumption 1 \( \vartheta^M \) is always a number in \((0, 1)\). It is thus a valid steady state if and only if the money premium is positive (otherwise the assumption \( \lambda > 0 \) is violated). Substituting \( \vartheta = \vartheta^M \) into equation (39), this condition is equivalent to

\[
\rho > \left( \frac{\bar{v} + a \phi (\bar{v} - \rho)}{\bar{v} + a (1 + \phi \bar{v})} \right)^2 \sigma^2 \iff \sigma < \sigma^I.
\]

The previous arguments show that the two steady state solutions are the only possible cases and they are mutually exclusive: the “I theory” steady state exists for parameters \( \bar{\sigma} \geq \bar{\sigma}^I \) and the “monetarist” steady state exists for parameters \( \bar{\sigma} < \bar{\sigma}^I \). In particular, there is always a unique monetary steady state. The remaining assertion made in Proposition 3, \( \vartheta = \max \{ \vartheta^I, \vartheta^M \} \), is easily verified by checking that \( \vartheta^I \geq \vartheta^M \iff \bar{\sigma} \geq \bar{\sigma}^I \).

### B.1.2 The Model of Section 5 with Banks

**Consumption-savings Choice and Proposition 4.** The arguments leading to Proposition 1 in the model without bankers solely rest on the fact that all agents’ consumption-wealth rate is \( \rho \), on goods market clearing and on the Tobin’s \( Q \) condition (21). All these conditions remain also valid in the general model with bankers (also bankers have logarithmic preferences with time preference rate \( \rho \) and thus choose optimally a constant consumption-wealth ratio of \( \rho \)). Consequently, Proposition 4 holds.

**Portfolio Choice.** As for the model without banks, I derive portfolio choice conditions using the stochastic maximum principle. Again, it is formally simpler to work with aggregate wealth as the numeraire good. In analogy to Lemma 2, the following Lemma expresses all returns in the new numeraire.

**Lemma 3.** *In terms of the aggregate wealth numeraire \( \tilde{q}_t K_t \), the return processes (12), (13),*
\((14), (15), (16)\) take the form

\[
d\hat{R}_t^k (v_t) = \left( \frac{\rho}{1 - \hat{v}_t} - \frac{\hat{\Sigma}_t (v_t)}{q_t^R} - \frac{\hat{\vartheta}_t \mu_t^\vartheta}{1 - \hat{v}_t} \right) dt - \frac{\hat{\vartheta}_t \sigma_t^\vartheta}{1 - \hat{v}_t} dZ_t + \hat{\sigma}_t d\tilde{Z}_t, \tag{40}
\]

\[
d\hat{R}_t^x = \hat{r}_t^x dt - \frac{\hat{\vartheta}_t \sigma_t^\vartheta}{1 - \hat{v}_t} dZ_t + \hat{\sigma}_t d\tilde{Z}_t, \tag{41}
\]

\[
d\hat{R}_t^{x,b} = \hat{r}_t^{x,b} dt - \frac{\hat{\vartheta}_t \sigma_t^\vartheta}{1 - \hat{v}_t} dZ_t + \beta \hat{\sigma}_t d\tilde{Z}_t, \tag{42}
\]

\[
d\hat{R}_t^m = \mu^\vartheta_t dt + \sigma^\vartheta_t dZ_t, \tag{43}
\]

\[
d\hat{R}_t^d = \left( i^d_t + \mu^\vartheta_t \right) dt + \sigma^\vartheta_t dZ_t, \tag{44}
\]

where \(\hat{r}_t^x\) is an expected return process determined in equilibrium.

\[\text{Proof.}\] Equations (40) and (43) are precisely as stated in Lemma 2 and the proof of their validity is identical word by word.

In the original consumption numeraire the risky claim return processes (13) and (14) have the same \(dZ_t\) loading as the capital return process (12). By Lemma 1, this property is preserved under a change of numeraire, so the \(dZ_t\) loadings in (41) and (42) must coincide with the one in (40), which is indeed the case. Lemma 1 also implies that the \(d\tilde{Z}_t\) loadings of return processes remain unaffected by the numeraire change. Consequently, equations (41) and (42) are correct.

For the remaining deposit return (44), note that \(dR^d_t = i^d_t dt + dR^m_t\) in the consumption numeraire and again using Lemma 1, this implies \(d\hat{R}^d_t = i^d_t dt + d\hat{R}^m_t\). \(\Box\)

The net worth evolutions of the two agent types in the new numeraire are\(^\text{81}\)

\[
\frac{d\hat{n}^h_t}{\hat{n}^h_t} = -\hat{c}^h_t dt + \hat{\theta}^{m,h}_t d\hat{R}^m_t + \hat{\theta}^{d,h}_t d\hat{R}^d_t + \hat{\theta}^{k,h}_t d\hat{R}^k_t (v_t) + \hat{\theta}^{x,h}_t d\hat{R}^x_t,
\]

\[
\frac{d\hat{n}^b_t}{\hat{n}^b_t} = -\hat{c}^b_t dt + \hat{\theta}^{m,b}_t d\hat{R}^m_t + \hat{\theta}^{d,b}_t d\hat{R}^d_t + \hat{\theta}^{x,b}_t d\hat{R}^x_t.
\]

\(^\text{81}\)Compare equations (8) and (10) for the consumption numeraire. Note that as in the solution for the model without bankers, the optimal investment rate choice \(\iota_t\) has already been used for Lemma 3.
The Hamiltonians are consequently

\[ H^h_t = e^{-\rho t} \log c^h_t - \xi^h_t \frac{c^h_t}{q_t K_t} \]

\[ + \xi^h_t \tilde{n}_t^h \left( i_d^i + \mu_i^d - \theta_t^{m,h} i^d_t + \theta_t^{k,h} \left( \frac{\rho}{1 - \theta_t} - \frac{\mathcal{Z}_t}{q_t} - \frac{\mu_i^d}{1 - \theta_t} - i_d^d \right) + \theta_t^{x,h} \left( \tilde{r}_t^x - \mu_i^d - i_d^d \right) \right) \]

\[ - \xi^h_t \xi^h_t \tilde{n}_t^h \sigma_t^d \left( 1 - \frac{\theta_t^{k,h} + \theta_t^{x,h}}{1 - \theta_t} \right) - \xi^h_t \xi^h_t \tilde{n}_t^h \left( \theta_t^{k,h} + \theta_t^{x,h} \right) \tilde{\sigma}_t \]

(45)

for households and

\[ H^b_t = e^{-\rho t} \log c^b_t - \xi^b_t \frac{c^b_t}{q_t K_t} \]

\[ + \xi^b_t \tilde{n}_t^b \left( i_d^i + \mu_i^d - \theta_t^{m,b} i^d_t + \theta_t^{x,b} \left( \tilde{r}_t^x - \mu_i^d - i_d^d \right) \right) \]

\[ - \xi^b_t \xi^b_t \tilde{n}_t^b \sigma_t^d \left( 1 - \frac{\theta_t^{k,b} + \theta_t^{x,b}}{1 - \theta_t} \right) - \xi^b_t \xi^b_t \tilde{n}_t^b \theta_t^{x,b} \beta \tilde{\sigma}_t \]

(46)

for bankers. In analogy to the model without banks, \( \xi^h_t \) and \( \xi^b_t \) can be interpreted as SDF processes for the two agents and can be written in the form \( (i \in \{h,b\}) \)

\[ \frac{d\xi^i_t}{\xi^i_t} = -r_t^i dt - \xi^i_t dZ_t - \tilde{\xi}^i_t d\tilde{Z}_t, \]

(47)

where \( \xi^i_t \) is the agent’s price for bearing aggregate risk and \( \tilde{\xi}^i_t \) is the agent’s price for bearing idiosyncratic risk.

By the stochastic maximum principle, households’ optimal choice of \( (c^h_t, \theta_t^{m,h}, \theta_t^{k,h}, \theta_t^{x,h}, u_t) \) must maximize \( H^h_t \). In their optimal choice, households have to respect the skin-in-the-game constraint

\[ \bar{\kappa} \theta_t^{k,h} + \theta_t^{x,h} \geq 0 \]

(48)

and the definition of velocity

\[ \frac{a}{q_t^h} \theta_t^{k,h} = v_t \left( 1 - \psi \right) \theta_t^{m,h} + \psi \left( 1 - \theta_t^{k,h} - \theta_t^{x,h} \right). \]

(49)

I ignore the remaining nonpositivity and nonnegativity constraints in the household’s problem as they will be automatically satisfied in equilibrium.
Similarly, bankers’ optimal choice of \((c_t^b, \theta_t^{m,b}, \theta_t^{x,b})\) must maximize the Hamiltonian \(H_t^b\). Again, I ignore the constraint \(\theta_t^{x,b} \geq 0\) because it will be automatically satisfied in equilibrium. However, for \(\psi = 0\) (narrow banking), one has to keep the constraint \(\theta_t^{m,b} \geq 0\) as otherwise bankers could still issue monetary claims by issuing directly outside money. Furthermore, I impose here the even stricter condition \(\theta_t^{m,b} = 0\) to be able to incorporate the model solution for the extension discussed in Section 7.1 that does not restrict \(\psi\) to be in \(\{0, 1\}\). Under the running assumption of \(\kappa_t(1 - \vartheta_t) \geq \eta_t\) (bankers are not so rich that they want to save in outside money in the aggregate), this additional restriction is inconsequential for the baseline model.\(^{82}\)

The following Lemma collects the relevant conclusions from the Hamiltonian maximization to solve the model.

**Lemma 4 (Individual Portfolio Choice).** Any optimal choice for households necessarily satisfies the following conditions

\[
\mu_t^{1-\eta} + \rho - \mu_t^d = \zeta_t^h (\sigma_t^{1-\eta} - \sigma_t^d) + \zeta_t^h \tilde{\sigma}_t^{n,h} + \lambda_t \nu_t, \tag{50}
\]

\[
\dot{\nu}_t = (1 - \psi) \lambda_t \nu_t, \tag{51}
\]

\[
\dot{\nu}_t - \mu_t^d \leq \zeta_t^h \left( -\frac{\sigma_t^d}{1 - \vartheta_t} \right) + \zeta_t^h \tilde{\sigma}_t + \lambda_t \nu_t, \tag{52}
\]

where the last condition holds with equality, whenever the constraint (48) does not bind, which is in equilibrium equivalent to \(\kappa_t < \bar{\kappa}\). Similarly, any optimal choice for bankers necessarily satisfies the following equations

\[
\mu_t^n + \rho - \mu_t^d - i_t^d = \zeta_t^b (\sigma_t^n - \sigma_t^d) + \zeta_t^b \tilde{\sigma}_t^{n,b}, \tag{53}
\]

\[
\dot{i}_t^x - \mu_t^d - i_t^d = \zeta_t^b \left( -\frac{\sigma_t^d}{1 - \vartheta_t} \right) + \zeta_t^b \tilde{\sigma}_t. \tag{54}
\]

In these conditions, \(\mu_t^{1-\eta}\) and \(\sigma_t^{1-\eta}\) denote the (geometric) drift and \(dZ_t\)-volatility of \(1 - \eta_t\), \(\tilde{\sigma}_t^{n,h}\) is the \(d\tilde{Z}_t\)-volatility of \(n_t^h\) and \(\tilde{\sigma}_t^{n,b}\) is the \(d\tilde{Z}_t\)-volatility of \(n_t^b\).

Furthermore, the prices of risk are given by

\[
\zeta_t^b = \sigma_t^n, \quad \zeta_t^h = \sigma_t^{1-\eta}, \quad \zeta_t^b = \tilde{\sigma}_t^{n,b}, \quad \zeta_t^h = \tilde{\sigma}_t^{n,h}. \]

\(^{82}\)This assumption is irrelevant unless bankers strictly prefer to simultaneously hold outside money and issue deposits. For \(\psi \leq 1\), they never strictly prefer such a choice, so that it is only relevant for \(\psi > 1\).
the price of transaction services $\lambda_t$ satisfies the condition

$$
\begin{align*}
\lambda_t &= \frac{1}{a} \bar{\gamma}'(v_t) v_t, \quad \alpha < \infty \\
\min\{\lambda_t, \bar{a} - v_t\} &= 0, \quad \alpha = \infty
\end{align*}
$$

(55)

and the idiosyncratic net worth volatility loadings satisfy in equilibrium the equations

$$
\tilde{\sigma}_t^{n,b} = \frac{\kappa_t}{\eta_t} (1 - \vartheta_t) \beta \tilde{\sigma}_t \\
\tilde{\sigma}_t^{n,h} = \frac{1 - \kappa_t}{1 - \eta_t} (1 - \vartheta_t) \tilde{\sigma}_t.
$$

Proof. The proof proceeds along the same lines as the Hamiltonian maximization in the model without banks discussed in Section B.1.1. I start with the first order conditions for households’ choice of $(c_h, \theta_{m,h}^t, \theta_{k,h}^t, \theta_{x,h}^t, v_t)$. Again, let $\xi_t^h \tilde{n}_t^h \lambda_t$ be the Lagrange multiplier on equation (49). Also, let $\xi_t^h \tilde{n}_t^h \nu_t$ the Lagrange multiplier on the additional constraint (48). The first-order condition with respect to the portfolio weights $\theta_{m,h}^t, \theta_{k,h}^t,$ and $\theta_{x,h}^t$ are (after rearranging and dividing by $\xi_t^h \tilde{n}_t^h$)

$$
\tilde{\sigma}_t^{n,b} = \frac{\kappa_t}{\eta_t} (1 - \vartheta_t) \beta \tilde{\sigma}_t \\
\tilde{\sigma}_t^{n,h} = \frac{1 - \kappa_t}{1 - \eta_t} (1 - \vartheta_t) \tilde{\sigma}_t.
$$

The first equation is precisely equation (51) in the assertion. The condition (52) in the assertion is simply a restatement of the third first-order condition: the multiplier $\nu_t$ is always nonnegative, so that one can equivalently write the condition without the $\nu_t$ term and replace “=” by an inequality “≤”. Clearly, this inequality holds with equality precisely when $\nu_t = 0$, which is equivalent to a nonbinding constraint.

To the derive the remaining household condition (50), multiply the first first-order condition

The first equation is precisely equation (51) in the assertion. The condition (52) in the assertion is simply a restatement of the third first-order condition: the multiplier $\nu_t$ is always nonnegative, so that one can equivalently write the condition without the $\nu_t$ term and replace “=” by an inequality “≤”. Clearly, this inequality holds with equality precisely when $\nu_t = 0$, which is equivalent to a nonbinding constraint.

To the derive the remaining household condition (50), multiply the first first-order condition
by $-\theta_{t}^{m,h}$, the second by $\theta_{t}^{k,h}$, and the third by $\theta_{t}^{x,h}$ and then sum them up

$$\begin{align*}
&-\theta_{t}^{m,h} \frac{d\bar{q}_t}{q^h_t} + \theta_{t}^{k,h} \left( \frac{\rho}{1 - \vartheta_t} \frac{\xi_t^{\bar{v}_t}}{q^h_t} - \frac{\mu_{t}^{d}}{1 - \psi_t} - \int_t^{d} \right) + \theta_{t}^{x,h} \left( \int_t^{x} - \mu_{t}^{d} - \int_t^{d} \right) \\
&= \varsigma_t \left( \theta_{t}^{k,h} + \theta_{t}^{x,h} \right) \left( \frac{-}\sigma_{t}^{d} \right) + \varsigma_t \left( \theta_{t}^{k,h} + \theta_{t}^{x,h} \right) \tilde{\sigma}_t + \psi \lambda_t v_t \\
&\quad + \lambda_t \left( \frac{\alpha}{q^h_t} \theta_{t}^{k,h} - \psi_t \left( 1 - \psi \right) \theta_{t}^{m,h} + \psi_t \left( 1 - \theta_{t}^{k,h} - \theta_{t}^{x,h} \right) \right) \\
&\quad - \nu_t \left( \tilde{\rho} \theta_{t}^{k,h} + \theta_{t}^{x,h} \right)
\end{align*}$$

The terms in the third and fourth lines are zero because either the respective constraint binds (then the second factor vanishes) or the respective multiplier is zero (and thus the first factor vanishes). A comparison of the expression on the left-hand side with the net worth evolution (41) and the return processes in Lemma 3 reveals that the left-hand side expression equals

$$\mu_{t}^{\hat{n},h} + \rho - \mu_{t}^{d} = \varsigma_t \left( \sigma_{t}^{\hat{n},h} - \sigma_{t}^{d} \right) + \varsigma_t \hat{\sigma}_{t}^{\hat{n},h} + \int_t^{d} + \psi \lambda_t v_t.$$

Now use that $\hat{n}_t^{h} = \frac{n_t^{h}}{q^h_t} = (1 - \eta_t) \frac{n_t^{h}}{N_t^{h}}$ and thus $\mu_{t}^{\hat{n},h} = \mu_{t}^{1-\eta} = \sigma_{t}^{1-\eta}$, $\hat{\sigma}_{t}^{\hat{n},h} = \hat{\sigma}_{t}^{1-\eta}$, and $\tilde{\sigma}_{t}^{\hat{n},h} = \tilde{\sigma}_{t}^{1-\eta}$. In addition, use the already proven deposit rate equation (51). Then it is easy to see that the previous equation is equivalent to equation (50) stated in the assertion.

For households’ prices of risk, one shows precisely as in Section B.1.1 that the first-order conditions for the maximization of (45) with respect to $c_t^{h}$ imply the SDF equation

$$\xi_t^{h} = e^{-\rho t} \frac{1}{\rho \hat{n}_t^{h}}.$$

Further, applying Ito’s formula to the right-hand side and comparing terms with the evolution (47) implies

$$\varsigma_t^{h} = \sigma_{t}^{\hat{n},h} = \sigma_{t}^{1-\eta} = \hat{\sigma}_{t}^{\hat{n},h} = \tilde{\sigma}_{t}^{\hat{n},h},$$

where in each case the validity of the second equality has already been justified earlier in this proof.

The equation (55) for the value of transaction services $\lambda_t$ is derived in precisely the same ways as in Section B.1.1 by maximizing the household Hamiltonian (45) with respect to $v_t$. 

81
Finally, households’ idiosyncratic net worth risk is

\[ \tilde{\sigma}_t^{n,h} = \left( \theta_t^{k,h} + \theta_t^{x,h} \right) \tilde{\sigma}_t = \frac{1 - \kappa_t}{1 - \eta_t} \left( 1 - \vartheta_t \right) \tilde{\sigma}_t, \]

where the second equality uses equations (17) and (18). This concludes the proof of all asserted equations relating to the household side.

For banker-related conditions, one proceeds analogously. By the remarks prior to this Lemma, \( \theta_t^{m,b} \) is constrained to be zero, so bankers have only two choice variables, \( c_t^b \) and \( \theta_t^{x,b} \), and do not face any constraints. The first-order condition for the maximization of the Hamiltonian (46) with respect to \( \theta_t^{x,b} \) is (after rearranging and dividing by \( \xi_t^{b,d_t} \))

\[ \tilde{r}_t^x - \mu_t^d - \tilde{r}_t^d = \xi_t^b \left( -\frac{\sigma_t^d}{1 - \vartheta_t} \right) + \xi_t^b \beta_t \tilde{\sigma}_t, \]

which is precisely equation (54). For a proof of the remaining equation (53), multiply this first-order condition with \( \theta_t^{x,b} \),

\[ \theta_t^{x,b} \left( \tilde{r}_t^x - \mu_t^d - \tilde{r}_t^d \right) = \xi_t^b \left( -\theta_t^{x,b} \frac{\sigma_t^d}{1 - \vartheta_t} \right) + \xi_t^b \theta_t^{x,b} \beta_t \tilde{\sigma}_t, \]

where the three expressions under the original terms follow as for households from the net worth evolution (42) together with the portfolio returns stated in Lemma 3. Again, \( \mu_t^{\hat{n},b}, \sigma_t^{\hat{n},b}, \tilde{\sigma}_t^{\hat{n},b} \) refer to the geometric drift and volatility loadings of the process \( \hat{n}_t \). Similarly as before, \( \hat{n}_t = \frac{n_t}{q_t \kappa_t} = \frac{\eta_t n_t}{N_t} \) implies \( \tilde{\mu}_t^b = \tilde{\mu}_t^n, \tilde{\sigma}_t^b = \tilde{\sigma}_t^n, \) and \( \tilde{\sigma}_t^{\hat{n},b} = \tilde{\sigma}_t^{\hat{n},b} \). Substituting these expressions into the previous equation yield equation (53).

To derive the price of risk expressions for bankers, one shows as for households that maximizing the Hamiltonian (46) with respect to \( c_t^b \) yields the condition

\[ \xi_t^b = e^{-\rho t} \frac{1}{\rho \hat{n}_t^b} \]

for the SDF process. Applying again Ito’s formula to the right-hand side and comparing terms with the evolution (47) implies then

\[ \xi_t^b = \sigma_t^{\hat{n},b} = \sigma_t^n, \tilde{\xi}_t^b = \tilde{\sigma}_t^{\hat{n},b} = \tilde{\sigma}_t^n. \]
Finally, bankers’ idiosyncratic net worth risk is

$$\tilde{\sigma}_{t}^{n,b} = \theta_t^{n,b} \tilde{\sigma}_t = \frac{\kappa_t}{\eta_t} (1 - \vartheta_t) \tilde{\sigma}_t,$$

where the second equality uses equation (18). This completes the proof of the lemma.

**Money Valuation Equation (Proposition 5).** The money valuation equation follows from the portfolio choice conditions in Lemma 4. Multiply equation (53) by $\eta_t$, equation (50) by $1 - \eta_t$, sum the two, use $\eta_t \mu_t^\vartheta + (1 - \eta_t) \mu_t^{1-\eta} = 0$, and solve for $\mu_t^\vartheta$

$$\mu_t^\vartheta = \rho - \left( \eta_t \varsigma_t^b (\sigma_t^\eta - \sigma_t^{\vartheta}) + (1 - \eta_t) \varsigma_t^h (\sigma_t^{1-\eta} - \sigma_t^{\vartheta}) + \eta_t \varsigma_t^{n,b} + (1 - \eta_t) \varsigma_t^{n,h} + \eta_t \lambda_t \nu_t \right).$$

(56)

This equation is already close to equation (5) and in particular, this form justifies the interpretation given in the main text of the terms as being the wealth-weighted sums of aggregate risk premia, idiosyncratic risk premia and liquidity premia. To obtain the equation stated in Proposition 5, substitute in the price of risk expressions, the expressions for idiosyncratic net worth risk and the deposit rate from Lemma 4. The aggregate risk premium term then becomes

$$\eta_t \varsigma_t^b (\sigma_t^\eta - \sigma_t^{\vartheta}) + (1 - \eta_t) \varsigma_t^h (\sigma_t^{1-\eta} - \sigma_t^{\vartheta}) = \eta_t \sigma_t^\eta (\sigma_t^\eta - \sigma_t^{\vartheta}) + (1 - \eta_t) \sigma_t^{1-\eta} (\sigma_t^{1-\eta} - \sigma_t^{\vartheta})$$

$$= \eta_t (\sigma_t^\eta)^2 + (1 - \eta_t) \left( \frac{\eta_t \sigma_t^\eta}{1 - \eta_t} \right)^2$$

$$= \eta_t (\sigma_t^\eta)^2 + (1 - \eta_t) \left( \frac{\eta_t \sigma_t^\eta}{1 - \eta_t} \right)^2$$

$$= \frac{\eta_t (\sigma_t^\eta)^2}{1 - \eta_t} = \frac{(\eta_t \sigma_t^\eta)^2}{\eta_t (1 - \eta_t)},$$

where the second equality uses $\eta_t \sigma_t^\eta + (1 - \eta_t) \sigma_t^{1-\eta} = 0$ and the third uses $\sigma_t^{1-\eta} = \frac{\eta_t \sigma_t^\eta}{1 - \eta_t}$ (follows from the previous equation). The idiosyncratic risk premium is

$$\eta_t \varsigma_t^{n,b} \tilde{\sigma}_t^{n,b} + (1 - \eta_t) \varsigma_t^{n,h} \tilde{\sigma}_t^{n,h} = \eta_t \left( \tilde{\sigma}_t^{n,b} \right)^2 + (1 - \eta_t) \left( \tilde{\sigma}_t^{n,h} \right)^2$$

$$= \eta_t \left( \frac{\kappa_t}{\eta_t} (1 - \vartheta_t) \beta \tilde{\sigma}_t \right)^2 + (1 - \eta_t) \left( \frac{1 - \kappa_t}{1 - \eta_t} (1 - \vartheta_t) \tilde{\sigma}_t \right)^2$$

$$= \left( \frac{\kappa_t^2 \beta^2}{\eta_t^2} + \frac{1 - \kappa_t^2}{1 - \eta_t} \right) (1 - \vartheta_t)^2 \tilde{\sigma}_t^2.$$
Finally, the liquidity premium is

\[ \eta_t i_t^d + (1 - \eta_t) \lambda_t v_t = \eta_t (1 - \psi) \lambda_t v_t + (1 - \eta_t) \lambda_t v_t = (1 - \psi \eta_t) \lambda_t v_t. \]

Consequently, all three premia are as stated in equation (5).

**Quantity Equation (Proposition 6).** The equation (55) stated in Lemma 4 for the determination of \( \lambda_t \) is identical to equation (29) in Proposition 6. To complete the proof of that proposition, it is therefore only left to show equation (28):

\[
v_t = \frac{a}{q_t^K} \theta_t^{k,h} + \psi \theta_t^{d,h} \quad \text{(equation 9)}
\]

\[
= \frac{a}{1 + \phi a} \frac{1 - \theta_t + \phi \rho}{1 - \theta_t} \theta_t^{k,h} + \psi \theta_t^{d,h} \quad \text{(Proposition 4)}
\]

\[
= \frac{a}{1 + \phi a} \frac{1 - \theta_t + \phi \rho}{1 - \theta_t} \frac{1 - \theta_t}{1 - \eta_t} + \psi \left( 1 - \frac{1}{1 - \eta_t} + \frac{\kappa_t (1 - \theta_t)}{1 - \eta_t} \right) \quad \text{(equations (17), (18), (19))}
\]

\[
= \frac{a}{1 + \phi a} \frac{1 - \theta_t + \phi \rho}{1 - \theta_t + \psi (\kappa_t (1 - \theta_t) - \eta_t)}.
\]

**Capital Allocation (Proposition 7).** Combining conditions (52) and (54) from Lemma 4 implies the inequality

\[ \varsigma_t^h \left( - \frac{\sigma_t \theta_t}{1 - \theta_t} \right) + \varsigma_t^h \tilde{\sigma}_t \lambda_t v_t + \lambda_t v_t \geq \varsigma_t^b \left( - \frac{\sigma_t \theta_t}{1 - \theta_t} \right) + \varsigma_t^b \beta \tilde{\sigma}_t + i_t^d \]

and this has to hold with equality, whenever \( \kappa_t < \bar{\kappa} \). Because the weak inequality \( \kappa_t \leq \bar{\kappa} \) has to hold always due to the skin-in-the-game constraint of households, the combined condition is a complementary slackness condition. By standard arguments, it is equivalent to the single equation

\[
\min \left\{ \left( \varsigma_t^h - \varsigma_t^b \right) \left( - \frac{\sigma_t \theta_t}{1 - \theta_t} \right) + \left( \varsigma_t^h - \beta \varsigma_t^b \right) \tilde{\sigma}_t + \lambda_t v_t - i_t^d, \bar{\kappa} - \kappa_t \right\} = 0.
\]

This is equation (30), except for the explicit presence of the prices of risk and the deposit rate in the version stated here. As for the money valuation equation, this version of the capital
allocation equation justifies the interpretation of the equation in the main text discussion following Proposition 7. To bring it into the form as stated in the proposition, one substitutes in the prices of risk and deposit rate from Lemma 4. With these replacements, the aggregate risk premium differential becomes

$$\left(\varsigma_t^h - \varsigma_t^b\right) \left(-\frac{\sigma_t^\varphi}{1 - \theta_t} \right) = \left(\varsigma_t^{1-\eta} - \varsigma_t^\eta\right) \left(-\frac{\sigma_t^\varphi}{1 - \theta_t} \right)$$

$$= -\sigma_t^\eta \left(\frac{\eta_t}{1 - \eta_t} + 1\right) \left(-\frac{\sigma_t^\varphi}{1 - \theta_t} \right)$$

$$= \frac{\sigma_t^\eta}{1 - \eta_t} \frac{\sigma_t^\varphi}{1 - \theta_t},$$

the idiosyncratic risk premium differential becomes

$$\left(\tilde{\varsigma}_t^h - \beta \varsigma_t^b\right) \tilde{\sigma}_t = \left(1 - \kappa_t \right) \left(\frac{\eta_t}{1 - \eta_t} \right) \left(-\frac{\sigma_t^\varphi}{1 - \theta_t} \right) \tilde{\sigma}_t$$

$$= \left(1 - \kappa_t \right) \left(\frac{\eta_t}{1 - \eta_t} \right) \left(-\frac{\sigma_t^\varphi}{1 - \theta_t} \right) \tilde{\sigma}_t,$$

and the liquidity premium differential becomes

$$\lambda_t \psi_t - \nu_t^d = \lambda_t \psi_t - (1 - \psi) \lambda_t \psi_t$$

$$= \psi \lambda_t \psi_t,$$

all in line with equation (30).

**Evolution of the Endogenous State Variable (Proposition 8).** \(\eta_t = \hat{N}_t^b\) is simply aggregate bank net worth measured in the total wealth numeraire. From the law of motion of banker net worth and the return processes, it is thus immediate that \(\eta_t\) follows an Ito process.

The drift of \(\eta_t\) is obtained by substituting the money valuation equation in the version (56) into the condition (53) from Lemma 4:

$$\mu_t^\eta = \varsigma_t^b (\sigma_t^\varphi - \sigma_t^\varphi) + \varsigma_t^b \tilde{\sigma}_t^{n,b} + \nu_t^d$$

$$- \left(\eta_t \varsigma_t^h (\sigma_t^{1-\eta} - \sigma_t^\varphi) + \eta_t \varsigma_t^h \tilde{\sigma}_t^{n,h} + (1 - \eta_t) \varsigma_t^h \tilde{\sigma}_t^{n,h} + \eta_t \nu_t^d + (1 - \eta_t) \lambda_t \psi_t\right)$$

$$= (1 - \eta_t) \left(\varsigma_t^h (\sigma_t^{1-\eta} - \sigma_t^\varphi) + \varsigma_t^b \tilde{\sigma}_t^{n,b} - \varsigma_t^h \tilde{\sigma}_t^{n,h} + \nu_t^d - \lambda_t \psi_t\right).$$
To bring this equation into the form stated in Proposition 8, substitute in the expressions for \( \xi_i^t, \xi_i^t, \tilde{\sigma}^{n,i}_t (i \in \{h,b\}) \) and \( i_t^d \) from Lemma 4. The aggregate risk term then becomes

\[
\xi^b_t (\sigma^\eta_t - \sigma^\vartheta_t) - \xi^h_t (\sigma^1_t - \eta^1_t - \sigma^\eta_t) = (\sigma^\eta_t)^2 - \left( \frac{\eta_t}{1 - \eta_t} \right)^2 \sigma^\eta_t + \left( \frac{\eta_t}{1 - \eta_t} + 1 \right) \sigma^\eta_t \sigma^\vartheta_t
\]

\[
= \left( \frac{1 - 2\eta_t}{1 - \eta_t} \right)^2 \sigma^\eta_t - \frac{1}{1 - \eta_t} \sigma^\eta_t \sigma^\vartheta_t
\]

\[
= \frac{1}{1 - \eta_t} \left( \frac{1 - 2\eta_t}{1 - \eta_t} \sigma^\eta_t - \sigma^\vartheta_t \right) \sigma^\eta_t,
\]

the idiosyncratic risk term becomes

\[
\xi^b_t \tilde{\sigma}^{n,b}_t - \xi^h_t \tilde{\sigma}^{n,h}_t = \left( \tilde{\sigma}^{n,b}_t \right)^2 - \left( \tilde{\sigma}^{n,h}_t \right)^2
\]

\[
= \left( \frac{\kappa_t}{\eta_t} \left( 1 - \vartheta_t \right) \beta \tilde{\sigma}_t \right)^2 - \left( \frac{1 - \kappa_t}{1 - \eta_t} \left( 1 - \vartheta_t \right) \tilde{\sigma}_t \right)^2
\]

\[
= \left( \frac{\kappa_t}{\eta_t} \beta \right)^2 - \left( \frac{1 - \kappa_t}{1 - \eta_t} \right)^2 \left( 1 - \vartheta_t \right)^2 \tilde{\sigma}_t^2
\]

and the liquidity term becomes

\[
i_t^d - \lambda_t v_t = (1 - \psi) \lambda_t v_t - \lambda_t v_t
\]

\[
= -\psi \lambda_t v_t.
\]

For the volatility of \( \eta_t \), note that it must coincide with the aggregate risk loading of \( \hat{n}_t^b \).
banker wealth measures in the total wealth numeraire. This loading is given by

\[
\sigma_t^\eta = \theta_t^{d,b} \sigma_t^\eta - \theta_t^{x,b} \frac{\partial_t \sigma_t^\eta}{1 - \vartheta_t}
\]

\[
= \sigma_t^\eta - \theta_t^{x,b} \left( \sigma_t^\eta + \frac{\partial_t \sigma_t^\eta}{1 - \vartheta_t} \right)
\]

\[
= \sigma_t^\eta - \frac{\kappa_t}{\eta_t} (1 - \vartheta_t) \frac{\sigma_t^\eta}{1 - \vartheta_t}
\]

\[
= -\frac{\kappa_t - \eta_t}{\eta_t} \sigma_t^\eta,
\]

which is equivalent to the equation stated in Proposition 8.

B.2 Numerical Solution Algorithm

In this appendix, I describe the numerical solution procedure of the model.

Computational of the Model Solution. A model solution can be completely described by

the two functions

\[ \vartheta, \kappa : (0, \infty) \times (0, 1) \rightarrow \mathbb{R}, \]

as these are sufficient to characterize all other equilibrium functions in Definition 1:84

1. \( \tilde{c}_b(s, \eta) = \tilde{c}_h(s, \eta) = \rho; \)

2. \( p, q \) and \( \iota \) are functions of \( \vartheta \) by Proposition 4;

3. \( \sigma^\eta(s, \eta) \) is determined by the risk generation equation in Proposition 11 as a function of

\( \vartheta(s, \eta), \partial_x \vartheta(s, \eta), \partial_\eta \vartheta(s, \eta) \) and \( \kappa(s, \eta) \);

4. using Ito’s formula, \( \sigma^\eta \) is given by

\[
\sigma^\eta(s, \eta) = \partial_x \vartheta(s, \eta) \sigma(s, \eta) + \partial_\eta \vartheta(s, \eta) \sigma^\eta(s, \eta) \eta;
\]

83 This follows directly from the definition of \( \hat{a}^b_t \) and it also follows from \( \sigma^h_t = \vartheta_t = \sigma^\eta_t \), see the proof of Lemma 4.

84 The money portfolio shares \( \theta^{m,b} \) and \( \theta^{m,h} \) are missing in the following list. This is because the algorithm is based on the assumption that only households hold outside money, i.e. \( \theta^{m,b} = 0, \theta^{m,h} = \frac{\vartheta}{1 - \vartheta} \). This is always the case if in the aggregate the running assumption \( \kappa(1 - \vartheta) \geq \eta \) holds. The validity of this inequality over the model’s ergodic distribution has to be verified once the model is solved.
5. $v(s, \eta)$ is determined by the quantity equation in Proposition 6 as a function of $\vartheta(s, \eta)$, $\kappa(s, \eta)$ and $\lambda(s, \eta)v(s, \eta) = \frac{\nu - 1}{\alpha} (v(s, \eta)/\bar{v})$ can then easily be computed;

6. $\mu^d(s, \eta)$ is a function of $\sigma^d(s, \eta)$, $\kappa(s, \eta)$, $\vartheta(s, \eta)$, $\lambda(s, \eta)v(s, \eta)$, $\bar{\sigma}(s)$, and $\eta$ by the money valuation equation in Proposition 5;

7. $\mu^u(s, \eta)$ is a function of $\sigma^u(s, \eta)$, $\sigma^d(s, \eta)$, $\kappa(s, \eta)$, $\vartheta(s, \eta)$, $\lambda(s, \eta)v(s, \eta)$, $\bar{\sigma}(s)$, and $\eta$ by Proposition 8.

This 7-step procedure explains how to use the given functions $\vartheta$ and $\kappa$ to compute both the function $\mu^g(s, \eta)$ and all the equilibrium functions in Definition 1 with the exception of $r^x$ and $i^d$. I exclude the functions $r^x$ and $i^d$ because they are not explicitly needed to compute the model solution and they are also not of independent interest.\footnote{It is not hard to compute these functions from the solved model using appropriate portfolio choice conditions, compare Lemma 4. For $r^x$, one also has to revert the numeraire change.} I now turn to the solution procedure that solves for $\vartheta$ and $\kappa$. The first of these functions, $\vartheta$, is determined by the dynamic equation (BSDE) stated in Proposition 5. I solve for it using an iterative procedure. The remaining function $\kappa$ is determined by the capital allocation condition (30), which is an algebraic equation that has to be solved in each iteration.

To obtain an iterative procedure for the dynamic equation, I transform the problem into a finite horizon problem with some arbitrary terminal condition for $\vartheta \in (0, 1)$.\footnote{For the numerical solutions in this paper, I have used the steady-state solution, but this should ultimately not matter because the convergence behavior is not very sensitive to the choice of the terminal condition, provided the PDE algorithm uses a monotone scheme and a sufficiently small time step.} Then the model equilibrium solution function $\vartheta$ on the finite time domain depends on three states, $s$, $\eta$, and $t$, and is thus represented by a function $\vartheta(s, \eta, t)$. Applying Ito’s formula to $\vartheta$ implies

$$
\begin{align*}
\text{d} \vartheta (s_t, \eta_t, t) = & \left( \partial_t \vartheta (s_t, \eta_t, t) + \mu_s(s_t) \partial_s \vartheta (s_t, \eta_t, t) + \mu^u(s_t, \eta_t, t) \partial_\eta \vartheta (s_t, \eta_t, t) \\
+ & \frac{1}{2} \left( \sigma_s(s_t) \right)^2 \partial_{ss} \vartheta (s_t, \eta_t, t) + \frac{1}{2} \left( \sigma^u(s_t, \eta_t, t) \eta_t \right)^2 \partial_{\eta\eta} \vartheta (s_t, \eta_t, t) \\
+ & \sigma_s(s_t) \sigma^u(s_t, \eta_t, t) \eta_t \partial_{s \eta} \vartheta (s_t, \eta_t, t) \right) dt \\
+ & \left( \sigma_s(s_t) \partial_s \vartheta (s_t, \eta_t, t) + \sigma^u(s_t, \eta_t, t) \eta_t \partial_\eta \vartheta (s_t, \eta_t, t) \right) dZ_t.
\end{align*}
$$

Equating the drift of this evolution with the drift expression from the BSDE for the variable,
\[ \mu^\vartheta(s_t, \eta_t, t), \] one obtains a representation as a partial differential equation,
\[
\begin{align*}
\partial_t \vartheta(s, \eta, t) + \mu_s(s) \partial_s \vartheta(s, \eta, t) + \mu^\eta(s, \eta, t) \partial_\eta \vartheta(s, \eta, t) \\
+ \frac{1}{2} (\sigma^\vartheta(s, \eta, t) \eta)^2 \partial_\eta \vartheta(s, \eta, t)
\end{align*}
\]
\[+ \sigma_s(s) \sigma^\eta(s, \eta, t) \eta \partial_\eta \vartheta(s, \eta, t) = \mu^\vartheta(\eta, t) x(\eta, t). \tag{57} \]

The model solution algorithm solves this partial differential equation backward in time, starting from the terminal guesses for \( \vartheta \) and iterating until the implied time derivative \( \partial_t \vartheta \) is approximately zero. If the procedure converges, then the resulting function (approximately) solves the original infinite-horizon equilibrium problem.

In sum, the model can be solved using the following algorithm:

1. **Preparation step.** Fix finite grids \( S = \{ s_i \mid i = 1, \ldots, n_s \} \), \( H = \{ \eta_i \mid i = 1, \ldots, n_\eta \} \) with \( 0 < s_1 < \cdots < s_{n_s} < 1 \) and \( 0 < \eta_1 < \cdots < \eta_{n_\eta} < 1 \) and a time increment \( \Delta t > 0 \). All functions are represented as matrices of tabled values on the product grid \( S \times H \) with linear interpolation off the grid. Also, choose a terminal guess \( (\vartheta(s, \eta))_{s \in S, \eta \in H} \) for the solution function of the dynamic equation on the grid.

2. For all grid points \((s, \eta) \in S \times H\), solve the static equations (30) for \( \kappa(s, \eta) \). In the process, it is necessary to follow the seven-step procedure outlined above to determine other variables in these equations as functions of \( \vartheta(s, \eta) \), \( \partial_s \vartheta(s, \eta) \), \( \partial_\eta \vartheta(s, \eta) \) and \( \kappa(s, \eta) \). Where needed, approximate derivatives \( \partial_s \vartheta \) and \( \partial_\eta \vartheta \) by finite differences using the matrix \( (\vartheta(s, \eta))_{s \in S, \eta \in H} \) from the previous iteration (this is an explicit method).

3. Use the seven-step procedure outlined above to compute \( (\mu^\vartheta(s, \eta))_{s \in S, \eta \in H}, (\sigma^\vartheta(s, \eta))_{s \in S, \eta \in H}, \) and \( (\mu^\vartheta(s, \eta))_{s \in S, \eta \in H} \). Also compute \( (\mu_s(s))_{s \in S, \eta \in H} \) and \( (\sigma_s(s))_{s \in S, \eta \in H} \).

4. Update the functions \( \vartheta \) on the grid \( S \times H \) by a single backward time step of length \( \Delta t \) in the linear PDEs (57) using an implicit Euler method, an upwind finite difference discretization for first derivatives, and the finite difference scheme described at the end of this section for second derivatives. To avoid the need for boundary conditions, replace \( \sigma^\eta(s, \eta_1) \eta_1 \) and \( \sigma^\eta(s, \eta_{n_\eta}) \eta_{n_\eta} \) (for all \( s \)) as well as \( \sigma_s(s_1) \) and \( \sigma_s(s_{n_s}) \) with zero.

---

87 Practically, these last two matrices can be pre-computed because they do not change over the iterations.

88 For “reasonable” initial guesses and grid choices, drift dynamics are always inward-pointing in all iterations so that there is no need to manipulate the drifts \( \mu^\vartheta \) and \( \mu_s \). One should check that replacing the outermost volatilities with zero results in only minor adjustments of their values. If this is not the case, grip points need to be placed...
5. If the update of \((\kappa(s, \eta))_{s \in S, \eta \in H}\) in step 2 and of \((\vartheta(s, \eta))_{s \in S, \eta \in H}\) in step 4 induces only small changes in all components, terminate. Otherwise, continue with step 2, using the updated matrix \((\vartheta(s, \eta))_{s \in S, \eta \in H}\) as a new terminal condition.

**Remark: Solution of the Model without Banks.** The previous solution procedure can also be used to solve the model without banks discussed in Section 4. In this case, one only has to solve for the one-dimensional function \(\vartheta(s)\) and sets everywhere in the solution approach above \(\eta = \kappa = \sigma^\eta = \mu^\eta = 0\). Practically, this case is much simpler because the problem reduces to solving a time-dependent PDE with only a single spatial dimension.

**Computation of the Ergodic Distribution.** Given the dynamics \(\mu^\eta(s, \eta), \sigma^\eta(s, \eta)\) of the endogenous state variable, which are computed in the solution algorithm above, I compute the ergodic density \(f(s, \eta)\) of the two-dimensional state \((s, \eta)\) by solving the stationary Kolmogorov forward equation

\[
0 = -\partial_{\eta} \left( \mu^\eta(s, \eta) f(s, \eta) \right) - \partial_{s} \left( \mu_s(s) f(s, \eta) \right) + \partial_{\eta} \left( \sigma_s(s) \sigma^\eta(s, \eta) f(s, \eta) \right) + \frac{1}{2} \partial_{\eta\eta} \left( (\sigma^\eta(s, \eta))^2 f(s, \eta) \right) + \frac{1}{2} \partial_{ss} \left( (\sigma_s(s))^2 f(s, \eta) \right)
\]

for a solution that satisfies the additional condition \(\int_0^\infty \int_0^1 f(s, \eta) d\eta ds = 1\). Practically, I do this by using the fact that the differential operator in the Kolmogorov forward equation is the adjoint of the differential operator in equation (57) without the time derivative, so that its discretization on the grid \(S \times H\) can be obtained by using the discretization for equation (57) in the PDE time step and taking the transpose. After discretization, the Kolmogorov forward equation becomes a homogeneous linear equation system. I use standard linear algebra methods to compute a base of the kernel of the matrix of that equation system, then verify that the kernel is one-dimensional and find the density by integrating the kernel vector along both dimensions (using a trapezoid method) and dividing by the total mass.

**A Monotone Finite Difference Scheme for the Second-order Terms.** My partial differential equation (57) contains second-order cross derivatives \(\partial_{\eta\eta}\vartheta\) which are notoriously problematic for finite difference methods because the most straightforward discretization schemes do not result in monotone schemes and are therefore inadequate to produce a convergent solution closer to the boundaries. For the boundary point \(s_{n_s}\) and a Heston model, this criterion is infeasible because volatility increases in \(s\). Instead, one should make \(s_{n_s}\) sufficiently large such that dynamics are very unlikely to ever end up there.
method. I circumvent this issue by using in each grid point a local coordinate rotation that eliminates the positive coefficient in front of the cross terms. This generically requires taking finite differences in directions that are not the original coordinate direction, which is achieved by an interpolation method that is guaranteed to put always nonnegative weight on all grid points and thus preserves monotonicity of the scheme.

To provide more detail, I simplify notation and give the two coordinate directions the generic names $x$ and $y$, the grids are denoted by $X = \{x_i \mid i = 1,\ldots,n_x\}$, $Y = \{y_i \mid i = 1,\ldots,n_y\}$. Let $\Upsilon := (x, y)^T$ denote the two-dimensional state vector and consider a fixed inner grid point $\Upsilon_0 := (x_i, y_j)$ with $1 < i < n_x$, $1 < j < n_y$. The goal is to create a (monotone) finite difference scheme for the differential operator

$$L(\Upsilon) := \frac{1}{2} \sigma_x(\Upsilon)^2 \partial_{xx} + \frac{1}{2} \sigma_y(\Upsilon)^2 \partial_{yy} + \sigma_x(\Upsilon) \sigma_y(\Upsilon) \partial_{xy}$$

at the grid point $\Upsilon_0$. Define the vector $\sigma_{\Upsilon,0} := (\sigma_x(\Upsilon_0), \sigma_y(\Upsilon_0))^T$. Then the operator $L_0 := L(\Upsilon_0)$ can be written more concisely as

$$L_0 = \frac{1}{2} \nabla^T_\Upsilon (\sigma_{\Upsilon,0} \sigma_{\Upsilon,0}^T) \nabla_\Upsilon,$$

where $\nabla_\Upsilon = (\partial_x, \partial_y)^T$ and I use the convention that $\partial_{z_1} \cdot \partial_{z_2} = \partial_{z_1 z_2}$, $z_1, z_2 \in \{x, y\}$, for the product of derivative operators. Because matrix multiplication is associative, one can also write this as

$$L_0 = \frac{1}{2} (\nabla^T_\Upsilon \sigma_{\Upsilon,0}) (\sigma_{\Upsilon,0}^T \nabla_\Upsilon) = \frac{1}{2} (\sigma_{\Upsilon,0}^T \nabla_\Upsilon)^2.$$

The advantage of this representation is that the scalar differential operator $\sigma_{\Upsilon,0}^T \nabla_\Upsilon = \sigma_x(\Upsilon_0) \partial_x + \sigma_y(\Upsilon_0) \partial_y$ can be interpreted as a first-order directional derivative $\partial_v$ in the direction $v := \sigma_{\Upsilon,0}$ (whereas $\partial_x$ is the directional derivative in the direction $(1, 0)^T$ and $\partial_y$ is the directional derivative in the direction $(0, 1)^T$). Specifically, $L_0 = \frac{1}{2} \partial_v$ is a second-order differential operator that takes derivatives only in the direction $v$ without any cross terms. One can therefore discretize $L_0$ by a standard one-dimensional second-order finite difference formula in the direction of $v$ and this will result in a monotone scheme:

$$L_0 \vartheta(\Upsilon_0) \approx \frac{\vartheta(\Upsilon_0) - \vartheta(\Upsilon_0)}{||\Upsilon_+ - \Upsilon_-||} \cdot ||\Upsilon_+ - \Upsilon_0|| + \frac{\vartheta(\Upsilon_0) - \vartheta(\Upsilon_0)}{||\Upsilon_+ - \Upsilon_-||} \cdot ||\Upsilon_0 - \Upsilon_-||,$$

\[\text{\cite{90}}\]

There is no need to discuss the discretization of second-order derivatives at boundary grid points because I set the volatilities at boundary grid points to zero, compare the algorithm description above.

\[\text{\cite{90}}\]

Formally, this is not conventional matrix multiplication for matrices over the real numbers but for matrices over the algebra of partial differential operators. Associativity holds of course also in this case.
where \( \Upsilon_+ := \Upsilon_0 + \Delta_+ \cdot v \), \( \Upsilon_- := \Upsilon_0 - \Delta_- \cdot v \) with some suitable positive numbers \( \Delta_+, \Delta_- > 0 \).

An issue with this approach is that the points \( \Upsilon_+ \) and \( \Upsilon_- \) are usually not on the grid \( X \times Y \) for any choice of \( \Delta_+ \) and \( \Delta_- \) unless \( v \) points in one of the coordinate directions. In the model presented in this paper, the latter is virtually never the case because both components of \( v = \sigma_T \cdot \Upsilon_0 \) are generally nonzero. This issue is overcome by using linear interpolation to compute \( \vartheta(\Upsilon_+\Upsilon_-) \) in the scheme. Specifically, I consider the rectangular area in the state space around \( \Upsilon_0 \) that is defined as the convex hull of the nine grid points \( (x_{i+k}, y_{j+\ell}) \) with \( k, \ell \in \{-1, 0, 1\} \). Call this area \( A \). The rays \( R_+ := \{ \Upsilon_0 + \Delta_+ \cdot v \mid \Delta_x \geq 0 \} \) and \( R_- := \{ \Upsilon_0 - \Delta_- \cdot v \mid \Delta_x \geq 0 \} \) each intersect the boundary of \( A \) at precisely one point, \( |R_+ \cap \partial A| = |R_- \cap \partial A| = 1 \). I use these intersection points as my points \( \Upsilon_+ \) (intersection \( R_+ \cap \partial A \)) and \( \Upsilon_- \) (intersection \( R_- \cap \partial A \)) in the finite-difference representation of \( L_0 \). Because \( A \) is a convex hull of finitely many points, it is closed, so in particular it contains all of its boundary \( \partial A \). Thus, \( \Upsilon_+ \) and \( \Upsilon_- \) can be written as convex combinations of the nine grid points \( (x_{i+k}, y_{j+\ell}) \). All these convex combinations define valid linear interpolation schemes with nonnegative weights on the function values \( \vartheta(x_{i+k}, y_{j+\ell}) \) on the grid. They would therefore in principle all result in monotone finite difference schemes. For my specific algorithm, I choose for each of the points \( \Upsilon_\pm \) the unique representation as a convex combination of the two grid points closest to \( \Upsilon_\pm \).

\[ \text{C Details on the Model Calibration} \]

In this appendix, I provide details on the model calibration briefly discussed in Section 5.3. The model has twelve parameters, the parameters \( b, \sigma_s \) and \( \bar{s} \) for the evolution of the exogenous risk state \( s \), the transaction cost parameters \( \alpha \) and \( \bar{v} \), capital productivity \( a \), the time preference \( \rho \), the capital adjustment cost parameter \( \phi \), the parameters \( \beta \) and \( \bar{\kappa} \) that determine the efficiency of banks’ diversification technology and the maximum share of capital that banks can hold, and the parameter \( \psi \) that determines the relative efficiency of bank deposits for making transactions. In addition, the choice of \( \psi > 1 \) is really a consequence of the mapping from a model with government bonds to the model without as described in Proposition 15, so that one has to invert this mapping when calibrating \( \psi \) and \( \bar{v} \) (all other parameters remain unaffected by the transformation in Proposition 15).

\[ \text{C.1 Calibration of} \ b, \sigma_s \text{ and} \ \bar{s} \]

I calibrate the exogenous process parameters \( b, \sigma_s \) and \( \bar{s} \) separately by matching the evidence on establishment-level idiosyncratic TFP shocks reported by Bloom et al. (2018). The data in
that paper is based on the Census panel of manufacturing establishments in the U.S. from 1972 to 2011. The authors use their empirical measure of idiosyncratic shocks to calibrate the $\sigma_{Z_{t-1}}$ process for the idiosyncratic TFP evolution (their equation (3))

$$\log z_{j,t} = \rho Z \log z_{j,t-1} + \sigma_{Z_{t-1}} \varepsilon_{j,t}$$

with i.i.d. standard normal shocks $\varepsilon_{j,t}$. I follow that approach in my model by interpreting my idiosyncratic capital shocks as productivity shocks to installed capital instead of taking the quantity variation literally. I assume that these shocks only affect the productivity of currently installed capital of the agent, but not the quality of newly produced capital or capital goods purchased from other agents after the shock has materialized. It is then easy to see that my model is isomorphic to a model with such idiosyncratic productivity shocks. In my model, the productivity $\tilde{k}_t$ of an installed unit of capital over a year evolves according to

$$\log \tilde{k}_t = \log \tilde{k}_{t-1} + \int_{t-1}^{t} \tilde{\sigma}_\tau d\tilde{Z}_\tau - \frac{1}{2} \int_{t-1}^{t} \tilde{\sigma}_\tau^2 d\tau,$$

which is an evolution that is up to the Ito correction term identical to (58) if one sets $\rho Z = 1$.91 I can therefore follow literally the calibration strategy of Bloom et al. (2018). Specifically, these authors report in their table VI the (time-series) mean, standard deviation, skewness and serial correlation of the interquartile range of the cross-sectional distribution of their estimated micro-level shocks (the empirical counterparts of $\sigma_{Z_{t-1}} \varepsilon_{j,t}$ across different $j$) at the annual frequency. They fit the parameters of a pre-specified Markov chain for $\sigma_{Z_{t-1}}$ to match these time-series properties of the cross-sectional interquartile range of idiosyncratic shocks in the model.

I follow that approach by defining the idiosyncratic productivity shock of an individual agent $\tilde{i}$ over a year by

$$e_t(\tilde{i}) := \int_{t-1}^{t} \tilde{\sigma}_\tau d\tilde{Z}_\tau(\tilde{i})$$

and choosing the parameters $b$, $\sigma_s$ and $\bar{s}$ of my exogenous process $s_t$ that determines $\tilde{\sigma}_t = \sqrt{s_t}$ so that for a simulated sequence of shocks $e_t(\tilde{i})$ (simulated both over time $t$ and across agents $\tilde{i}$), the simulated time series of the interquartile range of the distribution along the $\tilde{i}$-dimension of $\{e_t(\tilde{i})\}$ matches three of the four moments reported in table VI of Bloom et al. (2018), the mean, the standard deviation and the autocorrelation.92 Specifically, for each parameter combination

91 Bloom et al. (2018) work with a specification with $\rho Z < 1$, but this is not directly relevant for the calibration of the process $\sigma Z$ in their setting and $\tilde{\sigma}$ in mine.

92 Given that I have only three parameters available, it is generally impossible to match all four moments. I
(b, σ_s, ̄s), I draw use the M equally spaced quantiles (the quantiles \( \frac{m}{M+1} \) for \( m = 1, \ldots, M \)) of the known ergodic distribution of s (a gamma distribution) as starting points for \( s_0 \) and simulate \( M \) independent aggregate time series \( \{s^m_{k\Delta t}\} \) for \( k = 1, \ldots, K \) and \( m = 1, \ldots, M \) using a standard Euler scheme with time discretization \( \Delta t > 0 \) such that \( T = K\Delta t \) is an integer and \( \{k\Delta t \mid k = 1, \ldots, K\} \) contains all integers from 1 to \( T = K\Delta t \). I next simulate for each of the \( M \) aggregate time paths \( \{s^m_{k\Delta t}\}_{k=1}^{K} \) \( N \) independent time series \( \{\hat{e}^m_{k\Delta t}\} \) for \( k = 1, \ldots, K, n = 1, \ldots, N \) and \( m = 1, \ldots, M \) using an Euler discretization of the evolution \( d\hat{e}_t := \sigma_t d\tilde{Z}_t \) with initial condition \( \hat{e}_0 = 0 \). I then compute the annual shocks in line with definition (59) by taking differences \( e^m_{t,n} := \hat{e}_t - \hat{e}_{t-1} \) for \( t = 1, \ldots, T \). For each given \((t,m)\), I compute the interquartile range of the empirical distribution of \( \{e^m_{t,n}\}_{n=1,\ldots,N} \), then compute the mean, standard deviation and autocorrelation of the resulting interquartile range over time \( (t) \) for each given \( m \) and, finally, I average these measures over the independent paths \( m \). I then compare the resulting statistics with the targets and vary \((b, \sigma_s, ̄s)\) to achieve a good fit. The results of my estimation procedure are shown in Table 1.

<table>
<thead>
<tr>
<th>Target</th>
<th>Data</th>
<th>Model</th>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>39.28</td>
<td>39.21</td>
<td>( b )</td>
<td>0.150</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>4.89</td>
<td>4.98</td>
<td>( \sigma_s )</td>
<td>0.037</td>
</tr>
<tr>
<td>Autocorrelation</td>
<td>0.75</td>
<td>0.73</td>
<td>( ̄s )</td>
<td>0.085</td>
</tr>
</tbody>
</table>

Table 1: Calibration of exogenous risk process

C.2 Calibration of \( \alpha \)

I calibrate the parameter \( \alpha \) to be in line with the empirical money demand literature (Ball, 2001, 2012; Ireland, 2009; Lucas, 1988; Stock and Watson, 1993). Papers in this literature usually start from a reduced-form money demand equation of the form

\[
\log \left( \frac{M}{P} \right) = \beta_0 + \beta_y \log Y + \beta_i \hat{i} + \varepsilon,
\]

where \( M \) is a monetary aggregate, mostly M1, \( P \) is a price index, \( Y \) a measure of output, e.g. GDP, and either \( \hat{i} = \log i \), then \( \beta_i \) is an elasticity, or \( \hat{i} = i \), then \( \beta_i \) is a semi-elasticity, where \( i \) is some nominal interest rate or interest rate differential that is supposed to proxy the opportunity

chose to ignore the skewness both because it is difficult to match with a CIR process and because it appears to be the least important in the context of my model due to the fact that the cross-sectional wealth (or productivity of installed capital) distribution is irrelevant for aggregate dynamics in my model.
cost of holding money. The empirical literature is interested in the overall fit of this specification (usually estimated over the long-run as a cointegration relationship), the parameters $\beta_y$ – the income-elasticity of money demand – and $\beta_i$ – the interest (semi-)elasticity of money demand – and their stability over time.

In my model, a similar equation holds. Specifically, output is $Y_t = aK_t$, the value of broad money including bank deposits is $\vartheta_t + \kappa_t (1 - \vartheta_t) - \eta_t M_t K_t = \vartheta_t + \kappa_t (1 - \vartheta_t) - \eta_t M_t K_t$, where $m_t := \vartheta_t + \kappa_t (1 - \vartheta_t) - \eta_t \vartheta_t$ denotes the money multiplier. The opportunity cost of holding money is the money premium $\lambda_t v_t = a(\vartheta_t - 1) \alpha (\overline{v_t})^{-1}$. Using the equations $q_t^M = \vartheta_t \frac{1 + \phi a}{1 - \vartheta_t + \phi \rho}$ and $v_t = \frac{1 - \vartheta_t + \phi \rho}{1 + \varphi a \vartheta_t + \vartheta_t \rho}$ (from Propositions 1 and 6), it is immediate that $v_t = m_t \vartheta_t q_t^M K_t$. Consequently,

$$
\log \left( m_t^M q_t^M K_t \right) = \log K_t - \log \frac{1}{m_t^M q_t^M} = \log aK_t - \log \frac{a}{m_t^M q_t^M} + \log \bar{v} = -\log \bar{v} + \log Y_t - \frac{1}{\alpha} \log (\vartheta_t^{-1})
$$

Therefore, if one chooses the log specification for the interest rate $\hat{i} = \log i$, then a version of equation (60) with $\beta_y = 1, \beta_i = -\frac{1}{\alpha}$ holds exactly state by state.

Motivated by this observation, I calibrate the parameter $\alpha$ to match the empirical estimates of the interest elasticity or semi-elasticity of money. To compare my (constant-elasticity) money demand specification to the estimates of semi-elasticities, I use the linear approximation $\log \lambda_t v_t \approx \frac{\lambda_t v_t - \lambda_t^{ss} v_t^{ss}}{\lambda_t^{ss} v_t^{ss}}$ around the steady state level $\lambda_t^{ss} v_t^{ss}$, so that the model-implied semi-elasticity in steady state is $-\frac{1}{\alpha} \lambda_t^{ss} v_t^{ss}$. Ball (2001, 2012), Ireland (2009), Lucas (1988), and Stock and Watson (1993) estimate the money demand in the U.S. and report interest semi-elasticities of money demand in the range from $-0.01$ to $-0.1$ (for interest rates measured in percent), with the more recent papers (Ball, 2001, 2012; Ireland, 2009) reporting smaller absolute numbers. Ireland (2009) also estimates a log specification and reports for two specifications interest elasticities of $-0.09$ and $-0.06$, respectively. With my target for the money premium $\lambda_t^{ss} v_t^{ss}$ (discussed below), these estimates imply values for $\alpha$ between 10 and 100. I choose for my calibration a value at the lower end of this range, $\alpha = 10$. Larger values bring the quantitative

---

93 My model implies an income elasticity of 1, irrespective of the chosen parameters. This is also in line with most of the empirical money demand literature with the exception of Ball (2001), who admits, however, that the data can be equally explained by a unit elasticity and steady progress in the payments technology, an interpretation which I adopt here.
Table 2: Calibration to steady-state targets

<table>
<thead>
<tr>
<th>steady-state targets</th>
<th>model counterpart</th>
<th>value</th>
<th>calibrated parameters</th>
<th>parameter</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>bank leverage</td>
<td>( \theta^{x,b} )</td>
<td>12.5</td>
<td>( \rho )</td>
<td>0.047</td>
<td></td>
</tr>
<tr>
<td>money multiplier</td>
<td>( m )</td>
<td>4.11</td>
<td>( \beta )</td>
<td>0.062</td>
<td></td>
</tr>
<tr>
<td>share of nominal safe assets</td>
<td>( \vartheta )</td>
<td>0.1</td>
<td>( \bar{\kappa} )</td>
<td>0.244</td>
<td></td>
</tr>
<tr>
<td>proportion of monetary safe assets</td>
<td>( 1/\psi )</td>
<td>0.65</td>
<td>( \bar{v} )</td>
<td>0.235</td>
<td></td>
</tr>
<tr>
<td>money premium</td>
<td>( \lambda \psi )</td>
<td>0.01</td>
<td>( \psi )</td>
<td>1.539</td>
<td></td>
</tr>
</tbody>
</table>

model behavior closer to the cash-in-advance limit \((\alpha \to \infty)\) and increase the stabilizing effects of introducing narrow banking.

C.3 Calibration of \( \rho \), \( \beta \), \( \bar{\kappa} \), \( \bar{v} \), and \( \psi \)

I choose the time preference rate \( \rho \), the bank diversification ability \( \beta \), the skin-in-the-game parameter \( \bar{\kappa} \) which governs the maximum fraction of capital that can be held by banks, the velocity parameter \( \bar{v} \) in the transaction cost function, and the efficiency \( \psi \) of deposits in transactions to match the following steady state targets under fractional reserve banking: bank leverage, the money multiplier, the share of government debt that is non-monetary, the nominal wealth share \( \vartheta \) measured as the share of government debt as a fraction of total wealth, and the money premium. Table 2 summarizes the targets and the calibrated parameter values.

I first explain how the targets moments stated in Table 2 are measured in the data. I then discuss the model counterparts of these targets and derive an explicit relationship between these targets and the five parameters.

For four of my five targets, all but the money premium, I use data from the Financial Accounts of the United States maintained by the federal reserve board\(^{94}\) and data on monetary aggregates provided by the Federal Reserve Bank of St. Louis.\(^{95}\) For all my measures, I use quarterly data from 1981Q1 to 2019Q4 and measure my target statistics by taking a time-series average over this sample period.

The data counterpart of my banks in the model are institutions classified as “private depository institutions” or “money market funds” in the financial accounts. To measure bank leverage, I compute the asset-to-equity ratio implies by the series “Private depository institutions; total liabilities” (data series code FL704190005.Q) and “Private depository institutions; total assets” (data series code FL704190006.Q).

---

\(^{94}\)https://www.federalreserve.gov/releases/z1/default.htm

\(^{95}\)Accessed through FRED, https://fred.stlouisfed.org/
The money multiplier is the ratio of total money to outside money. As a measure of total money, I use the measure “money of zero maturity” provided by the St. Louis Fed (FRED code MZM). As a measure of outside money, I use the sum of the monetary base (FRED code BOGMBASE) and treasury securities held by depository institutions and money market funds, represented by the series “Private depository institutions; Treasury securities; asset” (FL703061105.Q) and “Money market funds; Treasury securities; asset” (FL633061105.Q) in the financial accounts. I add treasuries held by bank-like institutions because these represent money-like assets whose value is ultimately derived from government liabilities, precisely like outside money.

In the model, \( \vartheta \) is the fraction of total wealth that takes the form of government-provided nominally safe assets. For this reason, I choose the steady-state target \( \vartheta \) to be approximately in line with the share of government assets in the portfolio of the empirical counterparts of my agents in the model. I take these empirical counterparts to be depository institutions and money market funds for “bankers” and the private domestic nonfinancial sectors for “households”. In the following, I will refer to the set of these sectors simply as “model sectors”. As a measure of (domestic) real wealth \( q^K \) in the model, I start from the series “All sectors; U.S. wealth” (FL892090005.Q) and remove all components that do not correspond to holdings by model sectors (these are proprietors’ equity held by security brokers, FL662090003.Q, nonfinancial assets held by government entities, FL312010095.Q and FL212010095.Q, and U.S. financial claims on the rest of the world, FL882090265.Q). I remove further all corporate equities that are not held directly or indirectly by model sectors. Not directly held equities are equities held by the rest of the world (FL263064105.Q), equities held by government entities (FL213064103.Q and FL313064105.Q), and equities held by the domestic financial sector (FL793064105.Q), except for those held by private depository institutions (FL703064105.Q), which I add back. This correction for not directly held equities subtracts too much from my wealth measure because equities held by the financial sector include those held by mutual funds which are partially indirectly held by the model sectors through mutual fund shares. For this reason, I add the value of indirectly held equities by model agents back to my measure of real wealth. I compute the value of indirectly held equities by the model sectors using the same methodology that

\[96\] I exclude money market funds from this calculation, because the financial accounts simply list all their assets as liabilities, but using this procedure would overstate total bank leverage because money market funds partially invest in debt issued by depository institutions. In any case, adding the money market fund measure does not substantially change the calibration target.

\[97\] Money market funds do not hold equity securities.
the financial accounts use for their series “corporate equities indirectly held by households” (FL653064155.Q), but include all model sectors. Specifically, the value of indirectly held equities by model sectors is

\[(\text{equities held by mutual funds}) \cdot \frac{\text{mutual fund shares held by model sectors}}{\text{all mutual fund shares}}.\]

In this formula, equities held by mutual funds (FL653064100.Q) and all mutual fund shares (FL653164205.Q) are directly retrieved from the financial accounts, whereas “mutual fund shares held by model sectors” is defined as the sum of mutual fund shares held by households and nonprofit organizations (FL153064205.Q), mutual fund shares held by nonfinancial corporate business (FL103064203) and mutual fund shares held by private depository institutions (FL703064205.Q).

I define total government-provided nominal wealth held by model sectors as the sum of my measure of outside money described above and the value of all treasuries held by private domestic nonfinancial sectors.\(^98\) The latter in turn is defined as the difference between treasuries held by domestic nonfinancial sectors (FL383061105.Q) and treasuries held by state and local governments (FL213061105.Q). My target for \(\theta\) is then defined as\(^99\)

\[\theta = \frac{\text{government-provided nominal wealth held by model sectors}}{\text{real wealth held by model sectors} + \text{government-provided nominal wealth held by model sectors}}.\]

I deal with the discrepancy between total government-provided nominal wealth and outside money by assuming that only the outside money fraction provides transaction services. The results in Section 7.1 then imply that the parameter \(\psi\) must be chosen as

\[\psi = \frac{\text{government-provided nominal wealth held by model sectors}}{\text{outside money}}.\]

For my calibration of the money premium, I follow Beneš and Kumhof (2013) who calibrate this premium to 1%.

I next describe the model counterparts of these targets and derive explicit representations of the calibrated parameter values. The parameter \(\psi\) has already been expressed in terms of

---

\(^{98}\) Treasuries held by the empirical counterpart of “banks” are already included in the measure of outside money.

\(^{99}\) By using all government debt in my measure, I make the implicit assumption that all government debt is net wealth (or backed by taxes that are fully reflected in the valuation of real wealth). This is broadly in line with the evidence on primary surpluses in the U.S. which have been fluctuating around zero over the post-war period.
data moments. For the remaining parameters, I start from the steady-state conditions

\[ \rho = \left( \frac{\bar{\kappa} \beta}{\eta} \right)^2 (1 - \vartheta)^2 \hat{\sigma}^2 + (1 - \psi) \lambda \nu \]  
\[ \rho = \left( \frac{1 - \tilde{\kappa}}{1 - \eta} \right)^2 (1 - \vartheta)^2 \hat{\sigma}^2 + \lambda \nu \]  

Equations (61) and (62) express the parameters \( \beta, \hat{\kappa}, \rho \) and \( \nu \) in terms of the targets \( \theta_{x,b}, m^M, \vartheta, \lambda \nu \) and the parameters \( \alpha, \phi, \hat{\sigma} = \sqrt{s} \) that are chosen in other parts of the model.

\[ \hat{\kappa} = \frac{m^M - 1}{1 - 1/(\theta_{x,b} \vartheta)} \]  

Next, the money multiplier is the ratio of total money to outside money, \( m^M := \frac{\theta + \kappa(1 - \vartheta) - \eta}{\theta_{x,b}} \). Using that in steady state \( \kappa = \bar{\kappa} \), substituting in \( \eta = \frac{\hat{\kappa}}{\theta_{x,b}} (1 - \vartheta) \) and solving for \( \bar{\kappa} \) implies

\[ \bar{\kappa} = \frac{\theta^M - 1}{1 - 1/(\theta_{x,b} \vartheta)} \]  

Next, plug \( \eta = \frac{\bar{\kappa}}{\theta_{x,b}} (1 - \vartheta) \) into (62), which is a version of the money valuation equation that determines the target \( \vartheta \),

\[ \rho = \left( \frac{1 - \vartheta}{1 - \bar{\kappa} (1 - \vartheta)} \right)^2 + \lambda \nu. \]  

This is the equation for the money premium in the original model with government debt, not the transformed model that is solved numerically. Compare footnote 62 for the difference.
C.4 Calibration of $a$ and $\phi$

The choice of the parameters $a$ and $\phi$ is ultimately irrelevant for most of my numerical results, because they only affect $q^K, q^M$ and $\iota$ as stated in Proposition 1, but not the equilibrium functions $\vartheta(\eta), \kappa(\eta), \lambda(\eta)\varphi(\eta)$ and thus they are irrelevant for equilibrium dynamics and endogenous risk amplification.\textsuperscript{101} I choose $\phi = 5$ following the calibration of Di Tella and Hall (2020) who use an identical adjustment cost function in a business cycle model driven by risk shocks similar to mine. Given this choice of $a$, the calibrated value for $\rho$ and the calibration target $\vartheta$, I set $a$ so that the steady-state investment-to-output ratio is 20%, which is approximately in line with the data. By Proposition 4, this requires a choice of

$$a = \frac{\rho}{0.8 (1 - \vartheta) - 0.2\phi \rho} = 0.0673.$$ 

D Proofs

D.1 Proofs of Results in Section 4

The proofs of all results in Section 4 are contained in the model solution procedure for the simplified model without banks outlined in Appendix B.1.1.

D.2 Proofs of Results in Section 5.1

The proofs of all results in Section 5.1 are contained in the model solution procedure for the full model with banks outlined in Appendix B.1.2.

D.3 Proofs of Results in Section 5.2

I prepare the proofs of the results in Section 5.2 with a number of technical lemmas.

\textbf{Lemma 5.} Along any (monetary) equilibrium path without aggregate risk ($\sigma_s(s_t) = \sigma^\eta(s_t, \eta_t) = 0$ along the equilibrium path), $\eta_t \geq \eta^I := \frac{\beta \kappa}{1 - \gamma - \beta \kappa}$ implies $\kappa_t = \bar{\kappa}$ and $\mu^I_t \leq 0$. In addition, if $\eta_t > \eta^I$, then $\mu^I_t < 0$.

\textsuperscript{101}This may not be obvious because both variables do enter the quantity equation (28) in Proposition 6. Note however, that they only enter by as a multiplicative constant in the definition of velocity, yet when $\vartheta$ is chosen to match a steady-state level of the money premium as explained above, then the impact of such multiplicative constants on equilibrium dynamics disappears (a different choice for $a$ and $\phi$ simply leads to a different choice of $\bar{\vartheta}$ such that the money premium is unaffected).
Proof. Show first that \( \kappa_t = \bar{\kappa} \). Otherwise, \( \kappa_t < \bar{\kappa} \) and by the capital allocation condition (Proposition 7), in the absence of aggregate risk
\[
\frac{\kappa_t}{\eta_t} \beta^2 = \frac{1 - \kappa_t + \psi \lambda_t \psi_t}{1 - \eta_t} \geq \frac{1 - \kappa_t}{1 - \eta_t} \Rightarrow \kappa_t \geq \frac{\eta_t}{\eta_t + (1 - \eta_t) \beta^2}.
\] (67)

The function \( \eta \mapsto \frac{\eta}{\eta + (1 - \eta) \beta^2} \) is strictly increasing, thus \( \eta_t \geq \eta^f \) implies
\[
\frac{\eta_t}{\eta_t + (1 - \eta_t) \beta^2} \geq \frac{\eta^f}{\eta^f + (1 - \eta^f) \beta^2} = \frac{\bar{\kappa} \beta_k}{1 - \kappa + \beta \bar{\kappa}} + \frac{1 - \bar{\kappa}}{1 - \kappa + \beta \bar{\kappa}} \beta^2 = \frac{\bar{\kappa}}{\kappa + \beta (1 - \bar{\kappa})} > \bar{\kappa},
\]
where the last inequality follows from \( \beta < 1 \). Combining this with (67) yields \( \kappa_t > \bar{\kappa} \), contradicting \( \kappa_t \leq \bar{\kappa} \). Consequently, the assumption \( \kappa_t < \bar{\kappa} \) was false and \( \kappa_t = \bar{\kappa} \) must hold.

Next, for \( \kappa_t = \bar{\kappa} \) and no aggregate risk, the drift of \( \eta_t \) is given by (compare Proposition 8)
\[
\mu_t^\eta = (1 - \eta_t) \left( \frac{\bar{\kappa} \beta}{\eta_t} - \left( \frac{1 - \bar{\kappa}}{1 - \eta_t} \right)^2 (1 - \theta_t) \sigma_t^2 - \psi (1 - \eta_t) \lambda_t \psi_t \right) \leq (1 - \eta_t) \left( \frac{\bar{\kappa} \beta}{\eta_t} - \left( \frac{1 - \bar{\kappa}}{1 - \eta_t} \right)^2 (1 - \theta_t) \sigma_t^2 \right).
\] (68)

In addition \( \eta_t \geq \eta^f \) implies
\[
\left( \frac{\bar{\kappa} \beta}{\eta_t} \right)^2 - \left( \frac{1 - \bar{\kappa}}{1 - \eta_t} \right)^2 \leq \left( \frac{\bar{\kappa} \beta}{\eta^f} \right)^2 - \left( \frac{1 - \bar{\kappa}}{1 - \eta^f} \right)^2 = (1 - \bar{\kappa} + \beta \bar{\kappa})^2 - (1 - \bar{\kappa} + \beta \bar{\kappa})^2 = 0.
\] (69)

Combining this with (68) and using \( (1 - \eta_t) (1 - \theta_t) \sigma_t^2 > 0 \) implies \( \mu_t^\eta \leq 0 \). In addition, if \( \eta_t > \eta^f \), then the inequality in (69) is strict and consequently \( \mu_t^\eta < 0 \). \( \square \)

Lemma 6. Along any equilibrium path without aggregate risk, if \( \eta_t > 0, \kappa_t < \bar{\kappa} \), and in addition either \( \psi = 0 \) or \( \theta_t^{x,b} \geq 1 \), then \( \mu_t^\eta > 0 \).

Proof. In the absence of aggregate risk and for \( \kappa_t < \bar{\kappa} \), then the capital allocation condition (Proposition 7) implies
\[
\frac{\kappa_t}{\eta_t} \beta^2 (1 - \theta_t) \sigma_t^2 = \frac{1 - \kappa_t}{1 - \eta_t} (1 - \theta_t) \sigma_t^2 + \psi \lambda_t \psi_t.
\] (70)
The drift of $\eta$ is therefore (compare Proposition 8)\footnote{The only reason the lemma requires $\eta_t > 0$ is because the following expression does not make sense for $\eta_t = 0.$}

$$\mu^\eta_t = (1 - \eta_t) \left( \left( \frac{\kappa_t \beta}{\eta_t} \right)^2 - \left( 1 - \kappa_t \right)^2 \right) (1 - \vartheta_t)^2 \tilde{\sigma}_t^2 - \psi (1 - \eta_t) \lambda_t v_t$$

$$= (1 - \eta_t) \left( \left( \frac{\kappa_t}{\eta_t} - \frac{1 - \kappa_t}{1 - \eta_t} \right) \frac{1 - \kappa_t}{1 - \eta_t} (1 - \vartheta_t)^2 \tilde{\sigma}_t^2 + \left( \frac{\kappa_t}{\eta_t} (1 - \vartheta_t) - 1 \right) \psi \lambda_t v_t \right).$$

Equation (70) immediately implies $\kappa_t > \eta_t$ (due to $\beta < 1$) and thus $\frac{\kappa_t}{\eta_t} > \frac{1 - \kappa_t}{1 - \eta_t}$. Hence, the first summand on the right is positive. In addition, by assumption either $\psi = 0$ and then the second summand vanishes, or $\vartheta_t^{x,b} \geq 1$, which by equation (18) means $\frac{\kappa_t}{\eta_t} (1 - \vartheta_t) \geq 1$, and then the second summand is nonnegative. In either case, one obtains $\mu^\eta_t > 0$.

\textbf{Proof of Proposition 9.} When $s_t = \bar{s}$ is constant, then in particular $\sigma_s(s_t) = 0$ and by Proposition 11 (proven below, but its proof does not rely on this proposition), there is no aggregate risk, $\sigma^\eta_t = \sigma^\vartheta_t = 0$. By Lemma 5 then $\mu^\eta_t < 0$ whenever $\eta > \eta^f = \frac{\beta \bar{\kappa}}{1 - \kappa + \beta \bar{\kappa}}$. In particular, no such $\eta$ can be a steady state. Next, Lemma 6 that at any inner steady state $\eta^{ss} \in (0, 1)$ it must be the case that $\kappa^{ss} = \bar{\kappa}$. Otherwise, only $\kappa < \bar{\kappa}$ is possible and then the lemma immediately implies $\mu^\eta_t > 0$, which is inconsistent with a steady state, if $\psi = 0$ (under narrow banking). For $\psi = 1$ (fractional reserve banking), the same conclusion is valid, provided one can show that in any steady state necessarily $\theta_t^{x,b} \geq 1$. This is indeed the case: the steady-state version of equation (53), $\rho - i^d = (\tilde{\sigma}^{n,b})^2 = (\theta_t^{x,b})^2 \beta^2 \tilde{\sigma}^2$, implies $(\theta_t^{x,b})^2 = \frac{\rho - i^d}{\beta^2 \tilde{\sigma}^2}$ and for $\psi = 1$, $i^d = 0$ (equation (51)), so $(\theta_t^{x,b})^2 = \frac{\rho}{\beta^2 \tilde{\sigma}^2}$ and by Assumption 2, this implies $\theta_t^{x,b} > 1$.

In total, the previous results imply any (inner) steady state equilibrium must satisfy $\eta^{ss} \leq \eta^f$ and $\kappa^{ss} = \bar{\kappa}$. I show next that precisely one such equilibrium can exist and treat the cases $\psi = 0$ and $\psi = 1$ separately:

1. In the case $\psi = 0$, the steady-state version of the drift of $\eta$ (Proposition 8) becomes

$$0 = \eta^{ss} (1 - \eta^{ss}) \left( \left( \frac{\kappa^{ss} \beta}{\eta^{ss}} \right)^2 - \left( 1 - \kappa^{ss} \right)^2 \right) (1 - \vartheta^{ss})^2 \tilde{\sigma}^2$$

and there is precisely one solution $\eta^{ss} \in (0, 1)$ to this equation, which satisfies

$$\frac{\kappa^{ss} \beta}{\eta^{ss}} = \frac{1 - \kappa^{ss}}{1 - \eta^{ss}} \iff \frac{\kappa \beta}{\eta^{ss}} = \frac{1 - \bar{\kappa}}{1 - \eta^{ss}} \iff \eta^{ss} = \frac{\beta \bar{\kappa}}{1 - \kappa + \beta \bar{\kappa}} = \eta^f.$$
For these value of $\eta^{ss}$ and $\kappa^{ss}$, the remaining steady state equations that determine $\vartheta^{ss}$ and $v^{ss}$ are (Propositions 27 and 6)

$$\rho = \left( \frac{(1 - \bar{\kappa})^2}{1 - \eta^{ss}} + \bar{\kappa}^2 \beta^2 \right) (1 - \vartheta^{ss})^2 \bar{\sigma}^2 + \lambda^{ss} v^{ss}$$

$$v^{ss} = \frac{a}{1 + \phi a} \left( \frac{1}{\vartheta^{ss}} + \phi \rho \right)$$  \hspace{1cm} (71)

For $\eta^{ss} = \eta^I = \frac{\beta \kappa}{1 - \kappa + \beta \kappa}$, first equation simplifies to

$$\rho = (1 - \bar{\kappa} + \bar{\kappa} \beta)^2 (1 - \vartheta^{ss})^2 \bar{\sigma}^2 + \lambda^{ss} v^{ss}.$$  \hspace{1cm} (72)

Consider now two cases separately:

(a) For $\alpha = \infty$, condition (29) means that either $\lambda^{ss} = 0$, then equation (72) implies $\vartheta^{ss} = \vartheta^I$, or $v^{ss} = \bar{v}$, then equation (71) implies $\vartheta^{ss} = \vartheta^M$. Furthermore, in the former case, it must be that $v^{ss}$ implied by equation (72) for $\vartheta^{ss} = \vartheta^I$ is weakly smaller than $\bar{v}$ and thus $\vartheta^I \geq \vartheta^M$. In the latter case, it must be that $\lambda^{ss} v^{ss}$ implied by equation (72) is positive and therefore $\vartheta^M > \vartheta^I$. In both cases, this implies $\vartheta^{ss} = \max\{\vartheta^I, \vartheta^M\}$. Because $\eta^I = \eta^M$ under narrow banking, it is also true that $\eta^{ss} = \min\{\eta^I, \eta^M\}$. Finally, $\vartheta^{ss} = \max\{\vartheta^I, \vartheta^M\}$ is not only a formal solution to the steady state equations, but also satisfies $\vartheta^{ss} \in (0, 1)$ and corresponds thus to a valid monetary equilibrium. This is a simple consequence of Assumption 1. In total, this proves both the existence of a unique steady state and the explicit formulas stated in the proposition for $\alpha = \infty$ under narrow banking.

(b) For $\alpha < \infty$, condition (29) implies $\lambda^{ss} v^{ss} = \frac{\alpha - 1}{\alpha} \left( \frac{v^{ss}}{\bar{v}} \right)^\alpha$. Combining the two equations (71) and (72) therefore results in the single equation

$$\rho = (1 - \bar{\kappa} + \bar{\kappa} \beta)^2 (1 - \vartheta^{ss})^2 \bar{\sigma}^2 + \frac{\alpha - 1}{\alpha} \left( \frac{1}{\bar{v}} \frac{a}{1 + \phi a} \frac{1}{\vartheta^{ss}} + \phi \rho \right)^\alpha$$

for $\vartheta^{ss}$. The right-hand side is strictly decreasing in $\vartheta^{ss} \in (0, 1)$ and therefore there is at most one solution to the equation. For $\vartheta^{ss} \to 0$, the right-hand side diverges to $\infty > \rho$, whereas for $\vartheta^{ss} \to 1$, it converges to $\frac{\alpha - 1}{\alpha} \left( \frac{\phi a}{\bar{v} (1 + \phi a)} \right)^\alpha < \rho$, where the inequality follows from Assumption 1. Consequently, there must be precisely one solution $\vartheta^{ss}$.

2. In the case $\psi = 1$, it is simpler to work with the steady-state versions of equations (50)
and (53) from Lemma 4 instead of the steady-state money valuation equation and \( \eta \) drift. It is immediate from the derivations in Section B.1.2 that the two sets of equations are equivalent. The steady-state versions of equations (50) and (53) are for \( \psi = 1 \)

\[
\rho = \left( \frac{\tilde{\kappa} \bar{\beta}}{\eta^{ss}} \right)^2 (1 - \bar{\theta}^{ss})^2 \bar{\sigma}^2,
\]

\[
\rho = \left( \frac{1 - \tilde{\kappa}}{1 - \eta^{ss}} \right)^2 (1 - \bar{\theta}^{ss})^2 \bar{\sigma}^2 + \lambda^{ss} v^{ss}.
\] (73)

The first equation immediately implies \( \eta^{ss} = \frac{\tilde{\kappa} \bar{\beta} \bar{\sigma}}{\sqrt{\rho}} (1 - \bar{\theta}^{ss}) =: \eta^{ss} (\bar{\theta}^{ss}) \), and in particular, \( (\eta^{ss})' (\bar{\theta}^{ss}) < 0 \). In addition, by Proposition 6

\[
v^{ss} = \frac{a}{1 + \phi a} \frac{1 - \bar{\theta}^{ss} + \phi \rho}{(1 - \bar{\kappa}) \bar{\theta}^{ss} + \bar{\kappa} - \eta^{ss} (\bar{\theta}^{ss})}.
\] (74)

Now consider again two cases:

(a) Suppose \( \alpha = \infty \). Equation (74) represents \( v^{ss} \) as a strictly decreasing function of \( \bar{\theta}^{ss} \). There is consequently precisely one value for \( \bar{\theta}^{ss} \), call it \( \bar{\theta}^{M} \), such that \( v^{ss} = \bar{v} \) and \( v^{ss} \leq \bar{v} \Leftrightarrow \bar{\theta}^{ss} \geq \bar{\theta}^{M} \). In addition, after substituting \( \eta^{ss} (\bar{\theta}^{ss}) \) into equation (73), the first term is strictly decreasing in \( \bar{\theta}^{ss} \), so there is precisely one value for \( \bar{\theta}^{ss} \), call it \( \bar{\theta}^{I} \), that is consistent with \( \lambda^{ss} = 0 \) and \( \lambda^{ss} \geq 0 \Leftrightarrow \bar{\theta}^{ss} \geq \bar{\theta}^{I} \). Those two arguments together with condition (29) imply that \( \bar{\theta}^{ss} = \max\{\bar{\theta}^{M}, \bar{\theta}^{I}\} \) is the unique steady state solution. With the corresponding \( \eta \) values, \( \eta^{M} := \eta^{ss} (\bar{\theta}^{M}) \), \( \eta^{I} := \eta^{ss} (\bar{\theta}^{I}) \), we have then obviously \( \eta^{ss} = \min\{\eta^{M}, \eta^{I}\} \). It is only left to show that \( \bar{\theta}^{M}, \bar{\theta}^{I}, \eta^{M}, \eta^{I} \) take the asserted form. These are simple calculations. \( \bar{\theta}^{I} \) is defined by

\[
\rho = \left( \frac{1 - \tilde{\kappa}}{1 - \eta^{ss} (\bar{\theta}^{I})} \right)^2 (1 - \bar{\theta}^{I})^2 \bar{\sigma}^2 \Leftrightarrow \sqrt{\rho} - \tilde{\kappa} \beta \bar{\sigma} (1 - \bar{\theta}^{I}) = (1 - \tilde{\kappa}) (1 - \bar{\theta}^{I}) \bar{\sigma}
\]

\[
\Leftrightarrow (1 - \bar{\theta}^{I}) = \frac{\sqrt{\rho}}{(1 - \tilde{\kappa} + \tilde{\kappa} \beta) \bar{\sigma}}
\]

and this is clearly in line with the asserted form for \( \bar{\theta}^{I} \). \( \eta^{I} \) is then given by

\[
\eta^{I} = \eta^{ss} (\bar{\theta}^{I}) = \frac{\tilde{\kappa} \beta \bar{\sigma}}{\sqrt{\rho} (1 - \tilde{\kappa} + \tilde{\kappa} \beta) \bar{\sigma}} = \frac{\tilde{\kappa} \beta}{1 - \tilde{\kappa} + \tilde{\kappa} \beta}.
\] 104
Similarly, $\vartheta^M$ is defined by

$$\tilde{v} = \frac{a}{1 + \phi a} \frac{1 - \vartheta^M + \phi \rho}{\vartheta^M + \bar{\kappa} - \eta^s (\vartheta^M)}$$

$$\Leftrightarrow \tilde{v} \left( \frac{(1 - \bar{\kappa}) \vartheta^M + \bar{\kappa} - \frac{\bar{\kappa} \beta \tilde{\sigma}}{\sqrt{\rho}} (1 - \vartheta^M)}{1 + \phi a} (1 - \vartheta^M + \phi \rho) \right) = \frac{a}{1 + \phi a} \left( \frac{(1 - \bar{\kappa}) + \frac{\bar{\kappa} \beta \tilde{\sigma}}{\sqrt{\rho}}}{1 + \phi a} \vartheta^M = \frac{a}{1 + \phi a} (1 + \phi \rho) - \bar{\kappa} \tilde{v} \left( 1 - \frac{\beta \tilde{\sigma}}{\sqrt{\rho}} \right) \right)$$

$$\Leftrightarrow \vartheta^M = \frac{a (1 + \phi \rho) - \bar{\kappa} (1 + \phi a) \left( 1 - \frac{\beta \tilde{\sigma}}{\sqrt{\rho}} \right) \tilde{v}}{a + (1 + \phi a) \left( 1 - \bar{\kappa} + \frac{\bar{\kappa} \beta \tilde{\sigma}}{\sqrt{\rho}} \right) \tilde{v}}$$

which is up to minor rearrangements in the denominator as stated in the proposition.

Finally, $\vartheta^M$ is then

$$\vartheta^M = \eta^s \left( \vartheta^M \right) = \eta^s \left( \vartheta^M \right) = \frac{\bar{\kappa} \beta \tilde{\sigma}}{\sqrt{\rho}} \frac{1 + \phi a (\tilde{v} - \rho)}{a + (1 + \phi a) \left( 1 - \bar{\kappa} + \frac{\bar{\kappa} \beta \tilde{\sigma}}{\sqrt{\rho}} \right) \tilde{v}}.$$

Finally, that $(\eta^s, \vartheta^{ss})$ always represents not just a mathematical solution to the equations, but a valid inner monetary steady state equilibrium, if $\eta^s, \vartheta^{ss} \in (0, 1)$. For $\eta^s = \min \{ \eta^M, \eta^I \}$, this is always satisfied. In the case $\eta^s = \eta^M$, positivity is a direct consequence of Assumption 1. It is also clearly the case that $\vartheta^{ss} = \max \{ \vartheta^M, \vartheta^I \} < 1$. A steady state thus exists if

$$\vartheta^I > 0 \Leftrightarrow \tilde{\sigma} > \frac{\sqrt{\rho}}{1 - \bar{\kappa} + \bar{\kappa} \beta}$$

or if

$$\vartheta^M > 0 \Leftrightarrow \tilde{v} < \frac{a}{\bar{\kappa} \left( 1 - \frac{\beta \tilde{\sigma}}{\sqrt{\rho}} \right) \tilde{v}}.$$

In particular, one of these conditions is always satisfied for sufficiently large $\tilde{\sigma}$ or for sufficiently small $\tilde{v}$.

(b) For $\alpha < \infty$, because $\frac{1 - \vartheta^{ss}}{1 - \eta^s} = \frac{1 - \vartheta^{ss}}{1 - \eta^s (\vartheta^{ss})}$ is strictly decreasing in $\vartheta^{ss}$, so is the first term in equation (73). Furthermore, equation (74) represents $v^{ss}$ as a strictly decreasing function of $\vartheta^{ss}$ and by condition (29), $\lambda^s v^{ss}$ is a strictly increasing function of $v^{ss}$, so that also the second term in equation (73) must also be strictly decreasing in $\vartheta^{ss}$.
This implies, that there is at most one solution \( \vartheta^{ss} \) to the steady state equations (and if there is one, then the remaining equations determine \( \eta^{ss} \) and \( v^{ss} \) uniquely).

For \( \vartheta^{ss} \rightarrow 1 \), the right-hand side of equation (73) becomes \( \lambda^{ss} v^{ss} \) and \( v^{ss} \rightarrow \frac{a}{1+\phi a} \phi \rho \).

Thus, by the same arguments as for narrow banking, Assumption 1 implies that the right-hand side is smaller than \( \rho \). Consequently, any solution must satisfy \( \vartheta^{ss} < 1 \).

By continuity, there is precisely one solution, if in addition the right-hand side of equation (73) becomes larger than \( \rho \) for \( \vartheta^{ss} \rightarrow 0 \). This may not always be the case.

The limit for \( \vartheta^{ss} \rightarrow 0 \) is

\[
\left( \frac{1 - \bar{\kappa}}{1 - \bar{\kappa} \cdot \beta \bar{\sigma} / \sqrt{\rho}} \right)^2 \sigma^2 + \frac{\alpha - 1}{\alpha} \left( \frac{1}{\bar{\nu} \cdot 1 + \phi a \bar{\kappa} \left( 1 - \beta \bar{\sigma} / \sqrt{\rho} \right)} \right)^\gamma
\]

and this certainly larger than \( \rho \) if \( \bar{\sigma} \) is large or \( \bar{\nu} \) is small.

\( \square \)

**Proof of Proposition 10.** That \( \kappa \) is identical under both banking regimes follows directly from Proposition 9. In the case \( \alpha = \infty \), the assertion follows directly from the closed-form expressions stated in Proposition 9. It is therefore only left to consider the case \( \alpha < \infty \).

I start with the claim for \( \eta^{ss} \). Because \( \eta^{ss} = \frac{\bar{\kappa} \beta}{1 - \bar{\kappa} + \bar{\kappa} \beta} \) under narrow banking and \( \kappa_t = \bar{\kappa} \) for \( \eta_t \geq \frac{\bar{\kappa} \beta}{1 - \bar{\kappa} + \bar{\kappa} \beta} \) by Lemma 5, the drift of \( \eta \) is for any such \( \eta_t \) given by

\[
\mu_t^\eta = (1 - \eta_t) \left( \left( \frac{\bar{\kappa} \beta}{\eta_t} \right)^2 - \left( \frac{1 - \bar{\kappa}}{1 - \eta_t} \right)^2 \right) (1 - \vartheta_t)^2 \bar{\sigma}^2 - (1 - \eta_t) \psi \lambda_t v_t.
\]

The expression on the right is always negative, unless both \( \eta_t = \frac{\bar{\kappa} \beta}{1 - \bar{\kappa} + \bar{\kappa} \beta} \) and \( \psi \lambda_t v_t = 0 \). The latter is never the case for \( \psi > 0 \) and \( \alpha < \infty \), consequently \( \eta^{ss} \) must be smaller than \( \frac{\bar{\kappa} \beta}{1 - \bar{\kappa} + \bar{\kappa} \beta} \) under fractional reserve banking.

For the proof that \( \vartheta^{ss} \) is smaller under fractional reserve banking, consider the following
three equations which determine the steady-state values for $\vartheta$, $\eta$ and $v$.

\begin{align*}
v &= \frac{a}{1 + \phi \alpha \vartheta + \psi (\bar{\kappa} (1 - \vartheta) - \eta)} \\
rho &= \left( \frac{(1 - \bar{\kappa})^2}{1 - \eta} + \frac{\bar{\kappa}^2 \beta^2}{\eta} \right) \left( 1 - \vartheta \right)^2 \bar{\sigma}^2 + (1 - \psi \eta) i(v) \\
\left( \frac{\bar{\kappa} \beta}{\eta} \right)^2 \left( 1 - \vartheta \right)^2 \bar{\sigma}^2 &= \left( \frac{1 - \bar{\kappa}}{1 - \eta} \right)^2 \left( 1 - \vartheta \right)^2 \bar{\sigma}^2 + \psi i(v)
\end{align*}

where

\[ i(v) := \lambda v = \frac{\alpha - 1}{\alpha} \left( \frac{v}{\bar{v}} \right)^{\alpha}. \]

For all $\psi \leq 1/\bar{\kappa}$ (in particular for $\psi \leq 1$), equation (76) expresses $v$ as a strictly decreasing function of $\vartheta$ (for given $\eta$). Because the right-hand side of equation (76) is strictly decreasing in $\vartheta$ and strictly increasing in $\eta$, equations (76), (76) must then have for any given $(\eta, \psi)$ at most one solution $\vartheta = \vartheta_1(\eta; \psi)$ (at least locally around the existing steady state there is precisely one solution). The function $\vartheta_1$ is clearly decreasing in $\psi$ and it will be shown below that it is increasing in $\eta$.

In addition, combining equations (76) and (77) by eliminating the money premium $i(v)$ yields the single equation

\[ \rho = \left( \frac{(1 - \bar{\kappa})^2}{1 - \eta} + \frac{\bar{\kappa}^2 \beta^2}{\eta} \right) \left( 1 - \vartheta \right)^2 \bar{\sigma}^2 + \frac{1 - \psi \eta}{\psi} \left( \frac{\bar{\kappa} \beta}{\eta} \right)^2 - \left( \frac{1 - \bar{\kappa}}{1 - \eta} \right)^2 \left( 1 - \vartheta \right)^2 \bar{\sigma}^2. \]

The first term on the right is strictly decreasing in $\vartheta$ and for $\eta \leq \frac{\bar{\kappa} \beta}{1 - \bar{\kappa} + \bar{\kappa} \beta} =: \eta^I$ (only these $\eta$ are relevant in any steady state due to Lemma 5), also the second term is decreasing in $\vartheta$. Consequently, this equation has for any given $(\eta, \psi)$ at most one solution $\vartheta = \vartheta_2(\eta; \psi)$ (and again locally around the existing steady state there must be precisely one). Because the right-hand side is also strictly decreasing in both $\eta$ and $\psi$, the function $\vartheta_2$ must be (strictly) decreasing in both $\eta$ and $\psi$.

Because both functions $\vartheta_1(\cdot; \psi)$ and $\vartheta_2(\cdot, \psi)$ trace out the solution sets of $(\eta, \vartheta)$-pairs for two of the three steady-state equations, the steady-state level of $\eta$ is precisely the unique intersection point that satisfies $\vartheta_1(\eta^{ss}; \psi) = \vartheta_2(\eta^{ss}; \psi)$ and then $\vartheta^{ss} = \vartheta_1(\eta^{ss}; \psi) = \vartheta_2(\eta^{ss}; \psi)$. Because both $\vartheta_1$ and $\vartheta_2$ are strictly decreasing in $\psi$, then also the steady-state solution $\vartheta^{ss}$ must be strictly decreasing in $\psi$.\footnote{These are the quantity equation (28), the money valuation equation (27) and the drift of $\eta$ (Proposition 8).}
For $\psi$, and thus $\vartheta^{ss}$ must be lower under fractional reserve banking ($\psi = 1$) than under narrow banking ($\psi = 0$).

It remains to be shown that $\vartheta_1$ is indeed increasing in $\eta$ as claimed above. This follows from an application of the implicit function theorem to the combination of equations (67) and (68),

$$\frac{\partial \vartheta_1}{\partial \eta} = \frac{d}{d\eta} \left[ \frac{(1-\kappa)^2 + \kappa^2 \beta^2}{(1-\kappa)^2 + \kappa^2 \beta^2} \right] (1-\vartheta)^2 \sigma^2 + (1 - \psi \eta) i(v(\eta, \vartheta))$$

$$= \frac{\left(1 - \kappa \right)^2 - \left( \kappa \beta \eta \right)^2}{(1-\kappa)^2 + \kappa^2 \beta^2} \right] (1-\vartheta)^2 \sigma^2 + (1 - \psi \eta) i(v(\eta, \vartheta)) (\frac{\partial v}{\partial \eta})$$

and it is left to show that this is always positive. By equation (67), $\frac{\partial \log v}{\partial \eta} = \frac{\psi}{v + \psi (1-\vartheta - \eta)}$ and thus

$$(1 - \psi \eta) \frac{\partial \log v}{\partial \eta} = \frac{1 - \psi \eta}{v + \psi (1-\vartheta - \eta)}$$

For $\psi \leq 1/\kappa$, $(1 - \psi \kappa) \vartheta + \psi \kappa \in [\vartheta, 1]$ and thus $1 - \psi \eta \frac{1}{1 - \psi \eta} \leq 1 - \psi \eta \frac{1}{1 - \psi \eta} = 1$, which implies $(1 - \psi \eta) \frac{\partial \log v}{\partial \eta} \geq \psi$. For $\alpha > 2$, (78) is then indeed positive.

104While this should be obvious to any economist due to the clear analogy to supply-demand diagrams, here is a formal proof of the assumption of differentiability: if $(\eta^{ss})' (\psi) \geq 0$, then due to $\vartheta^{ss}(\psi) = \vartheta_2(\eta^{ss}(\psi); \psi)$, $\partial_{\eta} \vartheta_2, \partial_{\psi} \vartheta_2 < 0$, $(\vartheta^{ss})' (\psi) = \partial_{\eta} \vartheta_2(\eta^{ss}(\psi); \psi)$ $(\eta^{ss})' (\psi) + \partial_{\psi} \vartheta_2(\eta^{ss}(\psi); \psi) < 0$. If instead $(\eta^{ss})' (\psi) < 0$, then due to $\vartheta^{ss}(\psi) = \vartheta_1(\eta^{ss}(\psi); \psi)$, $\partial_{\eta} \vartheta_1 > 0, \partial_{\psi} \vartheta_1 < 0, (\vartheta^{ss})' (\psi) = \partial_{\eta} \vartheta_1(\eta^{ss}(\psi); \psi) (\eta^{ss})' (\psi) + \partial_{\psi} \vartheta_1(\eta^{ss}(\psi); \psi) < 0$. So in any case, $(\vartheta^{ss})' (\psi) < 0$.

105Applying that equation here means that the result $\partial_{\eta} \vartheta_1 > 0$ to be shown may only hold locally around the steady state, which is sufficient for the above argument for the monotonic relationship between $\psi$ and $\vartheta^{ss}$ to work out.
D.4 Proofs of Results in Section 5.3

Proof of Proposition 11. In any Markov equilibrium in the sense of Definition 4, \( \vartheta_t \) is a function of \( s_t \) and \( \eta_t \), \( \vartheta_t = \vartheta(s_t, \eta_t) \). Applying Ito’s formula to this equation yields for the volatility terms (dZ-loadings)

\[
\sigma_t^\vartheta \vartheta_t = -\partial_s \vartheta(s_t, \eta_t) \sigma_s(s_t) + \partial_\eta \vartheta(s_t, \eta_t) \sigma_\eta \eta_t.
\]

Furthermore, by Proposition 8,

\[
\sigma_\eta \eta_t = -(\kappa_t - \eta_t) \sigma_t^\vartheta.
\]

Combining the two equations yields

\[
\sigma_t^\eta \eta_t = (\kappa_t - \eta_t) \frac{\partial_s \vartheta(s_t, \eta_t)}{\vartheta(\eta_t, s_t)} \sigma_s(s_t) - (\kappa_t - \eta_t) \frac{\partial_\eta \vartheta(s_t, \eta_t)}{\vartheta(\eta_t, s_t)} \sigma_\eta \eta_t
\]

and solving for \( \sigma_\eta \eta_t \) results in the asserted equation.

Proof of Proposition 12. For the proof of the local version of the result note that whenever \( \lambda(s, \eta) > 0 \), then \( v(s, \eta) = \bar{v} \) by equation (29). Using the representation of velocity from Proposition 6 under narrow banking, this is equivalent to

\[
\frac{a}{1 + \phi a} \frac{1 - \vartheta(s, \eta) + \phi \rho}{\vartheta(s, \eta)} = \bar{v},
\]

or, after solving for \( \vartheta(s, \eta) \), \( \vartheta(s, \eta) = \bar{v}^M \). By assumption, \( \lambda(s, \eta) > 0 \) for all \((s, \eta) \) in a \(-\lambda\)-open neighborhood \( U \) of \((s_0, \eta_0) \), so by the previous argument \( \vartheta(s, \eta) = \bar{v}^M \) for all \((s, \eta) \) \( \in U \). Because \( U \) is a open set, this implies

\[
\partial_s \vartheta(s, \eta) = \partial_\eta \vartheta(s, \eta) = \partial_{ss} \vartheta(s, \eta) = \partial_{s\eta} \vartheta(s, \eta) = \partial_{\eta\eta} \vartheta(s, \eta) = 0
\]

for all \((s, \eta) \) \( \in U \). In particular, an application of Ito’s formula then implies \( \mu^\vartheta(s, \eta) = \sigma^\vartheta(s, \eta) = 0 \) for all such \((s, \eta) \), including \((s, \eta) = (s_0, \eta_0) \). In addition, the vanishing derivatives in combination with Proposition 11 imply \( \sigma^\eta(s_0, \eta_0) = 0 \). Finally, with \( \sigma^\eta(s_0, \eta_0) = 0 \), Lemma 5 and 6 are applicable and imply that the drift of \( \eta \) is negative whenever \( \eta_0 > \frac{\kappa \beta}{1 - \kappa + \kappa \beta} = \eta^M \) and positive, whenever \( \eta_0 < \eta^M \) and \( \kappa(s_0, \eta_0) < \bar{\kappa} \). For this reason, only the case \( \kappa(s_0, \eta_0) = \bar{\kappa} \) and \( \eta_0 \leq \eta^M \) has to be investigated. In this case, the drift of \( \eta \) reduces to (compare Proposition 8)

\[
\mu^\eta(s_0, \eta_0) \eta_0 = \eta_0 (1 - \eta_0) \left( \left( \frac{\bar{\kappa} \beta}{\eta_0} \right)^2 - \left( \frac{1 - \bar{\kappa}}{1 - \eta_0} \right)^2 \right) (1 - \vartheta(s_0, \eta_0))^2 \bar{\sigma}(s_0)^2
\]

109
and this is nonnegative for all $\eta_0 \leq \eta^M$ and strictly positive for $\eta_0 < \eta^M$. Hence, the system drifts locally towards $\eta^M$. This concludes the proof of the local stability result.

For the proof of the global stability result, we simply show that $\vartheta(s, \eta^M) = \vartheta^M$ and $\kappa(s, \eta^M) = \bar{\kappa}$ are valid equilibrium functions. This implies that when the system has reached $\eta = \eta^M$, then it remains there forever. Note first that for any fixed $\hat{\sigma} \leq \hat{\sigma}^I$, there is a multiplier $\lambda(s, \eta^M)$, so that with $v(s, \eta^M)$ defined in line with Proposition 6, these constants solve the steady-state versions of the money valuation equation (27), the capital allocation equation (30) and the drift of $\eta$ (Proposition 8), where “steady-state version” simply means that $\mu^\vartheta = \mu^n = \sigma^\vartheta = \sigma^n = 0$. This is a direct consequence of the steady state result stated in Proposition 9 and the fact that $\hat{\sigma} \leq \hat{\sigma}^I \iff \vartheta^M \geq \vartheta^I$ (compare the proof of Proposition 10). This argument therefore implies that if $\mu^\vartheta = \mu^n = \sigma^\vartheta = \sigma^n = 0$, then $\vartheta(s, \eta^M) = \vartheta^M$ and $\kappa(s, \eta^M) = \bar{\kappa}$ are valid equilibrium solutions. Furthermore, the proof of Proposition 9 immediately reveals that for the “monetarist” steady-state solution, the associated multiplier $\lambda(s, \eta^M)$ must always be positive. But by the local version of this proposition, a positive $\lambda(s, \eta^M)$ implies that the proposed equilibrium solution is indeed consistent with $\mu^\vartheta = \sigma^\vartheta = \sigma^n = 0$ and it is also consistent with $\mu^n = 0$, provided $\eta = \eta^M$, which is what has been assumed here. Thus, $\vartheta(s, \eta^M) = \vartheta^M$ and $\kappa(s, \eta^M) = \bar{\kappa}$ represent a valid equilibrium.

\[\square\]

D.5 Proofs of Results in Section 6

I prepare the proof of Proposition 13 with a simple Lemma about the expectation of logarithms of Ito processes.

Lemma 7. Let $X_t$ be a positive Ito process with evolution

$$dX_t/X_t = \mu^X_t \, dt + \sigma^X_t \, dZ_t + \tilde{\sigma}^X_t \, d\tilde{Z}_t,$$

where $Z$ and $\tilde{Z}$ are independent Brownian motions. Then,

$$\mathbb{E}_0 \left[ \int_0^\infty e^{-\rho t} \log X_t \, dt \right] = \frac{\log X_0}{\rho} + \mathbb{E}_0 \left[ \int_0^\infty e^{-\rho t} \frac{\mu^X_t}{\rho} - \frac{1}{2} \left( \frac{\sigma^X_t}{\rho} \right)^2 \, dt \right].$$
Proof. Applying Ito’s formula to $\log X$ implies

$$
d \log X = \frac{dX_t}{X_t} - \frac{1}{2} \frac{d\langle X \rangle_t}{X_t^2}
= \left( \mu_t^X - \frac{1}{2} \left( (\sigma_t^X)^2 + (\hat{\sigma}_t^X)^2 \right) \right) dt + \sigma_t^X dZ_t + \hat{\sigma}_t^X d\tilde{Z}_t.
$$

Writing this in integral form over the interval $[0,t]$ and taking expectations yields

$$
\mathbb{E}_0 [\log X_t] = \log X_0 + \mathbb{E}_0 \left[ \int_0^t \left( \mu_s^X - \frac{1}{2} \left( (\sigma_s^X)^2 + (\hat{\sigma}_s^X)^2 \right) \right) ds \right] + \mathbb{E}_0 \left[ \int_0^t \sigma_s^X dZ_s + \int_0^t \hat{\sigma}_s^X d\tilde{Z}_s \right] = 0 \quad \text{(79)}
$$

and the last term vanishes because it is the expectation of a martingale with initial value 0. Using this result, one obtains by simple calculations

$$
\mathbb{E}_0 \left[ \int_0^\infty e^{-\rho t} \log X_t dt \right] = \int_0^\infty e^{-\rho t} \mathbb{E}_0 [\log X_t] dt = \int_0^\infty e^{-\rho t} \log X_0 dt
$$

(Fubini’s Theorem)

$$
= \frac{\log X_0}{\rho} + \mathbb{E}_0 \left[ \int_0^\infty \int_s^\infty e^{-\rho s} \left( \mu_s^X - \frac{1}{2} \left( (\sigma_s^X)^2 + (\hat{\sigma}_s^X)^2 \right) \right) ds dt \right] = \frac{\log X_0}{\rho} + \mathbb{E}_0 \left[ \int_0^\infty e^{-\rho s} \frac{\mu_s^X}{\rho} \left( (\sigma_s^X)^2 + (\hat{\sigma}_s^X)^2 \right) ds \right].
$$

(Fubini’s Theorem)

Proof of Proposition 13. All agents in the model choose an identical consumption-wealth ratio of $\rho$, hence $c_t(i) = n_{i0}(i)$ for all $i \in \mathbb{I}$. Then, by definition of $\tilde{\eta}_t(i)$,

$$
c_t(i) = \tilde{\eta}_t(i) \rho \left( N_t^h + N_t^b \right) = \tilde{\eta}_t(i) \left( C_t^h + C_t^b \right) = \tilde{\eta}_t(i) (a - \iota_t) K_t,
$$

where the second equality follows from the fact that also the aggregate consumption-wealth ratio must be $\rho$ and the last equality is the aggregate resource constraint (6). Taking the logarithm

111
and integrating both over time and over states yields

\[ V_0(\tilde{i}) = E_0 \left[ \int_0^\infty e^{-\rho t} \log (a - \iota_t) \right] + E_0 \left[ \int_0^\infty e^{-\rho t} \log \tilde{\eta}_t(\tilde{i}) \right] + E_0 \left[ \int_0^\infty e^{-\rho t} \log K_t \right]. \tag{80} \]

Next, apply Lemma 7 to the processes \( \tilde{\eta}_t(\tilde{i}) \) and \( K_t \). These are Ito processes satisfying

\[
\frac{d\tilde{\eta}_t(\tilde{i})}{\tilde{\eta}_t(\tilde{i})} = \frac{d \left( n_t(\tilde{i}) / N_t^{t(\tilde{i})} \cdot \eta_t^{t(\tilde{i})} \right)}{n_t(\tilde{i}) / N_t^{t(\tilde{i})} \cdot \eta_t^{t(\tilde{i})}} \\
= \mu^{\eta, t(\tilde{i})} dt + \sigma^{\eta, t(\tilde{i})} dZ_t + \tilde{\sigma}_t^{n, t(\tilde{i})} d\tilde{Z}_t
\]

and

\[
\frac{dK_t}{K_t} = (\Phi (\iota_t) - \delta) dt,
\]

so by Lemma 7,

\[
E_0 \left[ \int_0^\infty e^{-\rho t} \log \tilde{\eta}_t(\tilde{i}) \right] = \frac{\log \tilde{\eta}_0(\tilde{i})}{\rho} + E_0 \left[ \int_0^\infty e^{-\rho t} \frac{\mu^{\eta, t(\tilde{i})} - \frac{1}{2} \left( \left( \sigma^{\eta, t(\tilde{i})} \right)^2 + \left( \tilde{\sigma}_t^{n, t(\tilde{i})} \right)^2 \right)}{\rho} dt \right],
\]

\[
E_0 \left[ \int_0^\infty e^{-\rho t} \log K_t \right] = \frac{\log \tilde{\eta}_0(\tilde{i})}{\rho} + E_0 \left[ \int_0^\infty e^{-\rho t} \frac{\Phi (\iota_t) - \delta}{\rho} dt \right].
\]

Substituting these expressions into equation (80) yields the decomposition of \( V_0(\tilde{i}) \) in the assertion. The explicit expressions for \( \tilde{\sigma}_t^{n, t(\tilde{i})} \) has already been derived in Lemma 4.

Proof of Proposition 14. I show that condition (33) restricts the planner’s choice in a way that \( \eta_0 \) is effectively the only relevant free control of the planner and it is chosen to maximize the objective stated in the assertion. To clarify the mathematically imprecise statement in the main text, the planner chooses Borel-measurable functions \( t: \mathbb{I} \to \{b, h\} \) and \( \tilde{\eta}_0 : \mathbb{I} \to [0, \infty) \),\(^{106}\) such that for the implied wealth share of bankers \( \eta_0 := \int_0^1 b(t) \tilde{\eta}_0(\tilde{i}) d\tilde{i} \), the equilibrium allocation at the aggregate state \((s_0, \eta_0)\) and wealth distribution \( (\tilde{\eta}_0(\tilde{i}))_{\tilde{i} \in I} \) leads to expected utilities \( (\tilde{V}(\tilde{i}))_{\tilde{i} \in I} \) that satisfy condition (33). Let \( \nu \) be the (probability) measure associated with the density \( u \), that is \( \nu(A) = \int_A u(\tilde{i}) d\tilde{i} \) for all \( A \in B(\mathbb{I}) \) (Borel subsets of \( \mathbb{I} \)) and for any given choice
t, define the implies sets of bankers and households,

\[ B := t^{-1}(\{b\}), \quad H := t^{-1}(\{h\}) = I \setminus B. \]

Note that \( B \) and \( H \) depend on \( t \), yet I suppress this dependence in the notation.

Let \( t \) and \( \tilde{\eta}_0 \) be any arbitrary admissible choice of the planner. By Proposition 13, utility of agent \( \tilde{i} \in I \) in permanent consumption units can be written as

\[ \exp(\rho V_0(\tilde{i})) = \tilde{\eta}_0(\tilde{i}) v^b(s_0, \eta_0) K_0 \]

and using condition (33), one can rewrite this equation as

\[ \int_I \exp(\rho V_0(\tilde{j})) \, d\tilde{j} \cdot u(\tilde{i}) = \tilde{\eta}_0(\tilde{i}) v^h(s_0, \eta_0) K_0. \]

Integrating equation (82) over \( B \) yields

\[ \nu(B) \int_I \exp(\rho V_0(\tilde{i})) \, d\tilde{i} = \int_B \tilde{\eta}_0(\tilde{i}) \, d\tilde{i} \cdot v^b(s_0, \eta_0) K_0 = \eta_0 v^b(s_0, \eta_0) K_0. \]

Similarly, integrating equation (82) over \( H \) yields

\[ \nu(H) \int_I \exp(\rho V_0(\tilde{i})) \, d\tilde{i} = \int_H \tilde{\eta}_0(\tilde{i}) \, d\tilde{i} \cdot v^h(s_0, \eta_0) K_0 = (1 - \eta_0) v^h(s_0, \eta_0) K_0. \]

Equations (83) and (84) have three implications. First, combining the two equations and solving for \( \omega := \nu(B) = 1 - \nu(H) \) (the share of total utility in permanent consumption units distributed to bankers) yields

\[ \omega = \frac{\eta_0 v^b(s_0, \eta_0)}{\eta_0 v^b(s_0, \eta_0) + (1 - \eta_0) v^h(s_0, \eta_0)}. \]

Second, combining each of the equations (83) and (84) with condition (33) yields an explicit representation of utility,

\[ \exp(\rho V_0(\tilde{i})) = \begin{cases} \eta_0 u(\tilde{i}) v^b(s_0, \eta_0) K_0, & \tilde{i} \in B \\ \frac{1-\eta_0}{1-\omega} u(\tilde{i}) v^h(s_0, \eta_0) K_0, & \tilde{i} \in H \end{cases} \]
and substituting $\omega$ from (85) into this equation implies

$$
\exp \left( \rho V_0 (\tilde{i}) \right) = u(\tilde{i}) \left( \eta_0 v^b (s_0, \eta_0) + (1 - \eta_0) v^h (s_0, \eta_0) \right) K_0.
$$

(86)

Third, comparing the utility representations (81) and (86) implies for the initial wealth distribution

$$
\tilde{\eta}_0 (\tilde{i}) = \begin{cases} 
\frac{\eta_0}{\omega} u(\tilde{i}), & \tilde{i} \in B \\
\frac{1 - \eta_0}{1 - \omega} u(\tilde{i}), & \tilde{i} \in H.
\end{cases}
$$

(87)

Equation (87) says that the choice of the wealth distribution $\tilde{\eta}_0$ is fully determined by the implications of the planner’s choices for $\eta_0$ and $\omega$ and the sets $B$ and $H$. Equation (85) furthermore implies that for all admissible choices, $\omega$ is already a function of $\eta_0$, so that by the previous statement, $\tilde{\eta}_0$ is already fully determined by the choice of $\eta_0$ and the sets $B$ and $H$. It is also not hard to see that conversely for any $\eta_0 \in [0, 1]$ and any Borel set $B$ such that $\nu(B) = \omega(\eta_0)$ with $\omega(\eta_0)$ as in equation (85), the wealth distribution $\tilde{\eta}_0$ defined by equation (87) together with $t(\tilde{i}) = b1_B(\tilde{i}) + h1_{\setminus B}(\tilde{i})$ are admissible planner choices that imply $\int_{t^{-1}(b)} \tilde{\eta}_0 (\tilde{i}) d\tilde{i} = \eta_0$. The set of admissible planner choices has therefore been reduced to a pair $(\eta_0, B) \in [0, 1] \times B(I)$ that satisfies the additional constraint $\nu(B) = \omega(\eta_0)$ (with $\omega(\eta_0)$ as in equation (85)). Clearly, for any $\eta_0 \in [0, 1]$ there is a set $B \in B(I)$ that satisfies this constraint, e.g. $B = [0, \tilde{i}_\omega]$ where $\tilde{i}_\omega$ is a $\omega$-quantile of $\nu$, $\nu([0, \tilde{i}_\omega]) = \omega$.\textsuperscript{107}

Finally, because agents’ utilities in the representation (86) do not depend on $B$ or the type assignment function $t$, but only on $\eta_0$ and the given quantities $u(\tilde{i})$ and $K_0$, the planner is indifferent between all choices of $B$ that lead to the same banker wealth share $\eta_0$. For this reason, the planner’s problem has been reduced to the choice of $\eta_0 \in [0, 1]$ to maximize some increasing function of the right-hand side of (86) for different $\tilde{i} \in I$. For all $\tilde{i} \in I$, (86) is maximized by solving the problem

$$
\max_{\eta_0} \left( \eta_0 v^b (s_0, \eta_0) + (1 - \eta_0) v^h (s_0, \eta_0) \right),
$$

so irrespective of the precise objective of the planner, the chosen value for $\eta_0$ always solves this maximization problem. Furthermore, comparing the maximum in this problem across different policy regimes (but for given utility density $u$) yields a Pareto ranking of the resulting planner allocations because a larger maximum implies higher expected utility for all agents $\tilde{i} \in I$ by

\textsuperscript{107}This exists, because $\nu$ is absolutely continuous with respect to the Lebesgue measure and thus its distribution function must be continuous.
equation (86). This completes the proof of the proposition.

D.6 Proofs of Results in Section 7

The proofs of all results in Section 7 are part of Appendix E, where I also present additional details regarding the model extensions briefly outlined in the main text.

E Details on the Model Extensions of Section 7

E.1 Model Extension with Government Debt

I briefly summarize here the changes to the agent problems and the equilibrium and define some notation for this extended model. In the end of this section, I present the proof of Proposition 15.

The household problem is as in the baseline model, except that the household chooses now not just a single money portfolio weight \( \theta_{m,h}^t \), but three portfolio weights \( \theta_{m,h}^t, \theta_{b,1,h}^t, \) and \( \theta_{b,2,h}^t \) for the three types of government debt. The net worth evolution (8) and for portfolio weights summing up to 1 have to be adjusted accordingly, the household faces a no-short-sale constraint for all three types of government debt, \( \theta_{m,h}^t, \theta_{b,1,h}^t, \theta_{b,2,h}^t \geq 0 \). The remaining parts of the household problem remain unaltered, including the definition of velocity \( v_t \) because bond holdings \( \theta_{b,1}^t \) and \( \theta_{b,2}^t \) do not contribute to nominal money balances. To avoid having to add a separate portfolio choice for narrow bank deposits (in the case \( \psi = 0 \)), I assume here for simplicity that \( \theta_{m,h}^t \) includes narrow bank deposits as there are perfect substitutes to outside money from the perspective of households.

The banker problem is largely unchanged, except that now bankers can also hold type-1 bonds, so there is an additional portfolio weight \( \theta_{b,1,b}^t \), an additional no-short-sale constraint \( \theta_{b,1,b}^t \), and the net worth evolution (10) has to be adjusted to reflect this additional asset.

In addition to bankers and households, this extended model also allows for the possibility of a competitive sector of narrow banks if \( \psi = 0 \). One can w.l.o.g. assume that these narrow banks are always present because under fractional reserve banking (\( \psi = 1 \)), bankers can perform all the activities of narrow banks and bankers are already perfectly competitive, so that adding an additional competitive sector of narrow banks does not affect equilibrium allocations or prices. The representative narrow bank has net worth \( N_{nb}^t \) that follows the evolution

\[
\frac{dN_{nb}^t}{N_{nb}^t} = - \frac{D_t}{N_{nb}^t} dt + \theta_{m,nb}^t dR_{ni}^m + \theta_{b,1,nb}^t dR_{b,1}^b,
\]
where $D_t \geq 0$ is a dividend payout process of the narrow bank, $\theta_t^{m,nb}$ is the combined portfolio weight in outside money (enters positively) and narrow bank deposits (enters negatively), so in particular, the narrow bank does not face a no-short-sale constraint for money, and $\theta_t^{b,1,nb}$ is the portfolio weight of the narrow bank in type-1 bonds. Like fractional reserve banks, the narrow bank cannot invest in type-2 bonds for unspecified exogenous reasons. The narrow bank maximizes some objective that depends positively on all future dividend payouts, $\{D_t\}_{t=0}^\infty$. The precise specification of that objective is irrelevant because perfect competition implies $D_t = 0$ in equilibrium.\footnote{108It is only important that the decision problem remains well-defined for $D_t = 0$.}

This model extension introduces two new return processes, $dR_t^{b,1}$ for type-1 bonds and $dR_t^{b,2}$ for type-2 bonds. In addition, the return process for outside money $dR_t^m$ changes because $M_t$ is no longer constant by the government budget constraint. All other return processes remain unaltered. The three return processes on the nominal asset are

\[
dR_t^m = \frac{d(1/P_t)}{1/P_t} = \left( \Phi_i(t_t) - \delta + \mu_t^{q,M} - \mu_t^M \right) dt + \sigma_t^{q,M} dZ_t
\]

\[
dR_t^{b,1} = i_{b,1}^t dt + \frac{d(1/P_t)}{1/P_t} = \left( i_{b,1}^t + \Phi_i(t_t) - \delta + \mu_t^{q,M} - \mu_t^M \right) dt + \sigma_t^{q,M} dZ_t
\]

\[
dR_t^{b,2} = i_{b,2}^t dt + \frac{d(1/P_t)}{1/P_t} = \left( i_{b,2}^t + \Phi_i(t_t) - \delta + \mu_t^{q,M} - \mu_t^M \right) dt + \sigma_t^{q,M} dZ_t
\]

where $q_t^M = \frac{M_t/P_t}{K_t}$ is as before the value of the money stock per unit of capital.

Market clearing conditions of the baseline model are unaltered, except for the outside money market clearing condition (19), which now has to be replaced with

\[
\theta_t^{m,b} N_t^b + \theta_t^{m,h} N_t^h + \theta_t^{m,nb} N_t^{nb} = q_t^M K_t.
\]

In addition to the baseline model, there are now two bond market clearing conditions,

\[
\theta_t^{b,1,b} N_t^b + \theta_t^{b,1,h} N_t^h + \theta_t^{b,1,nb} N_t^{nb} = b_1^1 q_t^M K_t,
\]

\[
\theta_t^{b,2,h} N_t^h = b_2^2 q_t^M K_t,
\]

where $b^M := 1-b_1-b_2^2$ is the fraction of government debt that is outside money. The adjustments to the equilibrium definition should then be straightforward and are omitted here.

The solution of this augmented model and the proof of Proposition 15 still need to be written.
up and will be added in a future version of this paper.

E.2 Model Extension with Monetary Policy

The changes to the model for Propositions 16 and 17 are either minimal or largely identical to the extension discussed in Appendix E.1, so that I can turn immediately to their proofs.

Proof of Proposition 16. Let in this generalized model again $q_i^M K_t$ be the value of the total money stock. With an arbitrary money supply, the value of a single unit of money is $1/P_t = q_i^M K_t/M_t$. The government budget constraint, $dM_t = i^m_t M_t dt$, means that the money growth rate must equal the nominal interest rate paid on money, $\mu_t^M = i^m_t$. This implies for the return on outside money

$$dR_t^m = \frac{i^m_t 1/P_t dt + d(1/P_t)}{1/P_t} = i^m_t dt + \frac{d(q_i^M K_t/M_t)}{q_i^M K_t/M_t} = \frac{d(q_i^M K_t)}{q_i^M K_t},$$

i.e. the return on outside money still satisfies equation (15) as in the baseline model.

Bank deposits continue to be nominally risk-free claims. Let again $i^d_t$ the nominal deposit rate and define $\hat{i}^d_t := i^d_t - i^m_t$ to be the spread between the deposit rate and the money rate. The return on deposits is then

$$dR_t^d = \frac{\hat{i}^d_t 1/P_t dt + d(1/P_t)}{1/P_t} = \hat{i}^d_t dt + \frac{d(q_i^M K_t/M_t)}{q_i^M K_t/M_t} = \frac{d(q_i^M K_t)}{q_i^M K_t},$$

and thus also the deposit return is formally identical to its counterpart in the baseline model (equation (16)) if one replaces $i^d_t$ with $\hat{i}^d_t$.

Clearly, no other return process in the model is affected by monetary policy and because all seigniorage revenues are paid out via interest payments on reserves, there are also no additional transfers to be paid to either agent type. Consequently, all other aspects of the model remain unchanged and thus the real side of the model is precisely the same as outlined and solved in the main text.

Proof of Proposition 17. The assertion follows directly from Proposition 15. Type-1 bonds and
outside money are essentially the same asset when bankers (under fractional reserve banking) or narrow banks (under narrow banking) can issue monetary deposits against them and therefore the share \( b^1 \) of government debt that takes the form of type-1 bonds is irrelevant as long as \( b^2 \) remains constant.

Before turning to the proof of Proposition 18, I discuss how the lending policy and the reserve requirement affect the formal model description. I start with the lending policy. The government makes nominal loans at interest rate \( i^{\ell}_t \geq i^m_t \). If \( L_t \geq 0 \) is the nominal quantity of these loans, then the government budget constraint becomes

\[
dM_t = \underbrace{i^m_t M_t dt}_{\text{nom. interest payments on reserves}} - \underbrace{i^{\ell}_t L_t dt}_{\text{nom. interest earned on loans}} + \underbrace{dL_t}_{\text{money created for new loans}}.
\]

Let \( M^0_t := M_t - L_t > 0 \) be the quantity of outside money that exists in excess of loans. It makes sense to define \( q^M_t := \frac{M^0_t}{P_t} \), because the remaining part of the money stock is not net wealth for private agents as there is an offsetting liability of bankers to the central bank. The law of motion of \( M^0_t \) is

\[
dM^0_t = \frac{dM_t - dL_t}{M^0_t} = i^m_t dt - \underbrace{(i^{\ell}_t - i^m_t)}_{=: \ell_t} L_t dt.
\]

These considerations imply that the return on outside money in this augmented model is\(^{109}\)

\[
dR^m_t = i^m_t dt + \frac{d(q^M_t K_t/M_t)}{q^M_t K_t/M_t} = \left( i^{\ell}_t - i^m_t \right) \ell_t dt + \frac{d(q^M_t K_t)}{q^M_t K_t}.
\]

This adds an additional term \( \left( i^{\ell}_t - i^m_t \right) \ell_t dt \) to the money return (15) in the baseline model, which is only relevant if both \( i^{\ell}_t > i^m_t \) and \( \ell_t > 0 \).

In addition to this changed money return, the presence of loans adds another asset to bankers’ portfolio choice. Bankers can now also choose a portfolio weight \( \theta^{\ell,b}_t \leq 0 \) on loans and the return process for loans is

\[
dR^l_t = i^{\ell}_t + dR^m_t,
\]

\(^{109}\)For the first equation, compare the proof of Proposition 16 above.
while the deposit return is as before

\[ dR^d_t = \delta t + dR^m_t, \]

where now in both return processes one has to take into account the altered money return \( dR^m_t \) relative to the baseline model.

In addition to these changes in the return processes, the market clearing condition for outside money changes. Because \( \vartheta_t = \frac{M^0_t/P_t}{M^0_t/P_t + q_t K_t} \), but the value of all outside money is \( M_t/P_t = (1 + \ell_t)M^0_t/P_t \), the right-hand side in equation (19) has to be replaced with \( (1 + \ell_t)\vartheta_t \).

These are the only changes to the model due to the lending policy. For fractional reserve banking, there is the following simple irrelevance result:

**Lemma 8** (Irrelevance of Lending Policy). Under fractional reserve banking, the equilibrium allocation is unaffected by the choice of \( \{\ell_t\} \).

**Proof.** Under fractional reserve banking, households are indifferent between holding monetary deposits and outside money, so the deposit rate must satisfy \( i^d_t = i^m_t \) by households’ portfolio choice between the two assets. Let now \( \{\ell_t\} \) be an arbitrary admissible central bank lending policy where admissible means \( \ell_t \geq i^m_t = i^d_t \) at all times. The latter inequality implies that central bank lending is never cheaper than deposit funding for bankers. As both types of funding have the same risk, bankers strictly prefer deposit funding when \( i^d_t > i^m_t \) (then \( \ell_t = 0 \)) and they are indifferent to absorb any quantity of central bank lending when \( i^d_t = i^m_t \) (then also \( i^d_t = i^m_t \)). In both cases, the product \( (i^d_t - i^m_t) \ell_t \) is thus zero, so that the return processes on money and deposits are precisely the same as in the baseline model without central bank lending. In addition, in the only case that lending is used at all in equilibrium, \( dR^d_t = dR^d_t = dR^m_t \), so that only the net position that agents hold in the three assets matter. But these net positions are unaffected by the value of \( \ell \). Consequently, all agents face effectively the same decision problems as in the model without lending (up to indifference to absorbing any quantity of central bank loans) and thus make the same choices.

With the previous result, it is easy to show the first part of Proposition 18:

**Proof of Proposition 18, Part 1.** To show that the set of allocations attainable under fractional reserve banking is contained in the set of allocations attainable under narrow banking, it is due to Lemma 8 sufficient to show that there is a lending policy that implements the fractional reserve banking allocation (in the absence of a lending policy). I claim that this indeed the case for the specific policy \( \ell_t = i^m_t \).
Following the model solution steps in Appendix B.1.2, there are only two places in the solution procedure where the banking regime (parameter $\psi$) explicitly matters, in Lemma 4 to determine the deposit rate $i_d^t$ that enters bankers’ portfolio choice, and in the definition of velocity (9) which affects the form of the quantity equation in Proposition 6.\textsuperscript{110} It is therefore sufficient to show that with the lending policy these conditions are the same as under fractional reserve banking ($\psi = 1$).

Start with bankers’ portfolio choice conditions in Lemma 4. These are derived under the assumption that deposits are the marginal funding source of bankers. If bankers marginally fund themselves at a different nominal rate, then this rate must appear in equations (53) and (54) and while the deposit rate is still determined by equation (51) this becomes a pure shadow rate as deposits are not used in equilibrium whenever $i_d^t$ is larger than the marginal funding rate for bankers. Because equation (51) implies $i_d^t - i_m^t = \lambda_t v_t > 0$ under narrow banking,\textsuperscript{111} the deposit rate is indeed always larger than the central bank lending rate $i_{\ell}^t = i_m^t$ and thus central bank lending is always the marginal funding source for bankers. As a consequence, the spread of banks’ funding rate over the reserve rate $i_m^t$ is zero and thus $i_d^t$ in equations (53) and (54) must be replaced with 0. The resulting equations are precisely the equations that are obtained under fractional reserve banking without a lending policy.

Next, due to $i_d^t > i_{\ell}^t$ deposits are not used in equilibrium, $\theta_{d,b}^t = \theta_{d,h}^t = 0$. Assuming w.l.o.g. that bankers do not hold outside money,\textsuperscript{112} all liabilities of bankers are to the central bank and thus the total real value of central bank loans is given by

$$L_t = \theta_{k,b}^t N_t^b - N_t^b \Rightarrow L_t = \frac{M_t^0}{P_t} \Rightarrow \theta_{t}^m \vartheta_t = \kappa_t (1 - \vartheta_t) - \eta_t.$$ From money market clearing, then

$$\theta_{m,h}^t = \frac{(1 + \ell_t) \vartheta_t}{1 - \eta_t} = \frac{\vartheta_t}{1 - \eta_t} + \frac{\kappa_t (1 - \vartheta_t) - \eta_t}{1 - \eta_t}.$$ Here, the first term is the same as $\theta_{t}^{m,h}$ and the second term is the same as $\theta_{t}^{d,h}$ in the model with

\textsuperscript{110}While $\psi$ enters many of the equilibrium equations stated in Section 5.1, all these occurrences of $\psi$ can be traced back to the deposit rate expression in Lemma 4 except for the effect on the quantity equation.

\textsuperscript{111}The Lemma simply says $i_d^t = \lambda_v v_t$, but this is because the baseline model assumes a constant money stock. By Proposition 16, all nominal rates in the baseline model must be replaced with the spreads above the reserve rate $i_m^t$ to obtain valid equations in the model with interest rate policy.

\textsuperscript{112}Bankers are indifferent to hold additional outside money on their balance sheets, but then also the quantity of loans $L_t$ has to be expanded by that amount, so that the net liability to the central bank is still as given in the following equation.
fractional reserve banking and no lending policy. Consequently, when substituting everything into equation (9) with \( \psi = 0 \) to derive the quantity equation along the same lines as in Appendix B.1.2, one obtains the same equation as for fractional reserve banking. This completes the proof that the lending policy \( i_t^\ell = i_t^m \) implements the fractional reserve banking allocation.

To conclude the proof, note that a lending policy with large \( i_t^\ell \) such that \( i_t^\ell > i_t^d \) at all times, implements the same allocation as in the model without lending under narrow banking. The argument is the same as in the proof of Lemma 8: bankers prefer deposit funding to central bank lending and therefore do not demand any additional funds from the central bank. Because the narrow banking allocation is different from the fractional reserve banking allocation, the set of implementable allocations is strictly larger under narrow banking. \( \square \)

I now turn to the reserve requirement (and assume again that there is no central bank lending). With this policy, bankers under fractional reserve banking can issue two types of deposits, monetary ones \( (\theta_t^m) \) and nonmonetary ones \( (\theta_t^n) \) and they face a reserve requirement for monetary deposits,

\[
\theta_t^m + \phi_t \theta_t^n \geq 0.
\]

Households absorb both types of deposits, but prefer monetary deposits because they enter nominal money balances,

\[
v_t = a \frac{\theta_t^k}{\theta_t^m + \theta_t^n}
\]

instead of equation (9). Monetary deposit portfolio weights \( \theta_t^m, \theta_t^n \) must of course satisfy an additional market clearing condition,

\[
\eta_t \theta_t^m + (1 - \eta_t) \theta_t^n = 0.
\]

The situation under narrow banking is the same, except that now bankers face an additional constraint \( \theta_t^m = 0 \), which effectively eliminates monetary deposits. A trivial consequence is the following lemma:

**Lemma 9.** Under narrow banking, the reserve requirement has no effect and the allocation is as in the baseline model presented in the main text.

I now discuss the changes under fractional reserve banking. Denote by \( i_t^n \) and \( i_t^m \) the deposit rates for nonmonetary and monetary deposits. Monetary deposits and outside money are perfect substitutes from the perspective of households so that households’ portfolio choice
immediately implies

\[ i_{t}^{md} \geq i_{t}^{m} \]

and with equality, whenever households hold outside money \((\theta_{t}^{m,h} > 0)\). Furthermore, one can assume w.l.o.g. that fractional reserve bankers are always marginal in both the monetary deposit and the outside money market. Then, \(i_{t}^{md} = i_{t}^{m}\) whenever the reserve requirement is not binding and \(i_{t}^{md} \geq i_{t}^{m}\) otherwise.

In addition, the same arguments as in Lemma 4 imply then that households’ portfolio choice between monetary and nonmonetary deposits determines the equilibrium value of \(i_{t}^{nd}\),

\[ i_{t}^{nd} - i_{t}^{md} = \lambda_{t}v_{t}. \] (89)

Consider next banks’ funding choice between monetary and nonmonetary deposits. By the previous condition, \(i_{t}^{md} \leq i_{t}^{nd}\), so bankers always prefer to fund themselves only through monetary deposits when their reserve requirement is not binding. When their reserve requirement is binding, then to fund one unit of resources with monetary deposits, the banker must issue \(\frac{1}{1-\rho_{t}}\) (real) units of monetary deposits and hold \(\frac{\rho_{t}}{1-\rho_{t}}\) additional (real) units of outside money. The effective nominal funding cost with monetary deposits is thus \(\frac{i_{t}^{md} - \rho_{t}i_{t}^{m}}{1-\rho_{t}}\). Thus, bankers’ portfolio choice between the two funding sources must satisfy

\[ \frac{i_{t}^{md} - \rho_{t}i_{t}^{m}}{1-\rho_{t}} \leq i_{t}^{nd} \]

and with equality whenever bankers use nonmonetary deposits. Combining this with condition (89) yields

\[ i_{t}^{md} \leq i_{t}^{m} + \frac{1-\rho_{t}}{\rho_{t}}\lambda_{t}v_{t} \]

and with equality, whenever nonmonetary deposits are used in equilibrium.

In total, the previous considerations imply

\[ i_{t}^{m} \leq i_{t}^{md} \leq i_{t}^{m} + \frac{1-\rho_{t}}{\rho_{t}}\lambda_{t}v_{t} \] (90)

and the first inequality holds with equality whenever the reserve requirement does not bind, the second inequality holds with equality whenever banks use nonmonetary deposits. It should be clear that except for the special cases \(\rho_{t} = 1\) or \(\lambda_{t} = 0\), only one of the two conditions can hold with equality.
With these preliminary considerations, the following two lemmas are almost immediate.

**Lemma 10.** If \( \varrho_t = 1 \) for all \( t \), then the equilibrium allocation under fractional reserve banking is the same as in the baseline model under narrow banking.

**Proof.** As has been argued previously in the proof of part 1 of Proposition 18, it is sufficient to show that bankers always use nonmonetary deposits as their marginal funding source in equilibrium, the deposit rate \( i_t^{md} \) on these deposits is determined as under narrow banking in Lemma 4, and the quantity equation is the same as in Proposition 6 for \( \psi = 0 \).

For the first claim, note that condition (90) together with \( \varrho_t = 1 \) imply \( i_t^{md} = i_t^m \). Then, condition (89) implies \( i_t^{md} - i_t^m = \lambda_t v_t \), which is indeed as in Lemma 4.

For the second claim, note that with \( \varrho_t = 1 \), the reserve requirement is binding, in particular, \( \theta_t^{md,b} = -\theta_t^{m,b} \). My monetary deposit market clearing then follows as \( \theta_t^{md,h} = \eta_t (1 - \eta_t) \theta_t^{m,h} + \eta_t \theta_t^{m,b} \), so that the numerator in the velocity expression (88) becomes

\[
\theta_t^{m,h} + \theta_t^{md,h} = \frac{1}{1 - \eta_t} \left( (1 - \eta_t) \theta_t^{m,h} + \eta_t \theta_t^{m,b} \right) = \frac{\theta_t}{1 - \eta_t},
\]

where the last equation uses money market clearing (19). This is precisely the same denominator as obtained under narrow banking in the baseline model. Consequently, the quantity equation in this model takes the same form.

**Lemma 11.** If \( \varrho_t = 0 \) for all \( t \), then the equilibrium allocation under fractional reserve banking is the same as in the baseline model without reserve requirements.

**Proof.** For \( \varrho_t \to 0 \), \( i_t^m + \frac{1 - \varrho_t}{\varrho_t} \lambda_t \varrho_t \to \infty \) (except if \( \lambda_t = 0 \), but then reserve requirements are obviously irrelevant because monetary and non-monetary deposits are the same). Because the reserve requirement does not bind for \( \varrho_t = 0 \), the first inequality in (90) must hold with equality, \( i_t^{md} = i_t^m \). But because \( i_t^m < \infty \), it can then not be true that the second inequality in (90) also holds with equality and hence banks do not use nonmonetary deposits. The allocation must therefore be identical to a model where banks obtain funding by issuing only monetary deposits without a reserve requirement. This is precisely the baseline model under fractional reserve banking presented in the main text.

The proof of the remaining part of Proposition 18 is an immediate consequence of the previous lemmas.

**Proof of Proposition 18, Part 2.** By Lemma 9, the only allocation implementable under narrow banking is the narrow banking allocation in the baseline model. Lemma 10 then implies that the
set of implementable allocations under fractional reserve banking is weakly larger than under narrow banking because the narrow banking allocation is also implementable under fractional reserve banking. Lemma 11 furthermore implies that this set must be even strictly large because the baseline fractional reserve banking allocation can also be implemented.

E.3 Model Extension with Bank Runs

In this appendix, I present the model extension with bank runs and give a precise definition of runs. As outlined in the main text, a systemic run in the model is an event in which all depositors simultaneously decide to withdraw their funds from the banking system at the same time. Assuming a vanishingly small (utility) cost of sudden withdrawals, such a run would only be individually rational, if banks were to default on their deposit liabilities conditional on a run occurring. I next derive a condition for this to be the case. If this condition is satisfied, an immediate systemic run and a continuation of the original (no-run) Markov equilibrium at the lower post-run wealth share of bankers is a valid continuation equilibrium. In these cases, I say that the economy is vulnerable to systemic runs.

To derive the vulnerability condition, denote by \( \eta^+ (s, \eta) \) bankers’ wealth share conditional on the pre-run state being \((s, \eta)\) and on a default of all banks. \( \eta^+ (s, \eta) \) can be positive because bankers may be able to rescue some wealth in default as assumed in the main text. After the run, the economy has to be in the Markov equilibrium without runs (as runs are assumed to be unexpected), so the value of outside money in the total wealth numeraire is \( \vartheta^+(s, \eta) := \vartheta (s, \eta^+ (s, \eta)) \). The value of the deposit liabilities of the bank is \( D (s, \eta) := \kappa (s, \eta) (1 - \vartheta (s, \eta)) - \eta \) before the run and as the deposit liability is fixed in nominal terms and the value of money changes by the factor \( \frac{1 - \vartheta^+}{1 - \vartheta} \), the value of deposit liabilities after the run is

\[
D^+ (s, \eta) := \frac{\vartheta^+ (s, \eta)}{\vartheta (s, \eta)} \kappa (s, \eta) (1 - \vartheta (s, \eta)) - \eta.
\]

Next, the value of bank assets before the run is \( A (s, \eta) := \kappa (s, \eta) (1 - \vartheta (s, \eta)) \). These assets are composed of risky claims to capital and thus fixed in capital units. The value of capital changes by the factor \( \frac{1 - \vartheta^+}{1 - \vartheta} \), so that the value of bank assets after the run must be

\[
A^+ (s, \eta) := \frac{1 - \vartheta^+ (s, \eta)}{1 - \vartheta (s, \eta)} \kappa (s, \eta) (1 - \vartheta (s, \eta)) = \kappa (s, \eta) \left( 1 - \vartheta^+ (s, \eta) \right).
\]

\[\text{113}\] Otherwise, there would be still sufficient bank assets left to pay all depositors after the run. Consequently, an individual depositor would be better off not running as this would save the cost of sudden withdrawal without inducing any losses on his deposit claims.
Out of these assets, assets of value $\eta^+ (s, \eta)$ have priority before deposits.\textsuperscript{114} Therefore, the bank defaults on deposits conditional on a run occurring, if deposit liabilities after the run exceed assets after the run less $\eta^+$, $D^+ (s, \eta) > A^+ (s, \eta) - \eta^+ (s, \eta)$, or equivalently,

$$\eta < \left( 1 - \frac{\vartheta (s, \eta)}{\vartheta^+ (s, \eta)} \right) \kappa (s, \eta) (1 - \vartheta (s, \eta)) + \frac{\vartheta (s, \eta)}{\vartheta^+ (s, \eta)} \eta^+ (s, \eta).$$

(91)

This inequality captures run vulnerability. If $\eta^+ (s, \eta) \approx 0$, systemic runs are only possible, if the first factor is positive, which requires $\vartheta^+ (\eta) > \vartheta (\eta)$, i.e. the relative value of money to capital is larger after the run than before the run. The reason is that there are no liquidation costs in the model. Therefore, the only way how a coordinated withdrawal of depositors can change the wealth position of bankers is through asset price changes. This is arguably the relevant feature of a systemic run as opposed to a purely idiosyncratic run on a single institution.

The “priority claim” $\eta^+ (s, \eta)$ of bankers changes the run vulnerability as follows. The direct effects is to increase the vulnerability threshold and thereby vulnerability through the second term in inequality (91). This is the dominant term when $\eta^+$ is large. The indirect effect is to decrease the vulnerability threshold and thereby vulnerability through the effect on $\vartheta^+$ (as $\vartheta$ is a decreasing function in $\eta$). This is the dominant effect for $\eta^+$ close to zero. As a consequence, the vulnerability threshold is first decreasing and then increasing in $\eta^+$ and equals precisely $\eta$ at the maximum value, $\eta^+ = \eta$.\textsuperscript{115} Thus, for each state $(s, \eta)$ there is at most one threshold value $\hat{\eta}^+ (s, \eta) \in [0, \eta)$ such that $\hat{\eta}^+ (s, \eta) = \eta$ and for all $\eta^+ (s, \eta) < \hat{\eta}^+ (s, \eta)$, the vulnerability condition (91) is satisfied whereas for all $\eta^+ (s, \eta) \geq \hat{\eta}^+ (s, \eta)$, it is not satisfied. If no such threshold $\hat{\eta}^+ (s, \eta)$ exists for a given state $(s, \eta)$, then the system is never vulnerable to a run there.

Figure 8 in the main text plots the threshold values $1 - \hat{\eta}^+ (s, \eta)/\eta$, whenever $\hat{\eta}^+ (s, \eta)$ exists (otherwise the region remains white). Because $1 - \eta^+ (s, \eta)/\eta = \frac{\eta - \eta^+ (s, \eta)}{\eta}$ is the proportional loss of bankers in the event of a run measured in the total wealth numeraire, the threshold value $1 - \hat{\eta}^+ (s, \eta)/\eta$ is the smallest proportional loss that makes the run self-fulfilling.\textsuperscript{116}

\textsuperscript{114}They do not literally need to be priority claims of the banker on the bank. This could also be the value of assets the banker holds outside the bank (measured in post-run total wealth units). Alternatively, there could be an ex-post bailout paid to the banking sector as in Mendo (2018) (although this would slightly change the default condition below).

\textsuperscript{115}This statement cannot be proven theoretically because it relies on assumptions on the functional form of the equilibrium solution $\vartheta$, but it can be verified ex post in any numerical example (and holds in any example investigated for this paper).

\textsuperscript{116}Strictly speaking, it is the infimum over all proportional losses $1 - \eta^+ (s, \eta)/\eta$ that are consistent with the run vulnerability condition (91). Because that condition is a strict inequality condition, $\eta^+ = \hat{\eta}^+$ is insufficient for run vulnerability.
E.4 A Microfounded Model for Money Demand in Production

In this appendix, I present a money demand specification based on a simple microfounded version of a cash-in-advance constraint in production that behaves qualitatively and quantitatively as the transaction cost specification used in the main text.\textsuperscript{117}

I modify the model setup as follows. Instead of final output goods, capital $k_t$ managed by each household produces an output flow $ak_t dt$ of a homogeneous, perishable intermediate good. In addition, each household receives infrequent production opportunities that allow the household to transform $y_t^I$ units of intermediate input goods into $y = v^{-1} f \left( \frac{y_t^I}{ak_t} \right) ak_t$ units of final goods. Here, $k_t$ is the quantity of capital the household manages, $y_t^I$ is the input choice of intermediate goods, $v > 0$ is a model parameter and $f$ is a production function that satisfies $f(x) = x$ for $x \leq 1$, $f'(x) > 0$ for all $x \geq 0$, $f''(x) < 0$ for all $x > 1$, and $f'(x) \to 0$ as $x \to \infty$. Production opportunities arrive at the jump times of a household-specific (idiosyncratic) Poisson process $\tilde{J}_t$ with constant intensity $v$.\textsuperscript{118} While idiosyncratic production opportunity risk is here directly tied to the household’s capital holdings for simplicity, this risk cannot be passed on to bankers via risky claims in the same way as idiosyncratic capital shocks $d\tilde{Z}_t$.

The key assumption is that the intermediate goods trade on a Walrasian market where trades have to be settled immediately using money.\textsuperscript{119} This implies that households have to hold money in order to exploit production opportunities once they arise. Let $P_t^I$ be the money price quoted on the intermediate goods market and $P_t$ the money price of final output goods. In the following, use the individual notation $\tilde{i} \in \mathbb{I}$ as in Section 6 for additional clarity and let $y_t^I(\tilde{i})$ denote the planned choice of intermediate input demand of agent $\tilde{i}$ conditional on $d\tilde{J}_t(\tilde{i}) = 1$. When an investment opportunity arrives for household $\tilde{i}$, exploiting it generates real profits

$$
\frac{P_t y_t(\tilde{i}) - P_t^I y_t^I(\tilde{i})}{P_t} = v^{-1} f \left( \frac{y_t^I(\tilde{i})}{ak_t(\tilde{i})} \right) ak_t(\tilde{i}) - \frac{P_t^I}{P_t} y_t^I(\tilde{i}) = v^{-1} \left( f \left( x_t^I(\tilde{i}) \right) - p_t x_t^I(\tilde{i}) \right) ak_t(\tilde{i})
$$

\textsuperscript{117}The counterparts of the figures presented in Section 5.3 are very similar for the specification presented here. These results are available upon request from the author.

\textsuperscript{118}The assumptions on $f$ and the arrival of production opportunities jointly imply that aggregate output of the final good remains $aK_t dt$ as in the baseline model in the main text.

\textsuperscript{119}In the spirit of the New Monetarist literature, one may assume that the intermediate goods market is open while other markets are closed and no trade records can be kept beyond the end of the (here infinitesimal) trading day. The assumption of a Walrasian market is uncommon for the New Monetarist literature and here used as a convenient shortcut to keep the model extension simple. At the expense of additional complexity, one could replace it with a search market that features bilateral bargaining.
with the definitions $x^I_t(\tilde{i}) := \nu y^I_t(\tilde{i}) / (ak_t(\tilde{i}))$ and $p^I_t := P^I_t / P_t$. When no investment opportunity arrived, the households simply sells the output flow $ak_t(\tilde{i}) \, dt$ from capital at the real intermediate goods price $p^I_t$. In total, these considerations imply that the return on capital is no longer given by (12) as in the main text, but instead by

$$dR^k_t(t, x^I_t) = \left( \frac{p^I_t a - \nu}{q^K_t} + \Phi(t) - \delta + \mu^q_K \right) dt + \sigma^q_K dZ_t + v^{-1} \left( f(x^I_t) - p^I_t x^I_t \right) a d\tilde{J}_t.$$  \hspace{1cm} (92)

Instead of the ad-hoc cash-in-advance constraint in production in the main text, households now face an intermediate goods input choice, effectively a choice of $x^I_t$, which is subject to the cash-in-advance constraint

$$P^I_t y^I_t \leq m_t \Leftrightarrow \frac{p^I_t a}{q^K_t} x^I_t \theta^k,h_t \leq v \left( \theta^m,h_t + \psi \theta^d,h_t \right),$$  \hspace{1cm} (93)

where as before $\psi = 1$ under fractional reserve banking and $\psi = 0$ under narrow banking.

The modified capital return (92) and cash-in-advance constraint (93) are the only changes to the household problem relative to the one described in Section 3.3. The banker’s problem remains entirely unaffected by this change in the transaction technology.

The introduction of intermediate goods market adds an additional market whose clearing condition determines the relative price $p^I_t$. More precisely, the total demand $Y^I_t dt$ for intermediate goods must satisfy

$$Y^I_t dt := \int y^I_t(\tilde{i}) d\tilde{J}_t(\tilde{i}) \, d\tilde{i} = \int y^I_t(\tilde{i}) d\tilde{i} \cdot v dt \Rightarrow Y^I_t = v \int y^I_t(\tilde{i}) d\tilde{i}.$$  

By scale invariance of households’ decision problems, all households choose the same value of $x^I_t$, such that $y^I_t(\tilde{i}) = x^I_t a k_t(\tilde{i})$. Substituting this into the previous equation yields

$$Y^I_t = x^I_t a \int k_t(\tilde{i}) d\tilde{i} = x^I_t a K_t.$$  

Total supply of intermediate goods is given by total production $aK_t dt$ and equating this with $Y^I_t dt$ implies that the intermediate goods market clears if and only if

$$x^I_t = 1.$$  \hspace{1cm} (94)
In particular, \( p^t \) adjusts in equilibrium such that all agents choose \( x^t = 1 \). Finally, \( x^t = 1 \) and \( f(1) = 1 \) imply that the aggregate production of final output goods is

\[
Y_t \, dt := \int_1 y_t(i) dJ_t(i) \, d\tilde{\nu} = \int_1 v^{-1} f \left( x^t(I(i)) \right) \, a_k(i) \, d\tilde{\nu} \cdot \, vdt = aK_t \, dt
\]
as in the baseline model. Consequently, the (final) goods market clearing condition remains the same as in the main text. The modifications (92) and (93) to the household problem and the additional clearing condition (94) are therefore the only changes to the model.

To solve this modified model, note that Propositions 4 (representation of \( q^K, q^M \) and \( \iota \)) and 7 (capital allocation condition) and the quantity equation (28) in Proposition 6 and the volatility of \( \eta \) in Proposition 8 remain unaffected by the modification in this section. The only equations that change are the money valuation equation (27) in Proposition 5, which has to be replaced with

\[
\mu^\varphi_t = \rho - \left( \frac{(\eta\sigma^\eta_t)^2}{\eta_t(1 - \eta_t)} + \left( \frac{(1 - \kappa t)^2}{1 - \eta_t} + \frac{\kappa_t^2 \beta^2}{\eta_t} \right)(1 - \tilde{\vartheta}_t)^2 \, \sigma^2_t + (1 - \psi \eta_t) \lambda_t \nu_t \right)
\]

\[
- \frac{\nu^h_t}{1 - \nu^h_t} \left( \tilde{\vartheta}_t + \psi (\kappa_t (1 - \tilde{\vartheta}_t) - \eta_t) \right) \lambda_t \nu_t
\]

(95)

and the condition (29) that determines the price of illiquidity \( \lambda \) in Proposition 6, which has to be replaced with the simpler condition

\[
\lambda_t \nu_t = \left( 1 - \nu^h_t \right) \left( \nu_t - \nu \right)^+.
\]

(96)

and the drift of the state variable \( \eta \) in Proposition 8, which has to be replaced with

\[
\mu^\eta_{\eta_t} = \eta_t \left( \frac{1}{1 - \eta_t} - \sigma^\eta_t \right) \sigma^\eta_t + \eta_t (1 - \eta_t) \left( \frac{\kappa_t \beta}{\eta_t} \right)^2 - \left( \frac{1 - \kappa_t}{1 - \eta_t} \right)^2 (1 - \tilde{\vartheta}_t)^2 \, \sigma^2_t
\]

\[
+ \eta_t (1 - \eta_t) \left( \nu \tilde{\nu}_t - \psi \lambda_t \nu_t \right)
\]

(97)

In these equations, the new variable \( \tilde{\nu}_t \) denotes the price of idiosyncratic jump risk that households require to bear idiosyncratic (compensated) Poisson risk \( dJ_t - vdt \).\textsuperscript{120} It is always a number between 0 (risk neutrality with respect to \( dJ_t \) shocks) and 1. These terms appear,\textsuperscript{120}

\textsuperscript{120} It is formally defined as the negative of the \( dJ_t \)-loading of \( \frac{d\tilde{\nu}_t}{\xi_t} \), where \( \xi_t = e^{-\rho t - \mu_t K_t} \) is again the household’s SDF.
because idiosyncratic production opportunities expose risk-averse households to additional idiosyncratic risk, so that they require an additional risk premium for holding capital risk.

While not very realistic, it is instructive to consider the situation in which households can share idiosyncratic production risk without being able to pool their money holdings when the intermediate goods market is open.\footnote{E.g., suppose households can write contracts that make payments contingent on idiosyncratic $dJ$ shocks when asset markets are open, but they cannot commit to deliveries in the intermediate goods market (where trades happen while asset markets are closed).} In this case, household’s wealth is no longer exposed to idiosyncratic production risk and thus $\tilde{\nu}_t^h = 0$. Equation (95) reduces to its counterpart in the baseline model and the only difference to the main text is that the money premium is determined by equation (96) instead of equation (29). Because these equations are qualitatively very similar, the model presented here can be viewed as an approximate microfoundation of the transaction cost technology used in the main text.