The short rate disconnect in a monetary economy *

Moritz Lenel
Princeton

Monika Piazzesi
Stanford & NBER

Martin Schneider
Stanford & NBER

December 21, 2018

Abstract

In modern monetary economies, most payments are made with inside money provided by payment intermediaries. This paper studies interest rate dynamics when payment intermediaries value short bonds as collateral to back inside money. We estimate intermediary Euler equations that relate the short safe rate to other interest rates as well as intermediary leverage and portfolio risk. Towards the end of economic booms, the short rate set by the central bank disconnects from other interest rates: as collateral becomes scarce and spreads widen, payment intermediaries reduce leverage and increase portfolio risk. Structural change induces low frequency shifts that mask otherwise stable business cycle relationships.

*Email addresses: lenel@princeton.edu, piazzesi@stanford.edu, schneidr@stanford.edu. We thank Oreste Tristani for helpful comments.

1 Introduction

Current research on monetary policy relies heavily on standard asset pricing theory. Indeed, it assumes the existence of real and nominal pricing kernels that can be used to value all assets. Moreover, the central bank's policy rate is typically identified with the short rate in the nominal pricing kernel. With nominal rigidities as in the New Keynesian framework, the central bank then has a powerful lever to affect valuation of all assets – nominal and real – and hence intertemporal decisions in the economy. Focus on this lever makes the pricing kernel a central element of policy transmission.

In spite of its policy relevance, empirical support for monetary asset pricing models has been mixed at best. Models that fit the dynamics of long duration assets, such as equity and long term bonds, often struggle to also fit the policy rate. This is true not only for consumption based asset pricing models that attempt to relate asset prices to the risk properties of growth and inflation, but also for more reduced form approaches such as arbitrage free models of the yield curve. This "short rate disconnect" is typically attributed informally to a convenience yield on short term debt.

This paper proposes and quantitatively assesses a new theory of the short rate disconnect that is based on the role of banks in the payment system. We start from the fact that short safe instruments that earn the policy rate are predominately held by intermediaries, in particular banks and money market mutual funds. We argue that these intermediaries are on the margin between short safe debt and other fixed income claims. We derive new asset pricing equations that relate the short rate to bank balance sheet ratios. We show that these equations account quite well for the short rate disconnect, especially at business cycle frequencies.

Our asset pricing equations follow from the fact that banks issue short nominal debt used for payments. In our model, leverage requires collateral, and the ideal collateral to back short nominal debt is in turn short nominal debt. When such debt becomes more scarce, its equilibrium price rises and the short interest rate falls. In particular, the market short rate disconnects from the short rate of the nominal pricing kernel used to value other assets, such as long term bonds or equity.

Empirically, our approach places restrictions on the joint dynamics of the yield curve and bank balance sheets that we evaluate with US data since the 1970s. Our measure of short rate disconnect is the spread between a "shadow" short rate from a term structure model estimated only with long term rates and the three month T-bill rate. This shadow spread consistently rises at the end of booms. As safe collateral becomes scarce, banks increase the share of risky collateral and thereby have a riskier portfolio overall. At the same time, banks lower risk by reducing their leverage, as our theory predicts.

The model also makes predictions for low frequency patterns. In particular, the 1980s saw a strong increase in the shadow spread that coincided with particularly low bank leverage, which our models predicts both qualitatively and quantitatively. Moreover, the financial crisis of 2008 which induced a large persistent increase in the share of safe short bonds together with a similarly persistent increase in leverage. While our model matches part of this co-movement, it cannot account for all of it without a change in banks' technology of producing deposits. However, incorporating regulatory changes that increase the operating cost of banks can account for a larger shift to safety and leverage.

Our results call into question the traditional account of how monetary policy is transmitted to the real economy. Systematic movement in the shadow spread suggests that the central bank does not control the short rate of the nominal pricing kernel. Its impact on intertemporal decisions of households and firms is thus less direct than what most models assume. Instead, the fit of our bank-based asset pricing equations suggest that transmission works at least to some extent through bank balance sheets. As a result, monetary policy and macroprudential policy are likely to both matter for the course of interest rates.

Formally, our model describes the behavior of a competitive banking sector that maximizes shareholder value subject to financial frictions. We capture the nonfinancial sector by two standard elements: a pricing kernel used by investors to value assets – in particular bank equity – and a broad money demand equation that relates the quantity of deposits to their opportunity cost. We also specify an incomplete asset market structure: banks can invest in reserves, short safe bonds that earn the policy rate, as well as a risky asset that stands in for other fixed income claims, such as loans, available to banks.

The key friction faced by banks is that delegated asset management is costly, and more so if it is financed by debt. We assume that a bank financed by equity only requires a proportional management fee per unit of assets. If the bank also issues deposits, this resource cost per unit of assets increases with bank leverage. One interpretation is that debt generates the possibility of bankruptcy, which entails deadweight costs proportional to assets. Since banks issue short nominal debt, they place a particular value on short nominal debt as collateral. It is this collateral benefit of short debt that generates the short rate disconnect in our model.

We then solve banks' optimization problem and evaluate their first order conditions. We show that there is no disconnect when bank assets are safe: banks only hold reserves and short nominal bonds. More generally, however, the collateral benefit of short bonds generates a wedge between the market short rate and the short rate in the nominal pricing kernel. This wedge is captured by the shadow spread which is high during times when banks have a large share of their portfolio invested in risky assets. During these times, banks do not have much good collateral and therefore place a particularly high value on short nominal bonds relative to other investors in the economy. The banks' optimization problem also implies that when the shadow spread is high, banks counteract the increase in risk on their asset side by reducing risk on their liability side. During these times, banks thus reduce their leverage.

To measure the positions of payment intermediaries, we consolidate bank balance sheets with those of money market funds. These funds also are regularly used for payments by households and corporations. The raw fact that provides evidence for our mechanism is that payment intermediaries have a portfolio share of safe assets as well as a leverage ratio that are both strongly negatively correlated with the shadow spread, both at business cycle frequencies and over longer periods. We define safe assets as assets with short maturity that are nominally safe (such as reserves, vault cash, and government bonds). We further define leverage as the ratio of inside money to total fixed income assets. To measure inside money, we use a broad concept of money that includes money market accounts.

Related literature

Our approach follows the spirit of consumption-based asset pricing pioneered by Breeden (1979) and Hansen and Singleton (1983): we test valuation equations that must hold in general equilibrium, without taking a stand on many other features of the economy, in particular the structure of the household sector and the technology and pricing policy of firms. Since we only require a pricing kernel and a money demand function, our approach is thus equally consistent with the supply side of a real business cycle and of the New Keynesian model: in both cases, the two elements can be derived from representative agent optimization. Our model is also consistent with heterogeneous agent models as long as there is a set of state prices used to evaluate shareholder value of banks.

The short rate disconnect has also been documented in Duffee (1996). The phenomenon is also well known in the literature on arbitrage-free term structure models, which struggle to fit the short end. Our explanation builds on the idea that bonds have a convenience yield. The idea is often formulized with a utility benefit from bonds (Patinkin (1956), Tobin (1963)), analogously to the utility benefit of money (Sidrauski (1967). Recent examples include Bansal and Coleman (1996), Krishnamurthy and Vissing-Jorgensen (2012) and Nagel (2016). Alternatively, bonds can relax constraints associated with making payments, similar to a cash-in-advance constraint for money (Clower 1967). For work along these lines, see Venkateswaran and Wright (2014) and Andolfatto and Williamson (2015).

We share the goal of a growing intermediary-based asset pricing literature to study the relationship between asset prices and balance sheet ratios that hold in equilibrium, without taking a stand on what the rest of the economy looks like. Examples are Brunnermeier and Pedersen (2009), He and Krishnamurthy (2013), Brunnermeier and Sannikov (2014), Adrien, Etula, and Muir (2014), Greenwood and Vayanos (2014), Koijen and Yogo (2015), Bocola (2016), Moreira and Savov (2017), He, Kelly, and Manela (2017), Muir (2017), Haddad and Sraer (2018), and Haddad and Muir (2018). A key difference of this paper and Piazzesi and Schneider (2018) to that literature is that our banks are not investors with their own utility function but instead firms that are owned by households and that maximize shareholder value. We thus endogenize households' decisions to hold some assets directly and others indirectly through banks.

While our theory is based on the scarcity of safe short assets available to banks. While reserves have made up an important share of banks' safe assets in recent years, the scarcity of short safe bonds we study is distinct from the scarcity of reserves. Indeed, the scarcity of reserves is measured by the spread between a short rate and the interest rate on reserves, which was positive before 2008 but negative thereafter. In contrast, the short rate disconnect we document is present both before and after 2008. Our paper is thus only tangentially related to work on bank liquidity management (for example, Bhattacharya and Gale 1987, Whitesell 2006, Cúrdia and Woodford 2011 Reis 2016, Bianchi and Bigio 2014, Drechsler, Savov, and Schnabl 2018, De Fiore, Hoerova, and Uhlig 2018.) Piazzesi and Schneider (2018) consider a model that incorporates both bank liquidity management and a scarcity of bank collateral as in the present paper and derive implications for monetary policy.

In the wake of the recent financial crisis, a growing literature studies monetary policy when banks face financial frictions. One strand assumes that banks have a special ability to lend, and hence add value via positions on the asset side of their balance sheets (for example, Cúrdia and Woodford 2010, Gertler and Karadi 2011, Gertler, Kiyotaki, and Queralto 2012, Christiano, Motto, and Rostagno 2014, Negro, Eggertsson, Ferrero, and Kiyotaki 2017, Brunnermeier and Sannikov 2016, Christiano, Motto, and Rostagno 2012, Del Negro, Eggertsson, Ferrero, and Kiyotaki 2017, and Brunnermeier and Koby (2018). These papers also distinguish assets priced by banks – for example bank loans – from assets priced by households, which include the policy instrument. Policy transmission depends on pass-through from the policy rate (which aligns with households' expected marginal rate of substitution) to the loan rate and hence to bank-dependent borrowers.

In contrast, our model features a short rate disconnect because the policy instrument is priced only by intermediaries. Our model thus says that policy transmission depends on the pass-through from the policy rate to the shadow rate – only the latter aligns with households' expected marginal rate of substitution). Piazzesi, Rogers, and Schneider (2018) show how the disconnect dampens the effects of policy in a New Keynesian model since the central bank no longer has a direct lever to affect households' marginal rate of substitution, and hence any intertemporal decisions of households.

Our paper assumes that banks have a special ability to provide inside money as a medium of exchange. We share this "liability centric" view of banking with e.g. Williamson (2012), Williamson (2016), Hanson, Shleifer, Stein, and Vishny (2015), and Di Tella and Kurlat (2017). As in these papers, banks' portfolio choice in our model is shaped by banks' ability to fund themselves with deposits. In our case, banks value short safe debt as particularly good collateral for inside money, which serves as the only medium of exchange.

2 A model of the short rate disconnect

We study an economy with a single consumption good and an infinite horizon. Competitive banks provide inside money that the nonfinancial sector – households and firms – values as a payment instrument. We do not model in detail what the nonfinancial sector does: Section 2.1 simply summarizes how that sector values assets including inside money. With this approach, we can focus on a model mechanism that is robust to what exactly the "real economy" looks like. Section 2.2 then lays out the problem of the banking system and Section 2.3 derives the key asset pricing conditions that must hold in equilibrium.

2.1 Environment and household preferences

Let M_{t+1} denote the real pricing kernel for the nonfinancial sector. It is a random variable that represents the date t value, in consumption goods, of contingent claims that pay off one unit of the consumption good in various states of the world at date $t+1$, normalized by the relevant conditional probabilities. For example, in an economy with a representative household, M_{t+1} is equal to the household's marginal rate of substitution between wealth at dates t and $t+1$. The price of any asset held by the nonfinancial sector in equilibrium is given by the present value of payoffs – in consumption goods – discounted with the pricing kernel. In particular, the value of a bank is given by the present value of its payout to shareholders, to be described below. Moreover, we think of this pricing kernel as determining real intertemporal decisions in the economy.

Since we are interested in nominal interest rates, it is helpful to introduce additional notation for the valuation of nominal claims. Let P_t denote the price of goods in terms of dollars and define the nominal pricing kernel as $M^{\$}_{t+1} = M_{t+1}P_t/P_{t+1}$. With this change of numeraire, $M^{\$}_{t+1}$ represents (normalized) date t values, in dollars, of contingent claims that pay off one dollar in various states of the world at date $t+1$. We also define a nominal one period safe interest rate by

$$1 = E_t\left[M^{\$}_{t+1}\right](1+i^S_t). \tag{1}$$

We refer to i^S_t, the short rate in the nominal pricing kernel, as the *shadow rate*.

We assume that the nonfinancial sector cannot borrow at the shadow rate. This assumption is sensible as long as private investors cannot issue perfectly safe debt. It implies that the shadow rate serves as an upper bound on the market nominal rate on short safe debt, denoted i_t^B. The two rates are equal only if the nonfinancial sector directly holds short safe debt. The short rate disconnect occurs when the market rate drops below the shadow rate. In this case, the nonfinancial sector perceives short nominal bonds as too expensive and does not hold them directly. As we will see, this scenario is consistent with equilibrium because banks may value short nominal bonds more than the nonfinancial sector.

Finally, consider the valuation of inside money, or deposits, by the nonfinancial sector. We assume the nonfinancial sector relies on deposits to make transactions and is therefore willing to accept an interest rate on deposits i_t^D that is below the shadow rate. The opportunity cost of money $i_t^S - i_t^D$ reflects the value of money for making payments. It is declining in real balances held by the rest of the economy: the marginal benefit of payment instruments is declining in the overall quantity held. Formally, we model the payment benefits as a decreasing convex "money demand" function v:

$$v_t(D_t/P_t) = \frac{i_t^S - i_t^D}{1 + i_t^S}, \tag{2}$$

where D_t denotes the dollar value of deposits, or inside money. The dependence on t here stands in for other forces that affect money demand, for example the level of consumption.

2.2 Payment intermediaries

Payment intermediaries provide inside money to the nonfinancial sector. In the U.S. economy, they consist not only of traditional depository institutions but also of money market funds. We consolidate all payment intermediaries and refer to them as "banks" for short.[1] Banks issue nominal deposits D_t to the rest of the economy and purchase assets worth A_t dollars to back those deposits. They maximize shareholder value. We allow shareholders to freely adjust equity every period and hence focus on a one period ahead portfolio and leverage choice.

Banks have access to three classes of assets: short safe debt that pays the market rate i_t^B, reserves and risky bonds. Reserves are short safe bonds that pay a nominal reserve rate i_t^M set by the central bank. Risky bonds deliver a stochastic real rate of return r_{t+1}^L. We describe a bank's portfolio by its share of reserves in total assets α_t^M as well as the share of other short safe bonds in assets α_t^B. We denote the real rate of return on the bank's asset portfolio by r_{t+1}^α – it is a weighted average of the returns on reserves, safe bonds and risky bonds. We also define

[1] In practice, money market mutual funds keep their assets at custodian banks and rely on the latter's access to Fedwire and other payment systems for their payment services. For an aggregate approach that distinguishes only between a payments intermediary and a nonfinancial sector, it thus makes to consolidate.

bank leverage at date t as the ratio of promised deposit payoffs to assets

$$\ell_t = \frac{D_t(1+i_t^D)}{A_t}. \tag{3}$$

All ingredients of the leverage ratio are known as of date t, so ℓ_t is part of the description of bank policy at date t.

Banks' technology is described by two cost functions. First, we introduce a cost of delegated portfolio management. The idea is that agency problems always entail costs, but that those are compounded when the value of assets falls short of the promised payoff on debt. We thus assume that, for each dollar of assets acquired at date t, the bank incurs an *asset management cost* of $k(\tilde{\ell}_{t+1})$ dollars at date $t+1$, where $\tilde{\ell}_{t+1}$ is an *ex post* measure of leverage, namely the ratio of deposits to the *stochastic* payoff on assets at $t+1$:

$$\tilde{\ell}_{t+1} = \frac{\ell_t}{(1+r_{t+1}^{\alpha})P_{t+1}/P_t}. \tag{4}$$

For given leverage chosen at date t, ex post leverage is high if the nominal return on assets in the denominator is low – a shortfall of assets relative to promised debt.

The function k is strictly increasing and convex in $\tilde{\ell}$.[2]. It starts at $k(0) > 0$: even an equity financed bank incurs some asset management cost. Leverage then raises costs at an increasing rate and a bank without equity is not viable. Convexity of the cost function thus effectively makes the bank more averse to risk than what would be implied by shareholders' pricing kernel M_{t+1} alone. This type of cost can be microfounded by a setup with bankruptcy costs: suppose, for example, banks incur a deadweight cost – a share of assets is lost in reorganization – whenever the return on assets falls below a multiple of debt.

Our second cost function captures the idea that reserves are liquid instruments that help banks meet liquidity shocks. Banks face such shocks because their debt is inside money used for payments. We assume that, for each dollar of deposit issued at date t, the bank incurs a *liquidity cost* of $f(m_t)$ dollars at date $t+1$, where m_t is the ratio of reserves to average depositors' transactions

$$m_t := \frac{\alpha_t^M A_t}{\zeta_t D_t}.$$

The average propensity to use deposits for payments ζ_t is known to the bank at date t. The function f is strictly decreasing and convex and converges to zero as m_t becomes large. The

[2] We impose no condition here to ensure that $\tilde{\ell}$ is below one so that bank equity is positive. Nevertheless, we focus throughout on interior solutions with that property. In our quantitative application, we specify a cost function that slopes up sufficiently quickly for banks to choose leverage below one, as in the data

presence of liquidity costs is not essential for the short rate disconnect to obtain. They are useful, however, to contrast the scarcity of short safe debt that gives rise to the short rate disconnect in our model to the scarcity of reserves that ended with quantitative easing programs.

At date t, a bank acquires A_t dollars worth of assets and issues D_t dollars worth of deposits; shareholders' equity is $A_t - D_t$. It chooses nonnegative assets, deposits as well as nonnegative balance sheet ratios α_t^M, α_t^B and ℓ_t with $\alpha_t^M + \alpha_t^B \leq 1$ in order to maximize the discounted value of payoffs

$$\left(E_t\left[M_{t+1}\left(1-k\left(\tilde{\ell}_{t+1}\right)\right)\left(1+r_{t+1}^{\alpha}\right)\right]-1\right)A_t/P_t+\left(1-E_t\left[M_{t+1}^{\$}(1+i_t^D)\right]-\zeta_t f(m_t)\right)D_t/P_t.$$

Here the portfolio weights α_t^M and α_t^B enter into the return on assets r_{t+1}^{α} and together with leverage determine m_t and ex post leverage $\tilde{\ell}_{t+1}$ according to equation (4). The first term is then the return on assets net of leverage costs and the second term is the interest payment on deposits plus liquidity costs. The bank's objective is homogeneous of degree one in its asset and liability positions – optimal policy determines only balance sheet ratios.

2.3 Bank optimization and bank Euler equations

Shareholder value maximization means that the bank compares returns on potential assets and liabilities to its cost of capital. In a setup with risk, the cost of capital is state-dependent and captured by shareholders' pricing kernel M_{t+1}. For each asset and liability position, the bank thus computes the risk-adjusted return. At an optimum, the risk adjusted return on each asset position has to be less than or equal one – otherwise the bank could issue an infinite amount of equity in order to buy the asset. If the risk-adjusted return is strictly below one, the bank holds zero units of the asset; while it would like to go short, it is not allowed to do so. The risk adjusted return thus has to be equal to one for all assets that the bank holds in equilibrium. Analogously, the risk adjusted return on deposits has to be larger than or equal to one – otherwise the bank would issue an infinite amount of deposits. Banks issue deposits if their risk adjusted return is equal to one.

A key feature of our model is that the asset management cost affects risk adjusted returns. To see this, consider for example the first order condition for assets A_t. Taking the derivative of shareholder value, we have that the risk adjusted overall return on bank assets must be equal to one:

$$E_t\left[M_{t+1}\left(1-k\left(\tilde{\ell}_{t+1}\right)+k'\left(\tilde{\ell}_{t+1}\right)\tilde{\ell}_{t+1}\right)\left(1+r_{t+1}^{\alpha}\right)\right]-\alpha_t f'\left(m_t\right)=1$$

The asset management cost enters in two ways. First, it proportionally lowers the return on assets – this is true even if leverage is zero. Second, an additional dollar of realized return has a *marginal collateral benefit* $k'\left(\tilde{\ell}_{t+1}\right)\tilde{\ell}_{t+1}$: it lowers ex post leverage and hence the asset management cost. In other words, backing deposits with assets makes deposit production cheaper.

Since all individual assets incur management costs and contribute collateral, the cost k enters all bank optimality conditions. To concisely write those conditions, we define the *bank pricing kernel*

$$M^B_{t+1} = M_{t+1}\left(1 - k\left(\tilde{\ell}_{t+1}\right) + k'\left(\tilde{\ell}_{t+1}\right)\tilde{\ell}_{t+1}\right). \tag{5}$$

Intuitively, this random variable describes how bank shareholders value contingent claims held inside the bank. There are two differences to the pricing kernel M_{t+1}: the proportional asset management cost is subtracted, whereas the marginal collateral benefit is added.

The bank pricing kernel clarifies what states of the world are "bad" for the bank (that is, high M^B_{t+1}), and hence what assets represent bad risks for the purposes of bank portfolio choice. Since the bank owes short nominal debt, it is entirely safe if and only if it is "narrow", that is, it holds only short nominal bonds or reserves. In this case, the leverage ratio $\tilde{\ell}_{t+1}$ as defined in equation (4) is constant across states at $t+1$. Indeed, for a narrow bank, the nominal return on bank assets in the denominator is a weighted sum of predetermined nominal interest rates. Short nominal debt is thus good collateral for the bank in the sense that it does not worsen its risk profile. More generally, states are even worse for the bank than for shareholders if the return on bank assets is low.

Using the real bank pricing kernel together with its nominal counterpart $M^{B,\$}_{t+1}$, we rearrange the bank first order conditions with respect to A_t, α^M_t and α^B_t to derive a set of "bank Euler equations". For each of the three available assets – risky bonds, safe short bonds and reserves – the Euler equation says that the risk adjusted expected return should be less or equal to one, with equality if the bank indeed holds the asset:

$$E_t\left[M^B_{t+1}(1+r^L_{t+1})\right] \leq 1, \tag{6}$$

$$E_t\left[M^{B,\$}_{t+1}\right](1+i^B_t) \leq 1, \tag{7}$$

$$E_t\left[M^{B,\$}_{t+1}\right](1+i^M_t) = 1 + f'(m_t). \tag{8}$$

The bank Euler equation for reserves must hold with equality in any equilibrium since only banks can hold reserves. Reserves differ from short safe bonds because of their marginal liquidity benefit $-f'(m_t)$. As a result, banks may wish to hold both in equilibrium: if the bank

Euler equation for bonds holds with equality, then

$$\frac{i_t^B - i_t^M}{1 + i_t^B} = -f'(m_t), \tag{9}$$

that is, the liquidity benefit is equated to the discounted spread between the bond rate and the reserve rate. As the quantity of reserves relative to deposits increases, as it has in recent years for most US banks, then the spread shrinks and may approach zero.[3]

Finally, consider the bank's first order condition with respect to deposits:

$$\frac{i_t^S - i_t^D}{1 + i_t^S} = E_t \left[M_{t+1}^{\$} k'\left(\tilde{\ell}_{t+1}\right) \left(1 + i_t^D\right) \right] + \zeta_t f(m_t) - \zeta_t f'(m_t) m_t. \tag{10}$$

The left hand side is the opportunity cost of deposits to the rest of the economy, or the value of the liquidity provided by deposits. The right hand side is the marginal cost of producing an additional unit of deposits. It consists of a marginal leverage cost as well as marginal liquidity cost. Competitive banks thus equate the price of inside money to its marginal cost.

The presence of the asset management and liquidity cost functions together with the liquidity benefit of deposits for households implies that our model has determinate interior solutions for leverage and portfolio weights. The choice of leverage works much like in the tradeoff theory of capital structure. On the one hand, deposits are a cheap source of funds for banks, since their interest rate is below the short rate in the nominal pricing kernel. On the other hand, issuing debt incurs leverage cost. An interior optimal leverage trades of the two forces. Moreover, portfolio choice is determinate because it affects portfolio risk and hence expected leverage cost. [4]

2.4 The short rate disconnect in equilibrium

We focus on equilibria such that the risky bond is priced by the nonfinancial sector pricing kernel. This might be because the nonfinancial sector can go both long and short in the bond, or alternatively that the outstanding quantity of bonds is so large that it is not only held by banks but is in part held directly. It follows that if the bank also holds risky bonds, then its pricing kernel must similarly price the risky bond. Since its pricing kernel is generally

[3]Piazzesi and Schneider (2017) present a model in which a counterpart of f is derived from banks' liquidity shock distribution. Their formulation implies a threshold for the ratio m_t beyond which f remains constant so that the spread is literally zero. They use this setup to distinguish the abundant reserve regime after 2008 with the scarce reserve regime prevalent before the financial crisis. In the present paper the focus is not on reserve management so this distinction is not critical.

[4]The four equations in (8) and (10) jointly restrict the three bank balance sheet ratios α_t^M, α_t^B and ℓ_t. An equilibrium in which the bank holds all assets thus requires that interest rates align to allow a solution.

different from that of shareholders, its balance sheet ratios must respond appropriately.

Importantly, however, equilibrium does not require that the short rate i_t^B equal the shadow rate i_t^S. To see this, we use the definition of the bank pricing kernel to rearrange the Euler equation for bonds as

$$\frac{1}{1+i_t^B} = \frac{1}{1+i_t^S} + E_t\left[M_{t+1}\left(-k\left(\tilde{\ell}_{t+1}\right) + k'\left(\tilde{\ell}_{t+1}\right)\tilde{\ell}_{t+1}\right)\right]$$

In general, there is a spread between the short rate and the shadow rate given by the risk adjusted difference between the marginal collateral benefit and the asset cost.

If the bank is narrow, that is, it holds no risky bonds, then ex post leverage $\tilde{\ell}_{t+1}$ is predetermined and the spread is zero. In other words, in an economy with narrow banks, there is no short rate disconnect. More generally, however, for a risky bank the asset management cost induces a wedge between the two interest rates. In the next section, we use a particular functional form for the cost function to work out its empirical implications.

3 Quantitative evaluation

In this section we connect the model to the data, and analyze its quantitative fit. We proceed in four steps: first, we provide empirical evidence about the short rate disconnect. We then make additional assumptions on the functional form of the operating cost function and the stochastic distribution of risky returns. These assumptions enable us to derive closed form equations for bank leverage and portfolio choice, which only depend on the shadow spread and the return variance of the risky claim. Next we develop data counterparts of payment intermediaries' leverage and portfolio choice, and compare whether, qualitatively, the model implied co-movements are given in the data. Finally, we estimate the model equations, which allows us to evaluate the model fit quantitatively and to estimate the latent risky return risk.

3.1 The short rate disconnect in the data

This paper argues that the interest rate on nominal safe bonds, such as T-bills, reflects the valuation by payment intermediaries who hold short safe bonds as collateral to back their liabilities. The collateral benefit lowers the observed short rate i_t^B relative to the shadow rate i_t^S that is consistent with the nominal pricing kernel of investors. To obtain a measure of the shadow spread, $i_t^S - i_t^B$, we need a measure of the shadow rate i_t^S.

Our measure of the shadow rate relies on results from Gurkaynak, Sack, and Wright (2007) who estimate forward rates from data on Treasuries. Citing concerns about market segmentation, their paper excludes all Treasury bills from the estimation (point *(iii)* on page 2297).

This exclusion is ideal for our purposes, because we can compute the 3-month rate off their estimated curve, which is consistent with investors' valuation of long Treasury bonds. In other words, we compute the shadow rate from the estimated curve in Gurkaynak, Sack, and Wright (2007).[5] This approach is similar to Greenwood, Hanson, and Stein (2015) who want to measure the convenience yield of T-bills relative to longer Treasury bonds.

Figure (1) plots our measure of short-rate disconnect $i_t^S - i_t^B$ as a black line. The 3-month T-bill rate is the grey line. The sample is quarterly data during the years 1973-2017. NBER recessions are shaded. Broadly speaking, the shadow spread moves with the level of short rates. In particular, the shadow spread consistently rises at the end of booms.

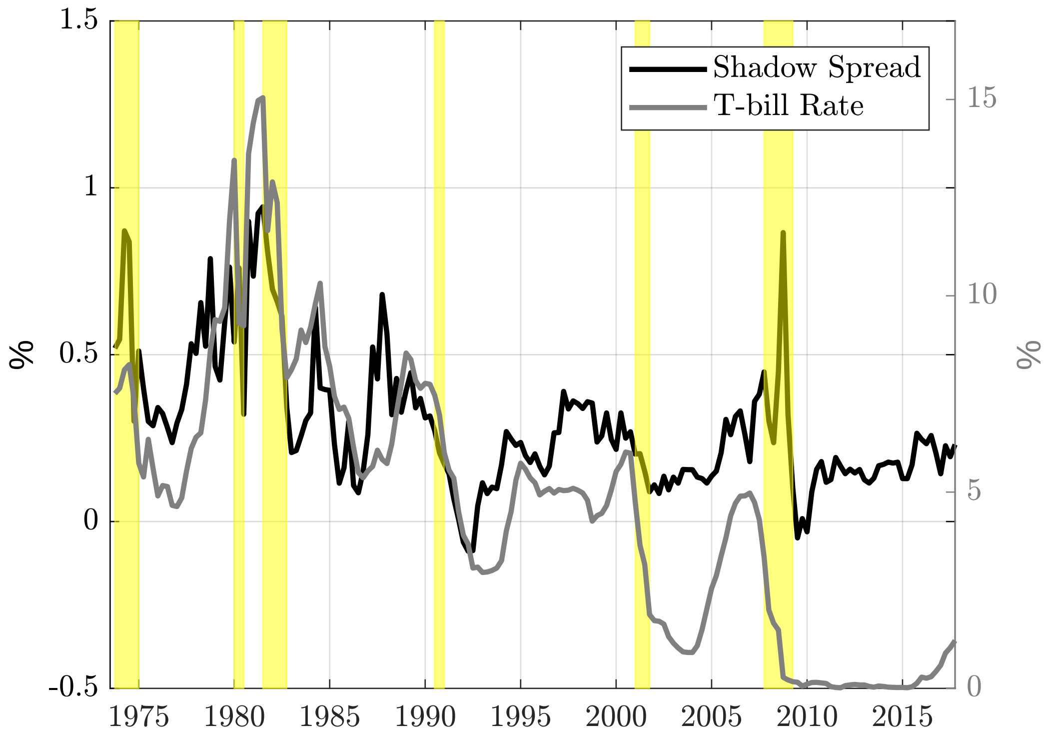

Figure 1: Shadow spread and 3-month T-bill rate. The black line is the difference between the shadow rate from equation i_t^S and the 3-month T-bill rate i_t^B with units measured along the left vertical axis. The grey line is the 3-month T-bill rate with units measured along the right vertical axis. NBER recessions are shaded.

We argue that the short rate disconnect is driven by payment intermediaries' valuation of short safe bonds as collateral. We now provide evidence that suggests that these intermediaries hold T-bills, while households do not hold them directly – only indirectly through

[5]To be precise, we evaluate equation (9) in their paper at maturity 1/4 years using their estimated parameter values.

intermediaries such as money market funds. First, intermediaries buy the lion share of T-bills that are issued by the U.S. Treasury Department on the primary market. While it is possible to buy T-bills directly from the Treasury through its website TreasuryDirect, data from the site shows that between 2008 and 2016, on average only 1.1% of all T-Bills sales went through TreasuryDirect directly to households, and only 1.6% was sold non-competitively in total. All the remaining T-bills were sold in a competitive auction process to primary dealers and other financial institutions. These statistics suggests that households do not buy T-bills in the primary market.

Our second source of information about T-bill holdings are data from the Financial Accounts of the United States. For some sectors of the economy, the Financial Accounts provide a split for holdings of Treasuries into short-term bills and holdings of long-term notes and bonds. The sectors for which we have these data are money market funds, insurance companies, mutual funds (since 2010), the monetary authority, and the rest of the world. Figure (2) depicts the composition of outstanding Treasury bills net of any holdings by the monetary authority and the rest of the world. The shaded areas represent the percentage of outstanding T-bills held by the various sectors. Nonfinancial corporations hold Treasuries mostly for in-house banking purposes; we assume that these holdings are mostly short term and include them into this composition. The top shaded area in the figure consists of T-bills held by "Others" – the remaining T-bills outstanding that are not accounted for by holdings of specific sectors.

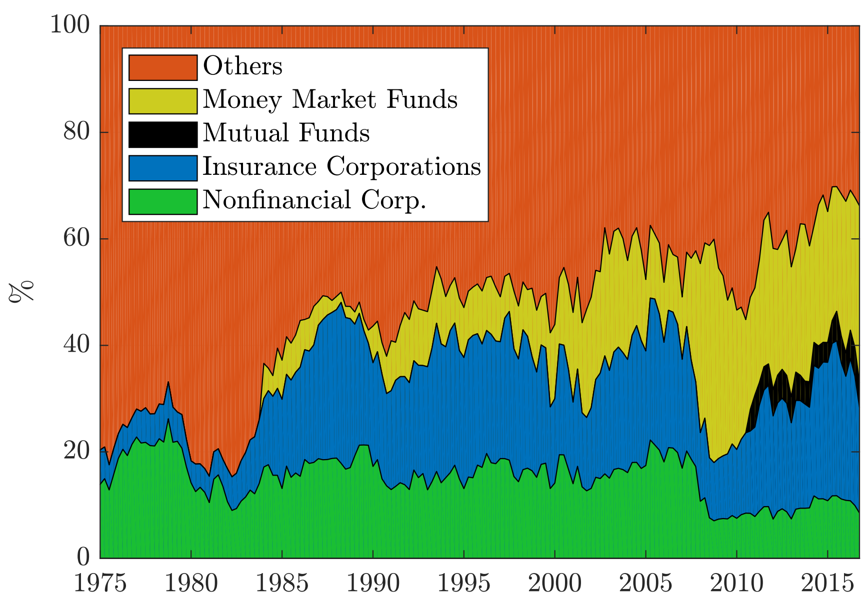

Figure 2: Holdings of T-Bills by Money Market Funds, Mutual Funds, Insurance Corporations, Nonfinancial Corporations and Others. Quarterly data from the Flow of Funds.

Figure (3) plots the time series of T-bill holdings by "Others" and money market funds. The other line shows all Treasury holdings of payment intermediaries — depository institutions, credit unions and banks – including Treasury holdings by money market funds. This figure thus illustrates that Treasury holdings by payment intermediaries are larger than the T-bill holdings that are unaccounted for in Figure (2). Moreover, these Treasury holdings share many of the movements as the series on unaccounted T-bill holdings. This evidence is consistent with payment intermediaries holding all these T-bills.

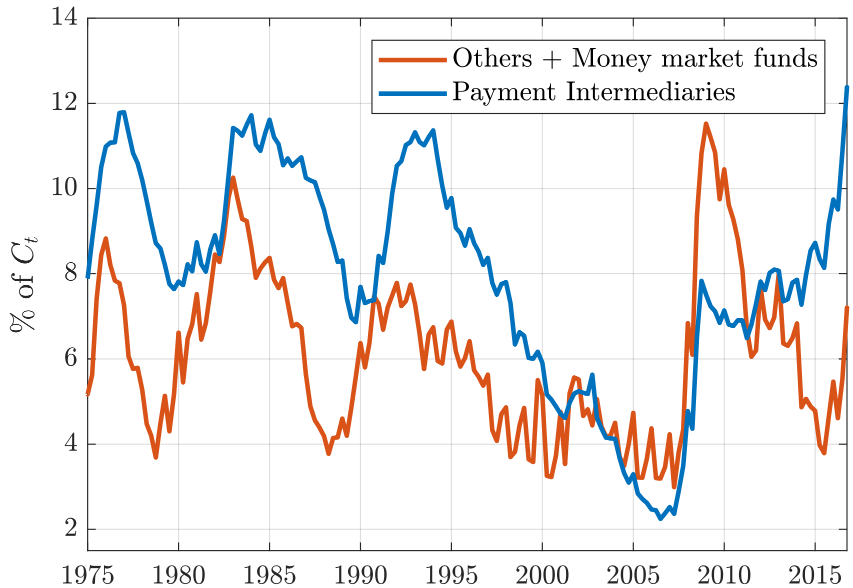

Figure 3: T-Bills held by Others and Money Market Funds, together with all Treasury holdings by Payment Intermediaries.

3.2 Derivation of model equations

To better understand bank choices, we make a functional form assumption on the asset management cost. In particular, we assume that $k(\tilde{\ell}_{t+1})$ is a power function plus a constant, so that

$$k(\tilde{\ell}_{t+1}) = b\left(\bar{k} + \tilde{\ell}_{t+1}^{\gamma}\right). \tag{11}$$

The parameter b scales the overall asset management cost, while an increase in $\bar{k}$ increases only the fixed management fee that is independent of leverage. The parameter γ governs the curvature of the cost function.

The pricing kernel of the bank can then be written as

$$M_{t+1}^{B,\$} = M_{t+1}^{\$}\left(1 - b\left(\bar{k} + (1-\gamma)\tilde{\ell}_{t+1}^{\gamma}\right)\right). \tag{12}$$

As long as k is convex ($\gamma > 1$), this pricing kernel is increasing in ex-post leverage $\tilde{\ell}_{t+1}$, so that the bank puts a higher value on assets that pay off more in states of the world in which its leverage is high.

It is helpful to decompose ex-post leverage $\tilde{\ell}_{t+1}$ into bank leverage ℓ_t at time t, which denotes the ratio of promised deposit repayment relative to the value of asset holdings in period t,

$$\ell_t = (1 + i_t^D) D_t / A_t, \tag{13}$$

and the stochastic nominal portfolio return

$$1 + r_{t+1}^{\alpha,\$} = (1 + r_{t+1}^{\alpha}) P_{t+1} / P_t. \tag{14}$$

We can then write ex-post leverage as $\tilde{\ell}_{t+1} = \ell_t / (1 + r_{t+1}^{\alpha,\$})$, which allows us to separate the bank's leverage decision, which sets ℓ_t, from its portfolio choice, that determines $r_{t+1}^{\alpha,\$}$.

We then summarize the bank's portfolio choice through its safe portfolio share $\alpha_t = \alpha_t^B + \alpha_t^M$ and approximate its portfolio return, $r_{t+1}^{\alpha,\$}$, as

$$1 + r_{t+1}^{\alpha,\$} \approx (1 - \alpha_t)(1 + r_{t+1}^{L,\$}) + \alpha_t (1 + i_t^B). \tag{15}$$

This approximation works well for our data sample, which is split into two different periods of reserves holdings: Before the end of 2007, banks hold a very small fraction of their portfolio in reserves, so that α_t^M is negligibly small. After 2007, banks hold larger amounts of reserves, but the spread between i_t^B and i_t^M disappears, so that a differentiation between reserve and bond shares becomes unnecessary. The latter observation is in line with our model since the marginal liquidity cost $f'(m_t)$ approaches zero as the ratio of reserves to average depositors' transactions m_t becomes large.

Once we make distributional assumptions on the risky return $r_{t+1}^{L,\$}$ we can use the two Euler equations for the risky bond and the safe bond to solve for the bank's leverage choice ℓ_t and its optimal safe asset portfolio share α_t. To do so, we assume that the risky return is log-normally distributed with variance σ_t^2. With this assumption we find the following set of results.

Proposition 1 *Given the functional form assumption (11), the return approximation (15) and a log-normally distributed risky return with variance σ_t^2, the bank's portfolio share of safe assets is given by*

$$\alpha_t \;\approx\; 1 - \frac{1}{\gamma \sigma_t^2} \log \left(1 + \frac{i_t^S - i_t^B}{b\bar{k}} \right),$$

which is decreasing in the shadow-bond spread, $i_t^S - i_t^B$, and increasing in the variance of the risky *return σ_t^2. The bank's leverage choice is given by*

$$\ell_t \approx \exp(\alpha_t i_t^B + (1-\alpha_t) i_t^S) \exp\left(-\frac{1}{2\sigma_t^2}\frac{1}{\gamma}\left(\log\left(1+\frac{i_t^S - r_t^B}{b\bar{k}}\right)\right)^2\right)\ell^*,$$

where $\ell^* = \left(\bar{k}/(\gamma-1)\right)^{1/\gamma}$. Leverage is decreasing in the shadow spread $i_t^S - i_t^B$, increasing in the variance of the risky return σ_t^2, and decreasing in the safe asset share α_t if the shadow spread is strictly positive, $i_t^S > i_t^B$.

The proof of Proposition 1 is in Appendix A. The proposition states that the optimal portfolio share of safe assets is increasing in payoff risk σ_t^2. Intuitively, an increase in the return risk of the risky claim makes it even worse collateral, such that the bank wants to hold less of it. An increase in the shadow spread however increases the cost of holding safe assets to back deposits and therefore lowers the safe portfolio share.

When the shadow spread $i_t^S - i_t^B$ goes to zero, the optimal safe portfolio share goes to one. In this case, optimal leverage is $\ell_t \approx \exp(i_t^B)\ell^*$, which defines the constant ℓ^* in the equation for optimal leverage as leverage of a safe bank.

The equation for optimal leverage in Proposition 1 is at first sight less intuitive, since it implies higher return risk σ_t^2 increases rather than decreases leverage. However, as discussed above, the bank holds in that case a larger share α_t of safe assets, which provide better collateral and enable the bank to increase its leverage. In the appendix we show that if we were to hold the safe portfolio share fixed, the result would be reversed, and as expected, higher risk would lower leverage. The same mechanism is at work when the shadow spread increases, which lowers the the safe asset share and thus the collateral quality of banks' asset holdings, and therefore leverage falls. An increase in the safe portfolio share will also lower $\exp(\alpha_t i_t^B + (1-\alpha_t)i_t^S)$, thereby dampening the effect on ℓ_t. When we estimate the model, we find that this effect is quantitatively small.

3.2.1 Stylized facts

Our model solution from the previous section makes two key predictions: First, the portfolio share α_t of safe assets is decreasing in the shadow spread and increasing in the variance of the risky asset return. Second, leverage ℓ_t is also decreasing in the shadow spread and increasing in the variance of the risky asset. The key intuition is that higher risk or lower collateral cost let the bank choose a safer portfolio, which in turn allows for higher leverage. In the following, we collect data counterparts on leverage by payment intermediaries and their portfolio weight on safe assets to test these model predictions.

Data In the model, a sector of payment intermediaries provides inside money D_t. When quantifying the model, we need to take a stance on the types of assets that we consider to be inside money, or payment instruments, in the data. We take a broad measure of money that includes money market accounts: money of zero maturity (MZM), a time series provided by the Federal Reserve Bank of St. Louis. An advantage of this series is its stable money-demand relationship to interest rates, as documented by Teles and Zhou (2005). Narrower definition of money which do not include money market accounts, such as M1, do not have a stable relationship.

This broader definition of payment instruments also guides our definition of payment intermediaries: we consolidate depository institutions and money market funds. To calculate total asset holdings of payment intermediaries, we use data from the U.S. Financial Accounts (Z.1), aggregating depository institutions (Table L.110) and money market funds (Table L.121). We add up their asset holdings, but subtract short term liabilities of depository institutions with a presumed seniority over deposits (commercial paper and repurchase agreements), because these assets cannot serve as collateral for deposits. We also remove money market checking and savings accounts to consolidate the two sectors and avoid double counting.

To find our data counterpart of leverage ℓ_t, we calculate the ratio of MZM and aggregate payment intermediary asset holdings. We need to multiply this ratio by the deposit interest rate, since we have defined ℓ_t as the ratio of promised repayment in the next period relative to current asset holdings. We use the MZM Own rate provided by the Federal Reserve Bank of St. Louis as our measure of the deposit rate. Our measure of leverage differs from other statistics of bank leverage discussed in the literature, in particular the ratio of bank liabilities to bank assets (or equivalently bank assets to bank equity). First, within the banking sector, we only consider depository institutions for our calculations, not a broader set of banks such as brokers and dealers. Second, we include money market funds, which hold for our purposes highly leveraged but safe portfolios. Third, and most importantly, we only consider deposits in the numerator of our leverage measure, not a broader set of liabilities.

The measure of safe assets aggregates the subset of those assets that are of short maturity and nominally safe. For depository institutions, we assume that vault cash, reserve and Treasury holdings fall into this category. For money market funds, we add holdings of Treasuries, municipal bonds and government agency debt. To the sum of those two measures we also add the net-repo holdings of both sectors, consistent with having subtracted repo liabilities from the total asset measure. The fraction of those safe assets relative to total asset holdings yields our time series of α_t.

As in Section 3.1, we use the interest rate on the 3-months T-bill as well as our 3-months shadow rate measure to calculate the shadow spread $i_t^S - i_t^B$. We evaluate the expression

$\exp(\alpha_t i_t^B + (1-\alpha_t) i_t^S)$ with these two rates as well as the safe portfolio share α_t.

Qualitative model fit The top panel of Figure (4) plots the time series of the safe portfolio share α_t in black against the shadow spread in grey over the sample 1975 to 2017. Even in the raw data, one can detect the negative co-movement between the two time series. The same can be said about the time series of leverage ℓ_t which is depicted in the bottom panel of the same figure. Qualitatively, our model gives predictions that are consistent with the data, namely that episodes of high shadow spreads are associated with a lower safe asset share on banks' balance sheet and lower bank leverage. This co-movement is also present in the period after the financial crisis of 2008, which sees an increase in both the safe asset share and bank leverage.

The evolution of leverage after 2008 highlights the differences between our leverage measure and for example the asset to equity ratio. While capital regulation has forced banks to lower their liability to asset ratio since 2008, the same is not true for the deposit to asset ratio. This observation does not rely on our definition of payment intermediaries, but also holds for commercial banks alone, whose liability to asset ratio has decreased from about 90% before the crisis to about 89% after, but whose deposit to asset ratio has increased from about 64% in the years preceding the crisis to more than 71% in 2017.[6] From our model's perspective, which focuses on the amount of assets that are available to back deposits, the latter is the relevant statistic.

While the co-movements of these three time series are at least qualitatively consistent with the model's mechanisms, these figures do not allow us to evaluate the model's quantitative success. In the next section we therefore study the empirical fit of the model, which will also allows us to back out the time series of return risk σ_t^2, that also affects the leverage and portfolio choice in the model.

[6]Data: Board of Governors of the Federal Reserve System, H.8 Assets and Liabilities of Commercial Banks in the United States.

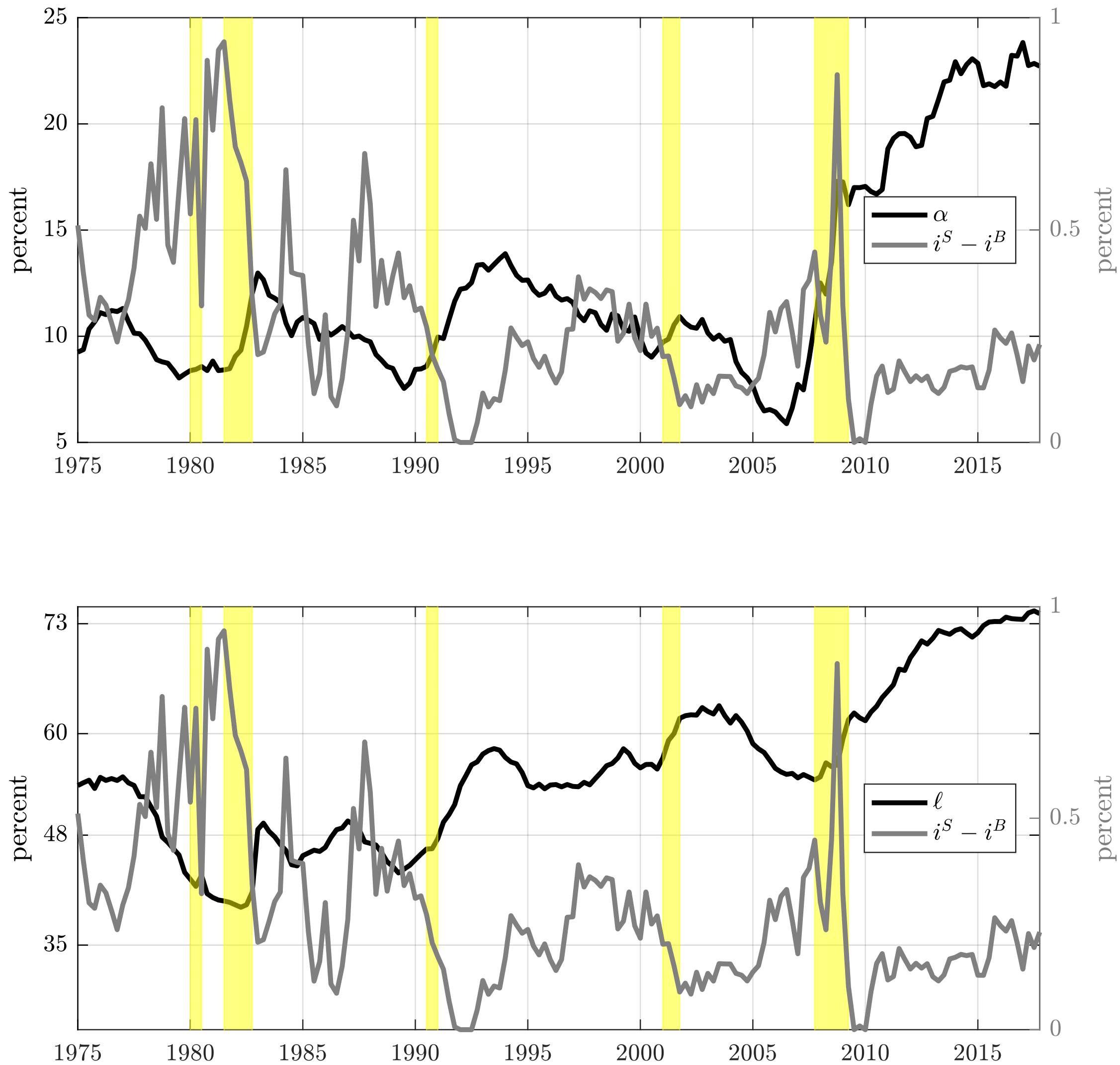

Figure 4: **Top panel:** Safe portfolio share (left axis) and shadow spread (right axis). **Bottom panel:** Leverage ℓ_t (left axis) and shadow spread (right axis). **Data:** i_t^B is the 3 month T-bill rate, i_t^S the shadow rate, data on leverage based on MZM (St. Louis Fed) and payment intermediary asset holdings measured from the U.S. Financial Accounts (Z.1).

3.2.2 Quantitative evaluation of model predictions

Section 3.2 derived two equations for the portfolio share and leverage in terms of the shadow-bond spread and the risky asset's return variance σ_t^2. While payoff risk is an unobserved latent factor, we can use the equation of the portfolio share to replace $\gamma\,\sigma_t^2$ in the leverage equation.

We then find that

$$\ell_t = \exp(\alpha_t i_t^B + (1-\alpha_t) i_t^S) \exp\left(-\frac{1}{2}(1-\alpha_t)\log\left(1+\frac{i_t^S - i_t^B}{b\bar{k}}\right)\right)\ell^*, \tag{16}$$

which states that leverage is, holding the portfolio share fixed, decreasing in the shadow spread, and, holding the spread fixed, increasing in the safe asset share. We can estimate the fit of this equation with data on α_t, ℓ_t, i_t^S, and i_t^B.

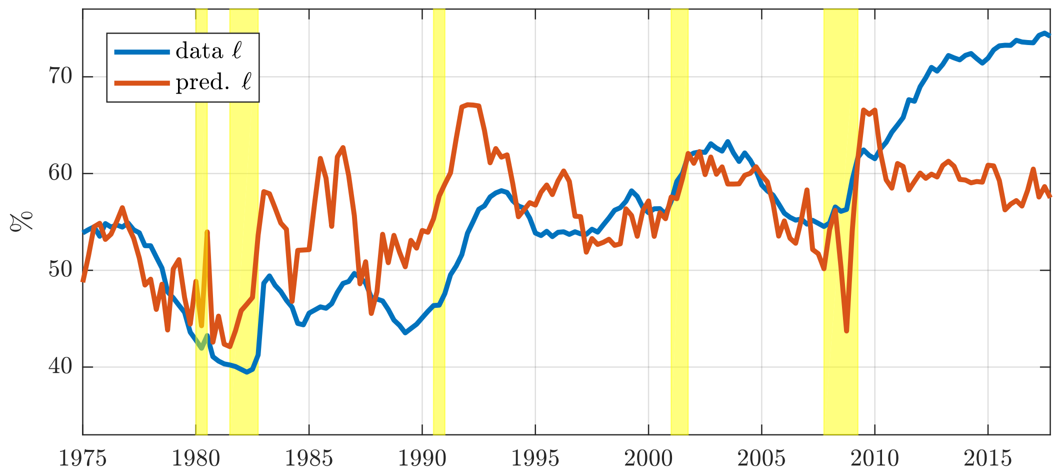

Figure 5: Leverage ℓ_t of payment intermediaries in the data (blue) and model predicted (red) as a function of $i_t^S - i_t^B$ given parameter estimates for $b\bar{k}$, ℓ^* and $\gamma\sigma_t^2$. **Data:** i_t^B is the 3 month T-bill rate, i_t^S the shadow rate, data on leverage based on MZM (St. Louis Fed) and payment intermediary asset holdings measured from the U.S. Financial Accounts (Z.1), see text and appendix.

Estimation We estimate the two parameters, $b\bar{k}$ and ℓ^*, by minimizing the sum of squared residuals of equation (16). The red line in the middle panel of Figure (5) depicts the time series of leverage predicted by the model. While the fit is far from perfect, we find that the model captures dynamics of leverage variation, at least up to the financial crisis in 2007. This can be seen even better when focusing on the cyclical component of leverage in both data in model. To do so, we use a bandpass filter on both the data and the series predicted by the model. The filter isolates business-cycle fluctuations that persist for periods between 1.5 and 8 years. The resulting cyclical components of the two series are shown in Figure (6). The correlation between the cyclical components of data and model is 68%. The deviations in the trend components of the two series could suggest that structural parameters of the banks' asset management cost function change over time. Given the regulatory changes in the banking environment, this is plausible, and in the next section we explore which parameter changes can, for example after the financial crisis of 2007, explain the observed deviations of model and data.

In terms of parameters, we estimate that the fixed fraction of management fees of banks, $b\bar{k}$, has an annualized value of 0.5%. The estimated level of optimal leverage ℓ^* for a safe bank is 67%. Without a structural change in parameters, the model is necessarily unable to fit the level of leverage after the crisis, since the level in the data post-2007 reaches close to 75%, but is, in the model, bounded by ℓ^*.

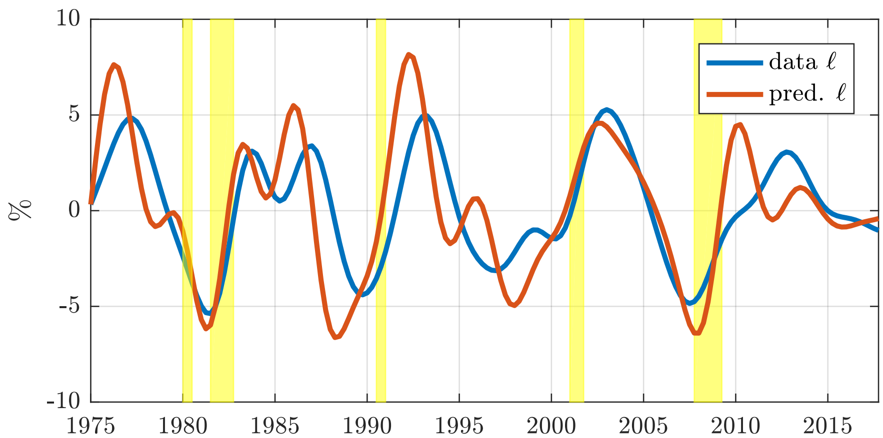

Figure 6: Bandpass filtered time series of data and model predicted ℓ_t.

3.2.3 Structural changes in banks' asset management cost

While our model is successful in capturing the cyclical components in the joint co-movement of safe asset share, leverage and the shadow spread, the overall level of model implied leverage shows at times larger deviations from the data, in particular post-2008. A natural extension of our analysis is to allow for structural changes in the banks' asset management cost function. To do so, we will re-estimate equation 16, but now allowing for time-variation in its two parameters, namely the fixed portion of management fees $\overline{bk}_t$, and the optimal level of leverage ℓ_t^* which banks choose when the shadow spread is zero. We are therefore interested in estimating

$$\ell_t = \exp(\alpha_t i_t^B + (1-\alpha_t) i_t^S) \exp\left(-\frac{1}{2}(1-\alpha_t)\log\left(1 + \frac{i_t^S - i_t^B}{\overline{bk}_t}\right)\right)\ell_t^* + \sigma_R \epsilon_t, \tag{17}$$

where ϵ_t is an independent, standard normally distributed measurement noise shock. We have to recover $\overline{bk}_t$ and ℓ_t^* as latent factors, and assume that both follow random walk processes, so that

$$\overline{bk}_t = \overline{bk}_{t-1} + \sigma_{bk}\eta_t^1$$
$$\ell_t^* = \ell_{t-1}^* + \sigma_\ell \eta_t^2$$

where η_t^1 and η_t^2 are independent shocks with a standard normal distribution.

Given the non-linear measurement equation, we use the unscented Kalman filter to back out the time series of $\overline{bk}_t$ and ℓ_t^*. While it would be possible to jointly estimate the stochastic parameters σ_{bk}, σ_ℓ and σ_R and the time series of the latent factors using maximum likelihood estimation, we find a calibration approach more sensible given the likely misspecification of our simple model. We choose to set σ_{bk} so that the annual standard deviation of the annualized management fee $\overline{bk}_t$ is $10bp$, while σ_ℓ is set so that the annual standard deviation of ℓ_t^* is 1%. Both choices are meant to ensure that these parameters will only vary slowly over time in order to not affect the cyclical fit of the model. We choose starting values $\overline{bk}_0$ and ℓ_0^* by minimizing the sum of squared residuals in the measurement equation 17, and find σ_R iteratively as the value that matches the resulting standard deviation of the residuals in the measurement equation.

The resulting time series of the latent factors are depicted in the upper panel of Figure 7. The thin black lines depict the backed out series of both latent factors. Since we want to make sure that these parameter changes reflect slow moving structural adjustments with no effect on the cyclical fit of the model, we use smoothing splines to further remove any higher frequency fluctuations in these estimates. The green line reflects the smoothed time series of fixed asset holding cost $\overline{bk}_t$, which is in annual terms initially slightly higher than 2%, but quickly declines in the late 70s to about 1%, a level that is roughly constant until the financial crisis of 2008, after which the cost is going back up. The red line denotes the time variation in optimal leverage ℓ_t^*, which banks would choose if the shadow spread was zero, i.e. if collateral were abundant. This series shares the overall dynamics of the evolution of $\overline{bk}_t$.

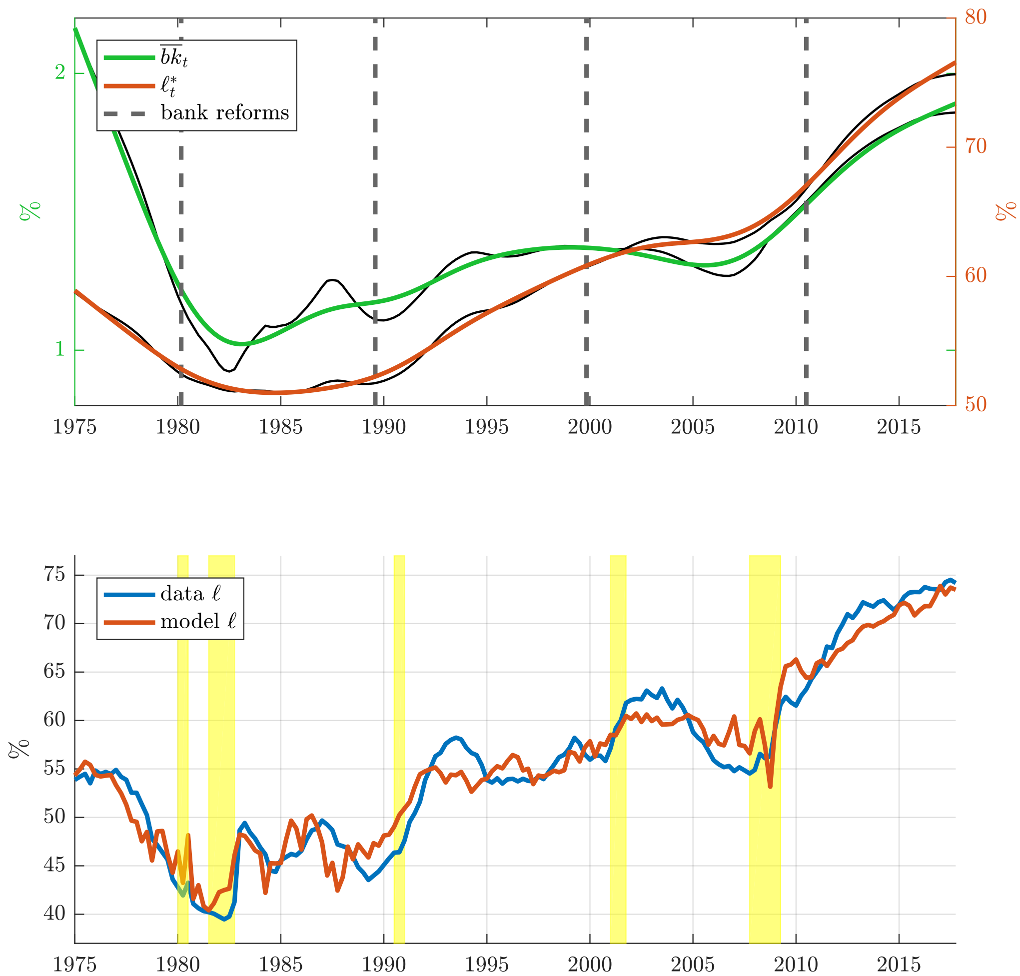

Figure 7: Top panel: Estimated time series of annual fixed asset management cost $\overline{bk}_t$ and maximum leverage ℓ_t^*. The thin black lines are the original results from the unscented Kalman filter. The colored lines are smoothed versions that we are using to evaluate the model fit in the bottom panel. The grey dotted lines mark the 1980 "Depository Institutions Deregulation and Monetary Control Act", the 1989 "Financial Institutions Reform and Recovery Act", the 1999 "Gramm-Leach-Billey Act" and the 2010 "Dodd-Frank Act". Bottom panel: Leverage of payment intermediaries in the data (blue) and model (red).

The lower panel of Figure 7 compares the new model fit to the data. As can be seen, the slow moving structural changes in the cost function parameters lead to an improved model fit over the whole sample, while maintaining the cyclical fit from the previous section. Importantly, the model is now able to match the increase in our leverage measure after 2008. The top panel shows that this increase in leverage is in our results driven by an increase in both $\overline{bk}_t$ and ℓ_t^*. Since $\ell^* = (k/(\gamma-1))^{1/\gamma}$, it is plausible to associate the joint movements between the two series as in increase in k_t, which also provides a rationale for why we might observe that tighter bank regulation after 2008 has led to an increase in leverage ℓ_t: if tighter bank regulation induces higher fixed asset management cost $b_t k_t$, both in absolute and relative terms through increases in k_t, holding assets inside the bank becomes more expensive so that households reduce asset holdings inside the bank by lowering bank equity. This economized production of deposits may also explain the increase of ℓ_t^* after the 1989 "Financial Institutions Reform and Recovery Act", which also tightened bank regulation. Overall we observe that break points in the two latent factor series are roughly associated with the four major bank reforms in the data, lending support to our idea of capturing structural changes in banks' operating cost.

Estimating return risk We can use our results on the time series of $\overline{bk}_t$ to find an implied measure of $\gamma\sigma_t^2$, To do so, we use the equation for the portfolio share to solve for $\gamma\sigma_t^2$ as

$$\gamma\sigma_t^2 = \frac{1}{1-\alpha_t}\log\left(1+\frac{i_t^S - i_t^B}{\overline{bk}_t}\right). \tag{18}$$

We can back out a value for γ by imposing that the average level of σ_t has to match a plausible level of return volatility in the data. We choose to match the average standard deviation of quarterly returns of the SP 500 stock index over the sample period ($\sigma = 7.7\%$), but this is only a choice of scaling that does not affect the dynamics of σ_t. We find that a curvature level of $\gamma = 31.7$ can match the return volatility target,

The resulting time series of return volatility σ_t is depicted in Figure 8. The series varies between 5% and 15%, apart from two outliers in the early 1990s and after the Great Recession of 2008-09 which are presumably driven by measurement noise in our estimate of the shadow spread. The volatility series spikes in episodes of distress in financial markets, namely during the second oil price shock in 1979, the recession episodes and banking crisis of the early 1980s, the stock market crash in 1987, the 1994 peso crisis, the 1997/98 episode of financial turmoil associated with Asia, Russia and LTCM and finally in the years leading up to to the financial crisis of 2007/08. As one would expect, estimated volatility is correlated with the shadow spread, since in times of higher risk, safe collateral becomes more valuable and the shadow spread widens.

Banks' operating cost With our estimate of γ at hand, we can also evaluate whether our estimated cost function is economically reasonable. We use the definition of $\ell_t^* = \left(\frac{\bar{k}_t}{\gamma-1}\right)^{1/\gamma}$ to derive a time series estimate of $\bar{k}_t$. We then find $b_t = \overline{bk}_t / b_t$, which let's us calculate asset

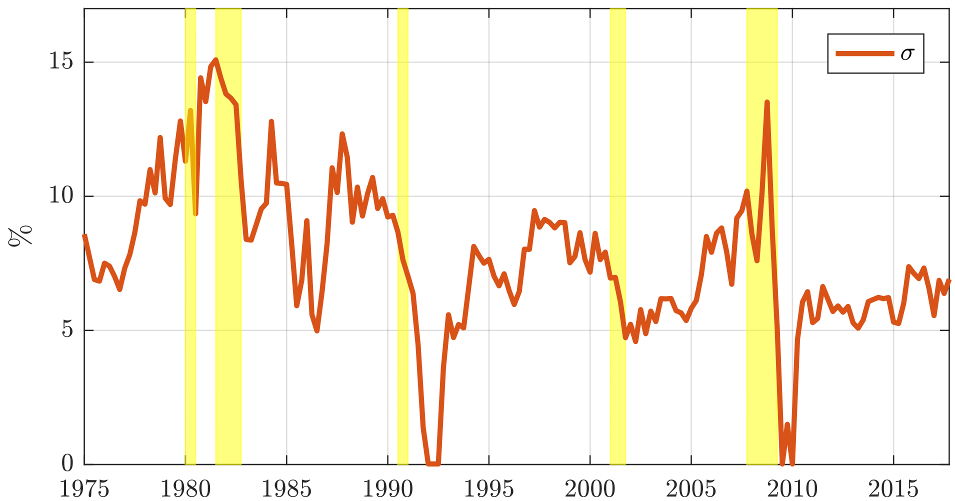

Figure 8: Estimated return volatility given the time series of the shadow spread, the backed out series of $\overline{bk}_t$ and a cost function curvature parameter of $\gamma = 31.7$, which is chosen to match the average quarterly return volatility of the SP 500 (7.7%).

management cost for any level of realized leverage $\tilde{\ell}_{t+1}$ as $k(\tilde{\ell}_{t+1}) = \overline{bk}_t + b_t\tilde{\ell}^{\gamma}_{t+1}$. Figure 9 depicts operating cost for historic levels of leverage, portfolio shares and asset returns, given the estimates γ, b_t, $\bar{k}_t$ and σ_t. The three lines depict asset management cost given the realization of the expected return (blue), the realization of a one standard deviation negative return shock (green) and the realization of a two standard deviation negative return shock (red). Periods in which cost are relatively robust to return shocks are either times of low leverage, high safe asset shares or low return volatility, or a combination of those factors. We find that our estimated cost function yields reasonable levels of asset management cost for plausible return scenarios. For even more negative return shocks the cost can quickly increase due to the high curvature in the cost function.

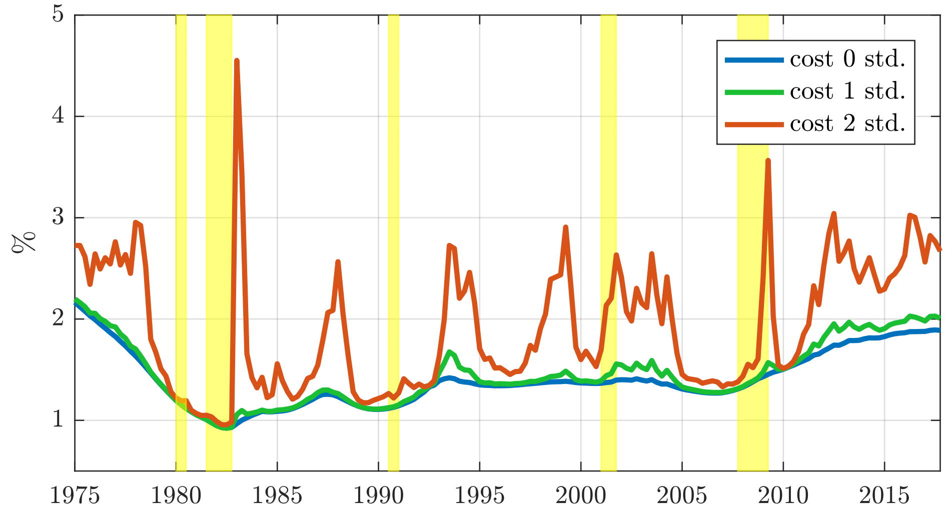

Figure 9: Estimated asset management cost given parameter estimates γ, b_t and $\bar{k}_t$ and given historic choices of leverage ℓ_t, portfolio shares α_t, bond returns, shadow spread and σ_t.

References

Adrien, T., E. Etula, and T. Muir (2014): *Journal of Finance*, LXIX, 2557–2596.

Andolfatto, D. and S. Williamson (2015): "Scarcity of safe assets, inflation, and the policy trap," *Journal of Monetary Economics*, 73, 70–92.

Bansal, R. and W. J. Coleman (1996): "A Monetary Explanation of the Equity Premium, Term Premium, and Risk-Free Rate Puzzles," *Journal of Political Economy*, 104, 1135–1171.

Bhattacharya, S. and D. Gale (1987): "Preference shocks, liquidity, and central bank policy," in *New Approaches to Monetary Economics: Proceedings of the Second International Symposium in Economic Theory and Econometrics*, ed. by W. A. Barnett and K. J. Singleton, Cambridge University Press, International Symposia in Economic Theory and Econometrics, 69–88.

Bianchi, J. and S. Bigio (2014): "Banks, Liquidity Management and Monetary Policy," SSRN Scholarly Paper ID 2496246, Social Science Research Network, Rochester, NY.

Bocola, L. (2016): "The Pass-Through of Sovereign Risk," *Journal of Political Economy*, 124, 879–926.

Breeden, D. T. (1979): "An intertemporal asset pricing model with stochastic consumption and investment opportunities," *Journal of Financial Economics*, 7, 265–296.

Brunnermeier, M. and Y. Koby (2018): "The Reversal Interest Rate," Working Paper, Princeton.

Brunnermeier, M. and Y. Sannikov (2014): "A macroeconomic model with a financial sector," *American Economic Review*, 379–421.

Brunnermeier, M. K. and L. H. Pedersen (2009): "Market Liquidity and Funding Liquidity," *Review of Financial Studies*, 22, 2201–2238.

Brunnermeier, M. K. and Y. Sannikov (2016): "The I Theory of Money," SSRN Scholarly Paper ID 2827562, Social Science Research Network, Rochester, NY.

Christiano, L., R. Motto, and M. Rostagno (2012): "Financial Factors in Economic Production," Working paper, Northwestern.

Christiano, L. J., R. Motto, and M. Rostagno (2014): "Risk Shocks," *American Economic Review*, 104, 27–65.

Clower, R. (1967): "A reconsideration of the microfoundations of monetary theory," *Western Economic Journal*, 1–9.

Cúrdia, V. and M. Woodford (2010): "Credit Spreads and Monetary Policy," *Journal of Money, Credit and Banking*, 42, 3–35.

——— (2011): "The central-bank balance sheet as an instrument of monetarypolicy," *Journal of Monetary Economics*, 58, 54–79.

De Fiore, F., M. Hoerova, and H. Uhlig (2018): "Money Markets, Collateral and Monetary Policy," Working Paper 25319, National Bureau of Economic Research.

Del Negro, M., G. Eggertsson, A. Ferrero, and N. Kiyotaki (2017): "The Great Escape? A Quantitative Evaluation of the Fed's Liquidity Facilities," *American Economic Review*, 107.

Di Tella, S. and P. D. Kurlat (2017): "Why are Banks Exposed to Monetary Policy?" SSRN Scholarly Paper ID 3082257, Social Science Research Network, Rochester, NY.

Drechsler, I., A. Savov, and P. Schnabl (2018): "A Model of Monetary Policy and Risk Premia," *The Journal of Finance*, 73, 317–373.

Duffee, G. R. (1996): "Idiosyncratic Variation of Treasury Bill Yields," *The Journal of Finance*, 51, 527–551.

GERTLER, M. AND P. KARADI (2011): "A model of unconventional monetary policy," *Journal of Monetary Economics*, 58, 17–34.

GERTLER, M., N. KIYOTAKI, AND A. QUERALTO (2012): "Financial crises, bank risk exposure and government financial policy," *Journal of Monetary Economics*, 59, 17–34.

GREENWOOD, R., S. G. HANSON, AND J. C. STEIN (2015): "A Comparative-Advantage Approach to Government Debt Maturity," *The Journal of Finance*, 70, 1683–1722.

GREENWOOD, R. AND D. VAYANOS (2014): "Bond Supply and Excess Bond Returns," *Review of Financial Studies*, hht133.

GURKAYNAK, R., B. SACK, AND J. WRIGHT (2007): "The U.S. Treasury yield curve: 1961 to the present," *Journal of Monetary Economics*, 54, 2291–2304.

HADDAD, V. AND T. MUIR (2018): "Do Intermediaries Matter for Aggregate Asset Prices?" Working Paper, UCLA.

HADDAD, V. AND D. A. SRAER (2018): "The Banking View of Bond Risk Premia," *Working Paper*.

HANSEN, L. P. AND K. J. SINGLETON (1983): "Stochastic Consumption, Risk Aversion, and the Temporal Behavior of Asset Returns," *Journal of Political Economy*, 91, 249–265.

HANSON, S. G., A. SHLEIFER, J. STEIN, AND R. W. VISHNY (2015): "Banks as Patient Fixed-Income Investors," *Journal of Financial Economics*, 117, 449–469.

HE, Z., B. KELLY, AND A. MANELA (2017): "Intermediary asset pricing: New evidence from many asset classes," *Journal of Financial Economics*, 126, 1–35.

HE, Z. AND A. KRISHNAMURTHY (2013): "Intermediary Asset Pricing," *American Economic Review*, 103, 732–770.

KOIJEN, R. S. J. AND M. YOGO (2015): "The Cost of Financial Frictions for Life Insurers," *American Economic Review*, 105, 445–475.

KRISHNAMURTHY, A. AND A. VISSING-JORGENSEN (2012): "The Aggregate Demand for Treasury Debt," *Journal of Political Economy*, 120, 233–267.

MOREIRA, A. AND A. SAVOV (2017): "The Macroeconomics of Shadow Banking," *The Journal of Finance*, 72, 2381–2432.

MUIR, T. (2017): "Financial Crises and Risk Premia," *Quarterly Journal of Economics*, 765–809.

NAGEL, S. (2016): "The Liquidity Premium of Near-Money Assets," *The Quarterly Journal of Economics*, 131, 1927–1971.

NEGRO, M. D., G. EGGERTSSON, A. FERRERO, AND N. KIYOTAKI (2017): "The Great Escape? A Quantitative Evaluation of the Fed's Liquidity Facilities," *American Economic Review*, 107, 824–857.

PATINKIN, D. (1956): *Money, Interest, and Prices: An Integration of Monetary and Value Theory*, Row, Peterson.

PIAZZESI, M., C. ROGERS, AND M. SCHNEIDER (2018): "Money and Banking in a New Keynesian Model," Working Paper, Stanford.

PIAZZESI, M. AND M. SCHNEIDER (2018): "Payments, Credit and Asset Prices," *Working Paper*.

REIS, R. (2016): "Funding Quantitative Easing to Target Inflation," Tech. Rep. 1626, Centre for Macroeconomics (CFM).

SIDRAUSKI, M. (1967): "Rational choice and patterns of growth in a monetary economy," *American Economic Review*, 534–544.

TELES, P. AND R. ZHOU (2005): "A stable money demand: Looking for the right monetary aggregate," *Economic Perspectives*, 50–63.

TOBIN, J. (1963): "Commercial Banks as Creators of 'Money'," Tech. Rep. 159, Cowles Foundation for Research in Economics, Yale University.

VENKATESWARAN, V. AND R. WRIGHT (2014): "Pledgability and Liquidity: A New Monetarist Model of Financial and Macroeconomic Activity," *NBER Macroeconomics Annual*, 28, 227–270.

WHITESELL, W. (2006): "Interest rate corridors and reserves," *Journal of Monetary Economics*, 53, 1177–1195.

WILLIAMSON, S. (2016): "Scarce collateral, the term premium, and quantitative easing," *Journal of Economic Theory*, 164, 136–165.

WILLIAMSON, S. D. (2012): "Liquidity, Monetary Policy, and the Financial Crisis: A New Monetarist Approach," *American Economic Review*, 102, 2570–2605.

A Functional form derivations

This section derive the closed form solutions for leverage ℓ_t and portfolio share α_t. We start from decomposing stochastic ex post leverage $\tilde{\ell}_{t+1}$ into ex ante leverage

$$\ell_t = (1 + i_t^D) D_t / A_t \tag{19}$$

and the nominal risky return

$$1 + r_{t+1}^{\alpha,\$} = (1 + r_{t+1}^{\alpha}) P_{t+1} / P_t, \tag{20}$$

so that

$$\tilde{\ell}_{t+1} = \frac{\ell_t}{1 + r_{t+1}^{\alpha,\$}}. \tag{21}$$

We can now use the Euler equations for the risky asset and the safe bond to solve for the pre-determined component of leverage ℓ_t and the safe portfolio share α_t. Given the functional form assumption, we rewrite the Euler equations for the risky asset and the safe bond as

$$\begin{aligned} b(\gamma-1)\ell_t^{\gamma} &= b\bar{k} E_t \left[M_{t+1}^{\$} (1 + r_{t+1}^{\alpha,\$})^{-\gamma} (1 + r_{t+1}^{L,\$}) \right]^{-1} && (22) \\ 1 &= (1 + i_t^B) \left(\frac{1 - b\bar{k}}{1 + i_t^S} + b(\gamma-1)\ell_t^{\gamma} E_t \left[M_{t+1}^{\$} (1 + r_{t+1}^{\alpha,\$})^{-\gamma} \right] \right) && (23) \end{aligned}$$

Plugging the first equation into the second equation we find

$$(1 + i_t^B) \left(\frac{1 - b\bar{k}}{1 + i_t^S} + b\bar{k} \frac{E_t \left[M_{t+1}^{\$} (1 + r_{t+1}^{\alpha,\$})^{-\gamma} \right]}{E_t \left[M_{t+1}^{\$} (1 + r_{t+1}^{\alpha,\$})^{-\gamma} (1 + r_{t+1}^{L,\$}) \right]} \right) = 1 \tag{24}$$

To solve for α in closed form we use the usual small return approximation $1 + r_{t+1} \approx \exp(r_{t+1})$ and furthermore assume that the nominal risky asset return, defined as

$$1 + r_{t+1}^{L,\$} = (1 + r_{t+1}^{L})(1 + \pi_{t+1}), \tag{25}$$

is log-normally distributed with parameters μ_t and σ_t^2 under the household's nominal risk-neutral measure, which is defined as

$$E_t^* \left[z_{t+1} \right] = E_t \left[\frac{M_{t+1}^{\$}}{E_t [M_{t+1}^{\$}]} z_{t+1} \right], \tag{26}$$

for some random variable z_{t+1}. From

$$\exp(i_t^S) = E_t^* \left[\exp(r_{t+1}^{L,\$})\right] = \exp(\mu_t + 0.5\sigma_t^2) \tag{27}$$

we know that $\mu_t + 0.5\sigma_t^2 = i_t^S$.

We then can then rewrite

$$\begin{aligned}
\frac{E_t\left[M_{t+1}^{\$}\exp(r_{t+1}^{\alpha,\$})^{-\gamma}\right]}{E_t\left[M_{t+1}^{\$}\exp(r_{t+1}^{\alpha,\$})^{-\gamma}\exp(r_{t+1}^{L,\$})\right]} &= \frac{E_t^*\left[\exp(-\gamma r_{t+1}^{\alpha,\$})\right]}{E_t^*\left[\exp(-\gamma r_{t+1}^{\alpha,\$})\exp(r_{t+1}^{L,\$})\right]} \\
&= \frac{E_t^*\left[\exp(-\gamma(\alpha_t i_t^B + (1-\alpha_t)r_{t+1}^L))\right]}{E_t^*\left[\exp(-\gamma(\alpha i_t^B + (1-\alpha)r_{t+1}^L) + r_{t+1}^L)\right]} \\
&= \exp(-(\mu_t + 0.5\sigma_t^2))\exp(\gamma(1-\alpha))\sigma_t^2).
\end{aligned}$$

We plug the above into equation (24) to find

$$\exp(i_t^B)\left((1-b\bar{k})\exp(-i_t^S) + b\bar{k}\exp(-i_t^S)\exp(\gamma(1-\alpha))\sigma_t^2)\right) \;=\; 1 \tag{28}$$

so that we can solve for the safe asset share α_t as

$$\alpha_t \;\approx\; 1 - \frac{1}{\gamma\sigma_t^2}\log\left(1 + \frac{i_t^S - i_t^B}{b\bar{k}}\right).$$

A higher return variance σ_t^2 of the risky asset and more curvature γ in the bank's leverage cost function increases the safe portfolio share. A higher shadow spread $i_t^S - i_t^B$ lowers the safe portfolio share.

We then rearrange the risky bond Euler equation to solve for leverage

$$\begin{aligned}
\ell_t \;=\; & \left(\frac{\bar{k}}{\gamma-1}\right)^{1/\gamma}\exp(\alpha i^B + (1-\alpha)\mu_t \\
& + 1/\gamma(i_t^S - \mu_t - 0.5(1+\gamma^2(1-\alpha)^2 - 2\gamma(1-\alpha))\sigma_t^2)
\end{aligned}$$

which yields

$$\ell_t \;\approx\; \exp\left(i_t^S - \alpha\left(i_t^S - i_t^B\right)\right)\exp(-0.5\gamma(1-\alpha_t)^2\sigma_t^2))\left(\frac{\bar{k}}{\gamma-1}\right)^{1/\gamma},$$

where we have dropped the Jensen term $\exp(0.5(1-\alpha_t)\sigma_t^2))$, assuming that it is sufficiently close to 1.[7] We later verify this assumption using our estimates for $b\bar{k}$ and the range of plausible values for γ, which connect to the Jensen term through

$$\exp(0.5(1-\alpha_t)\sigma_t^2)) = \left(1 + \frac{i_t^S - i_t^B}{b\bar{k}}\right)^{\frac{1}{2\gamma}}. \tag{29}$$

When holding the portfolio share α_t fix, leverage is decreasing in return risk. When plugging in for α_t from above we find the solution from the main text, where leverage is now increasing in risk, as we take into account that the collateral quality is increasing when risk increases.

[7] We try to explicitely estimate γ by including the Jensen term given the solution for α_t, but find that the estimation routine can numerically not differentiate between large values of γ, for which the Jensen term becomes too close to 1.