Banks, Low Interest Rates, and Monetary Policy Transmission*

Olivier Wang†
MIT

Job Market Paper

December 30, 2018

Please check http://economics.mit.edu/grad/o_wang/research for updates.

This paper studies the effect of low interest rates on financial intermediation and the transmission of monetary policy. Using U.S. bank- and branch-level data, I document two new facts: first, the long-run decline in bond rates has not been fully passed through to loan rates; second, the short-run pass-through of policy rates to loan rates is lower at lower rates. To explain these facts, I build a model in which banks provide both credit and liquidity, and the nominal interest rate affects the *composition* of bank interest income between loan and deposit spreads. In the long run, a decline in the equilibrium real rate r^* compresses deposit spreads but increases loan spreads. In the short run, the sensitivity of output to monetary shocks is dampened relative to a benchmark with perfect pass-through, and even more so the lower r^* is: I find a dampening that grows from 20% to 32% as r^* falls from 3% to -1%. A higher inflation target can redistribute from depositors to borrowers and enhance monetary policy transmission.

*I am deeply indebted to Iván Werning, Ricardo Caballero, Alp Simsek and David Thesmar for their guidance and support. I am grateful to Daron Acemoglu, Marios Angeletos, Adrien Auclert, Martin Beraja, Olivier Blanchard, Gabriel Chodorow-Reich, Emmanuel Farhi, Juliette Fournier, Emily Gallagher, Arda Gitmez, Daniel Greenwald, Jonathon Hazell, Zhen Huo, Eben Lazarus, Antoine Levy, Pierre Mabille, Pooya Molavi, Andrés Sarto, Pari Sastry, Ludwig Straub, Rob Townsend, Pierre-Luc Vautrey, James Vickery, Nathan Zorzi, and especially Arnaud Costinot and Jean Tirole, as well as participants at the Stanford SITE 2018 and Chicago Fed Rookie conferences and multiple MIT seminars for very helpful comments and discussions.

†Email: o_wang@mit.edu. I thank the Alfred P. Sloan Foundation and the Macro Financial Modeling (MFM) project for supporting this research. This paper has been prepared by the author under the Lamfalussy Fellowship Program sponsored by the ECB. Any views expressed are only those of the author and do not necessarily represent the views of the ECB or the Eurosystem. All remaining errors are my own.

1 Introduction

Prolonged low interest rates in advanced economies have spurred concerns about declining bank profitability and its macroeconomic consequences. By compressing banks' net interest margins, low rates might lead to weaker balance sheets that hinder intermediation (Committee on the Global Financial System 2018); but lower bank margins could also benefit banks' customers. Moreover, monetary policy transmission through the banking sector may be impaired at low rates: in the extreme case of negative nominal rates implemented in Europe, the pass-through of policy rates to deposit and lending rates appears limited.[1] While nominal rates are currently positive in the U.S., they are likely to remain low in the coming years.[2]

In this paper, I study how persistently low rates affect financial intermediation and the transmission of monetary policy through banks. I first document two new facts about long-run and short-run interest rate pass-through at U.S. banks. I then explain these facts by building a new model of imperfect asset substitutability, featuring commercial banks with a dual role. On the liability side, banks earn deposit spreads by providing liquidity services. On the asset side, banks earn loan spreads by lending to borrowers with no access to capital markets. I argue that due to the *interaction* between the two roles of banks, low rates shift the costs of financial intermediation from depositors—who pay lower deposit spreads—to borrowers—who pay higher loan spreads—and weaken the transmission of monetary policy, even above the zero lower bound.

I motivate my approach with two new facts on the pass-through of bond rates to retail rates at U.S. banks. The first fact shows that the long-run decline in bond rates has not been fully transmitted to the loan rates faced by firms and consumers. Using data on bank income from the Call Reports, I find that over the past 20 years, the maturity-adjusted spread between the return earned by U.S. banks on loans and Treasuries has doubled, in spite of similar credit risk and lower operating costs. Meanwhile, the total spread between the return banks earn on loans and the return they pay on deposits has remained stable. The second fact is about the short-run pass-through of monetary policy shocks. Using data from a large panel of U.S. bank branches, I find that the pass-through of market rates to deposit and loan rates is not only incomplete, as documented in a vast literature, but is also lower at low interest rates.[3]

[1] See among others Heider, Saidi and Schepens (2018) in the Euro area, Eggertsson, Juelsrud and Wold (2017) in Sweden, Basten and Mariathasan (2018) in Switzerland, Hong and Kandarac (2018) in Japan.

[2] For instance, Gourinchas and Rey (2018) predict an average U.S. short-term real rate of -1.37% or -2.35% (depending on the estimation method) for the period 2015-2025. See also Del Negro, Giannone, Giannoni and Tambalotti (2018) for recent estimates of r^* as defined by Laubach and Williams (2015).

[3] Incomplete short-run deposit pass-through has been studied by Hannan and Berger (1991), Neumark and Sharpe (1992), and more recently Driscoll and Judson (2013), Yankov (2014), and Drechsler et al. (2017). Incomplete short-run loan pass-through has been widely documented in the U.S. (Berger and Udell 1992) and in

I then develop a macroeconomic model of financial intermediation to understand these facts. The key mechanism is that nominal interest rates affect the *composition* of bank income. Banks earn the sum of two spreads, reflecting the two sides of the balance sheet. On the asset side, banks earn loan spreads—the difference between loan and bond rates not explained by maturity, credit risk and operating costs. On the liability side, banks earn deposit spreads—the difference between bond and deposit rates. Both spreads can persist due to limits to arbitrage in the form of financial constraints or market power.

The level of nominal interest rates comes into play because savers can substitute between private liquidity (deposits) and public liquidity (currency). Since the nominal interest rate is the opportunity cost of public liquidity, a lower nominal rate makes public liquidity more attractive. Under the plausible condition that money and deposits are sufficiently substitutable (gross substitutes), the equilibrium deposit spread, which is the opportunity cost of private liquidity, must also fall for deposits to be held. Thus lower nominal rates reduce not only public seigniorage, but also the "private seigniorage" earned by banks from deposit creation. If bank lending capacity is high enough relative to credit demand—the *unconstrained lending regime*—the decline in private seigniorage has no consequences for borrowers. If, however, bank lending capacity is already low—the *constrained lending regime*—lower deposit profits spill over to banks' asset side, and loan spreads must rise to reflect the tighter financial constraints on banks.

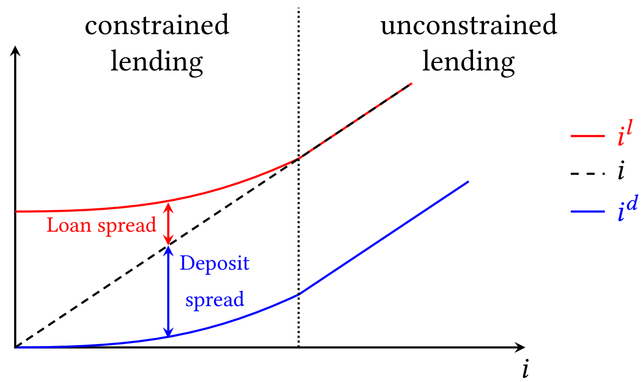
To sum up, if private and public liquidity are gross substitutes, then a decrease in the nominal rate compresses deposit spreads, but widens loan spreads in the constrained lending regime, as depicted in Figure 1. I show theoretically that the behavior of loan and deposit spreads has implications for monetary policy in both the long run, when prices are flexible and the economy reaches its steady state, and in the short run, in the presence of nominal rigidities.

In the long run, for a given inflation target, the economy is in the unconstrained lending regime when the steady state real rate r^* is high, and in the constrained regime when r^* is low enough. Inflation is superneutral in the unconstrained regime, but not when lending is constrained. At high r^*, banks can sustain their long-run required return on equity (determined by entry and exit and the costs of equity issuance) with deposit spreads alone. A classical dichotomy then holds: an increase in inflation inefficiently raises the opportunity cost of liquidity but leaves consumption allocations unchanged.

At low r^*, however, the steady state deposit spread is too low. Banks' retained earnings

Europe (Mojon 2000, De Bondt 2002). Papers on loan pass-through in the U.S. often use as a proxy for loan rates the "bank prime loan rate" reported in release H.15 by the Federal Reserve (see for instance Karagiannis, Panagopoulos and Vlamis (2010) and Rocheteau, Wright and Zhang (2018)). However, as I explain in Section 2.1, the prime loan rate does not reflect the actual rates offered by banks.

Figure 1: Nominal loan rate i^l, bond rate i (45° line) and deposit rate i^d as functions of the nominal bond rate i.



Note: When the nominal interest rate i is low, bank lending is constrained, and a further decrease in i reduces the deposit spread but increases the loan spread.

drop, which depletes their equity. As banks are financially constrained, their lending capacity becomes too low relative to loan demand, and a loan spread opens up to clear the credit market. This process ends when the loan spread is high enough to stabilize banks' earnings and compensate the lower deposit spread. In this regime, a higher inflation target relaxes banks' financial constraints, to the benefit of their borrowers: inflation is not superneutral because it redistributes from depositors, whose opportunity cost of liquidity rises, to borrowers, whose borrowing costs fall.

Turning next to the short run, I add nominal rigidities to explore how incomplete pass-through of bond rates to retail deposit and loan rates also alters the transmission of monetary policy to output. I thereby turn my setting into a tractable heterogeneous agents New Keynesian model with financial frictions. Monetary policy transmission to output is unaffected by the presence of banks and incomplete deposit pass-through in the unconstrained lending regime. By contrast, the sensitivity of output to monetary shocks is dampened in the constrained lending regime, and the more so the lower r^*.

I first use a two-period version of the model to give the key short-run intuitions through closed-form formulas relating deposit pass-through, loan pass-through and the sensitivity of aggregate output to monetary policy. Loan and deposit markets are entangled through banks' balance sheets. Lower deposit pass-through implies lower loan pass-through, which in turn dampens output sensitivity. Under a mild condition on the substitutability between money and deposits, deposit pass-through is lower at lower nominal rates, consistent with my second motivating fact on short-run pass-through. As a result, loan pass-through and the transmission to output are also weaker at low rates.

In addition to these aggregate predictions, the model sheds light on the conflicts of interest that can occur between banks and the rest of the economy at low interest rates. Rate hikes always contract total output in my model: there is no "reversal" of monetary policy. However, rate hikes do have ambiguous effects on bank profits, because monetary shocks have opposite effects on the two components of bank interest income, loan spreads and deposit spreads. Which effect dominates depends on both the health of banks and the level of interest rates when the monetary shock happens. As deposit *rates* are less responsive to policy rates at lower rates, the positive effect of rate hikes on deposit *spreads* becomes stronger, making banks more likely to benefit from those rate hikes. Consistent with this state-dependent relation between the level of interest rates and the sensitivity of bank profits to monetary shocks, I find using high-frequency data that, while unexpected rate hikes hurt U.S. bank stock returns on average, this negative effect is muted or even reversed at low rates.

Next, I quantify the mechanisms uncovered in the two-period model by calibrating the full dynamic model to the U.S. banking sector. The key moments I match are loan and deposit spreads, banks' reported repricing maturity structure of loans, and my estimate of the deposit rate pass-through. Two benchmarks are useful to highlight the macroeconomic role of banks. To control for the "redistribution channel" of monetary policy that arises with heterogeneous agents, I compare my model to a Modigliani-Miller economy, in which all assets are perfect substitutes.[4] To separate the role of loans and deposits, I also also introduce a "credit frictions only" economy where credit frictions remain, but deposits provide no liquidity services. In my baseline calibration where r^* equals 3%, the interest-elasticity of output is 20% lower than in the Modigliani-Miller benchmark, and 8% lower than in the "credit frictions only" benchmark.

My first counterfactual exercise varies the steady state real rate r^* by changing households' discount factor (changes in productivity growth have the same effect). I find that, through lower deposit and loan pass-through, monetary policy transmission is further dampened at low interest rates: when r^* declines from 3% to -1%, the dampening grows from 20% to 35% (see Table 1). By contrast, the interest-elasticity of output barely changes with r^* in both the Modigliani-Miller and "credit frictions only" benchmarks. Therefore, the effect of r^* stems solely from interactions between loan and deposit markets.

I then study the interaction between steady state policies and monetary policy transmission. As in the long run, a higher inflation target can offset the impact of a lower r^* because the steady state *nominal* rate is what matters for *real* spreads. Thus increasing the inflation target not only reduces loan spreads (and increases bank equity) but also enhances monetary policy transmission. This comes at the cost of higher liquidity costs for savers on both money

[4] Auclert (2017) shows that monetary policy has redistributive effects when unhedged interest rate exposures are correlated with marginal propensities to consume, even absent any financial friction.

Table 1: Interest-elasticities, Modigliani-Miller benchmark ("MM") vs. full model ("Banks").

	$r^* = 3\%$		$r^* = -1\%$	
	MM	Banks	MM	Banks
Deposit rate	1	0.7	1	0.4
Loan rate	1	0.8	1	0.5
Output	1.8	1.4	1.8	1.2

Note: Interest-elasticities are defined as $\frac{d\log z_0/d\log R_0}{[d\log z_0/d\log R_0]_{\text{RANK}}}$ for each variable z, where R_0 is the date-0 real interest rate and $[d\log z_0/d\log R_0]_{\text{RANK}}$ is the interest-elasticity of z in the corresponding representative agent New Keynesian model.

and deposits. Second, my model allows to address interactions between financial regulation and monetary policy. I find that holding steady state rates fixed, tighter capital requirements also dampen monetary policy transmission.[5]

An important channel through which monetary policy can affect banks is the revaluation of long-term assets. In my last counterfactual exercise, I evaluate the macroeconomic consequences of banks' maturity mismatch by introducing a third benchmark, in which banks' book equity is fully hedged against monetary shocks. At the baseline average loan duration, revaluation effects make loan pass-through 40% higher than if banks were fully hedged, which increases the interest-elasticity of output by 8%. In the absence of hedging, a tightening of monetary policy has a negative impact on lending capacity through the downward revaluation of outstanding long-term loans, thus the loan spread must increase to clear the loan market. Consequently, the pass-through to loan rates and output is stronger when banks do not hedge interest rate risk.

Related literature

The first contribution of the paper is to study the long-run impact of declining rates on financial intermediation. My long-run results are connected to a recent body of work that tries to measure the (relative) importance of banks' lending and deposit-taking businesses, in terms of profits, returns or stock market valuations. Egan, Lewellen and Sunderam (2017) apply structural methods from production function estimation to the banking sector and find meaningful synergies between deposit-taking and lending in the cross-section of U.S. banks. Schwert (2018) finds that banks earn a large premium over the market price of credit risk.

[5]As I abstract from risk, I cannot speak to the trade-off between the benefits of regulation and potential costs in terms of monetary policy transmission.

Begenau and Stafford (2018) take the opposite view and argue that banks make losses on both sides of their balance sheets, while Drechsler (2018) disagrees. My paper takes no stand on the overall profitability of banks; instead, I examine how the composition of bank profits varies with the level of interest rates.[6]

My work also connects to the recent literature on banks' exposure to monetary policy. Drechsler, Savov and Schnabl (2018) and Di Tella and Kurlat (2017) turn the conventional view of banks' interest rate risk upside down by proposing that maturity mismatch is not a source of interest rate risk, but a hedge, given the behavior of deposit spreads. As nominal interest rates fall, they argue, banks earn lower deposit spreads, thus they find it optimal to hold long-term bonds whose value rises. This mechanism is consistent with the stability of banks' net interest margins and might explain why Begenau, Piazzesi and Schneider (2015) find that U.S. banks do not use derivatives to hedge against interest rate risk. I propose and document a complementary explanation for the stability of net interest margins based on the offsetting behavior of loan and deposit spreads. Accounting for two kinds of spreads allows my model to generate more stable net interest margins than previous work that focuses exclusively on deposit spreads—such as Di Tella and Kurlat (2017)—or on credit spreads—such as Gertler and Karadi (2011). Moreover, unlike hedging through long-term Treasuries, hedging through loan spreads has direct consequences for the cost of credit faced by banks' borrowers.[7]

The second focus of the paper is the transmission of monetary policy through banks. My empirical results on the state-depence of interest rate pass-through connect to work investigating the time-varying effects of monetary policy, for instance by Boivin and Giannoni (2006), Galí and Gambetti (2009), and Boivin, Kiley and Mishkin (2010). The results of this literature (that uses aggregate data and VARs) are ambiguous, partly owing to the high level of aggregation: I use micro-data on bank rates instead, to highlight a specific channel affecting the monetary transmission mechanism. My short-run theoretical results are most closely related to Drechsler, Savov and Schnabl (2017). In their model, banks with market power over deposits optimally contract deposit supply following a monetary tightening in order to earn a higher deposit spread. If it is costly to replace deposits with wholesale funding, loan supply contracts as a side effect. I follow their lead in putting deposits at center stage, but highlight a complementary mechanism: higher bank profits on deposits may boost banks' lending capacity. My model remains within the extant macro-finance literature by assuming perfect

[6] Aiyagari, Braun and Eckstein (1998) and English (1999) find a positive cross-country correlation between inflation and financial sector size in developed economies. Both papers offer a similar explanation: there are only bonds and money (no deposits), but households choose to make purchases with money or use the transaction services provided by banks. In a high inflation regime, households substitute towards bank transaction services.

[7] In Section 2.1, I explain why it is not enough to suppose that banks earn income from deposit provision and maturity transformation to match the data.

competition with financial constraints.[8] Thus incomplete pass-through to loan and deposit rates follows from general equilibrium effects instead of monopoly pricing.[9] In Appendix D, I discuss how the combination of financial constraints and substitution between money and deposits allows my model to match both the findings of Drechsler et al. (2018) and those of Gomez, Landier, Sraer and Thesmar (2016), who show that the "income gap" (the difference between short-term interest income and expense) predicts the response of bank lending to monetary shocks. If banks with "stickier" deposit rates hold assets with longer durations to hedge interest rate risk, then these banks have a higher duration gap but a relatively lower income gap, and thus contract lending by more in response to rate hikes.

The implementation of negative rates in Europe has led to a growing literature that examines the costs and benefits of this new policy. Rognlie (2016) models the trade-off between the expansionary effects of negative rates and the inefficient subsidy to money they imply. Eggertsson, Juelsrud and Wold (2017) show how in the presence of credit frictions, negative rates might not even be expansionary. Bonds and deposits are perfect substitutes in their model so there is no deposit spread, but the pass-through to loan rates breaks down once deposit rates hit zero. Brunnermeier and Koby (2018) go even further, by arguing that there exists a "reversal rate" under which a marginal decrease in the policy rate becomes contractionary. They emphasize market power in loan markets, noting that "one of the most striking features of [their] reversal result is that it does not rely on stickiness of the deposit rate". Relative to this literature, my paper puts incomplete deposit pass-through at its heart to study the less extreme case of positive, but potentially low, nominal rates. Moreover, in both the long and short run, I concentrate on the effect a persistently low real rate r^*, unlike Rognlie (2016) and Eggertsson et al. (2017) who consider transitory shocks that push the economy against the zero lower bound, and unlike Brunnermeier and Koby (2018) who analyze the non-linear effects of large monetary shocks holding r^* fixed.

In terms of modeling, I follow closely Gertler and Karadi (2011), who embed the banking model of Gertler and Kiyotaki (2010) in a standard New Keynesian DSGE. They only model the credit side of banks in order to evaluate quantitative easing policies, while I incorporate liquidity provision as a central function of banks, as well as maturity mismatch, viewed as a crucial link between banks and conventional monetary policy. In an alternative approach,

[8] I build on Gertler and Kiyotaki (2010) for the asset side of banks and Di Tella and Kurlat (2017) for the liability side.

[9] Within a pure market power framework, it is difficult to explain why the loan spread is higher at low interest rates in the long run. If anything, low interest rates should erode banks' market power on deposits and therefore lead them to increase both deposit and loan supplies, which would reduce loan spreads. Some form of financial constraint on banks is thus needed to explain why low profitability on the deposit side may hurt lending. While I could combine such a constraint with the standard model of imperfect competition among banks, abstracting from market power altogether makes my contribution clearer.

Goodfriend and McCallum (2007) and Cúrdia and Woodford (2016) model credit spreads as arising from a loan production function instead of financially constrained banks.

The seminal papers on the "bank lending channel" of monetary policy, such as Bernanke and Blinder (1988; 1992) and Kashyap and Stein (1995), relied on reserve requirements. I share Van den Heuvel (2002)'s emphasis on capital requirements instead. His model only addresses the credit side of banks and is cast in a partial equilibrium setting to zoom in on banks' portfolio choices. Bianchi and Bigio (2017) show how loan spreads depend on central bank policies such as the rate paid on reserves, in a model where banks face withdrawal shocks and search frictions in interbank markets lead them to hold precautionary reserves.[10] Piazzesi and Schneider (2018) discuss the interplay between inside and outside money in a two-layered model: inside money (deposits) facilitates end-user transactions, as in this paper, but outside money (reserves) is used in interbank transactions. Hence there is no substitution or competition between inside and outside money. I focus on the interactions between this role and the lending business. Zentefis (2018) shows that the pass-through of monetary policy to loan rates can break down when banks have too little capital to compete with each other in a Salop model. Gerali et al. (2010) build a DSGE model with a banking sector with sluggish adjustment of retail rates due to Calvo frictions in rate-setting. I confirm the sluggishness empirically, but assume it away in the model to concentrate on the pass-through at a one-year horizon.

The current environment has given rise to many empirical studies on the effects of low interest rates on banks, especially in Europe. Claessens et al. (2018) study the relationship between interest rates and net interest margins. Altavilla, Boucinha and Peydro (2017) examine how standard and non-standard monetary policy measures affect European banks' profitability. In contemporaneous work, Ampudia and den Heuvel (2018) also explore the state-dependent impact of monetary shocks on bank stock prices in Europe. Their results are similar to those I find for U.S. banks in Section 5.3.

My paper is mainly about banks, but the presence of heterogeneous agents connects my model to recent work on monetary policy in incomplete markets, such as Werning (2015), Auclert (2017), and Kaplan, Moll and Violante (2018). This strand of the literature features a different kind of market incompleteness, namely Bewley-Huggett-Aiyagari models with idiosyncratic risk and precautionary savings. I simplify the household side to focus on banks, but my paper demonstrates how several insights can be transposed to macro-finance models.

Finally, this paper studies inside and outside liquidity. Since Holmström and Tirole (1998), several papers have pointed out the crowding-out effect of higher public liquidity supply on

[10] In Appendix D.3, I extend my baseline model to incorporate excess reserves. As in Bianchi and Bigio (2017), this allows me to study whether excess reserves can crowd out lending.

private liquidity provision, e.g., Krishnamurthy and Vissing-Jorgensen (2012), Greenwood, Hanson and Stein (2015), and Nagel (2016). Crowding-out can have social benefits when private liquidity provision entails negative externalities like fire sales. By contrast, I highlight that crowding-out may also have a cost in terms of higher credit spreads.

Outline

Section 2 presents three key stylized facts on U.S. banks that motivate my model, presented in Section 3. In Section 4, I illustrate the interactions between banks, monetary policy and financial frictions, when prices are flexible and bank capital reaches its steady state ("the long run"). In Section 5, I consider the transmission of monetary shocks by introducing nominal rigidities and aggregate demand ("the short run"). I derive analytical results in a two-period setting, before turning to a quantitative analysis in Section 6. Appendix D covers extensions of the baseline model.

2 Evidence on long-run and short-run pass-through in the U.S.

In this section, I document two new facts about U.S. banks. First, the secular decline in bond rates has not been fully transmitted to loan rates faced by consumers and firms. The reason is a shift in the composition of bank interest income: deposit spreads have shrunk while loan spreads have widened. Second, in the short run, the pass-through of policy rates to retail deposit and loan rates is lower at low rates.

Data.

I use three main data sources:

Bank data. I obtain quarterly income and balance sheet data for all U.S. commercial banks from the Call Reports. I use the period 1997Q2-2018Q2, for which the reports contain detailed information on the repricing maturity structure of assets and liabilities.

Retail rates. Weekly data on loan and deposit rates are collected across U.S. bank branches by RateWatch. My sample runs from 1998 to 2018 for deposits, and 2000 to 2018 for loans. Following Drechsler et al. (2017), I restrict attention to branches that actively set rates. I use representative products in each category of deposits that appears as "liquid assets" in the

Survey of Consumer Finance: checking deposits, savings deposits, and money market deposit accounts. For loan rates, I use three types of short-term consumer loans: adjustable rate mortgages (with 1 year maturity), personal loans (24 months), and auto loans (36 months).

Monetary shocks. The series for unanticipated monetary shocks are from Nakamura and Steinsson (2018). Shocks are defined as changes in market expectations of the Fed funds rate (over the remainder of the month, because Fed funds futures settle on the average rate over the month) in a 30-minute window around FOMC anouncements.[11] The sample is all regularly scheduled FOMC meetings from 01/01/2000 to 3/19/2014, excluding the peak of the financial crisis from July 2008 to June 2009.

2.1 Long-run incomplete pass-through of lower bond rates

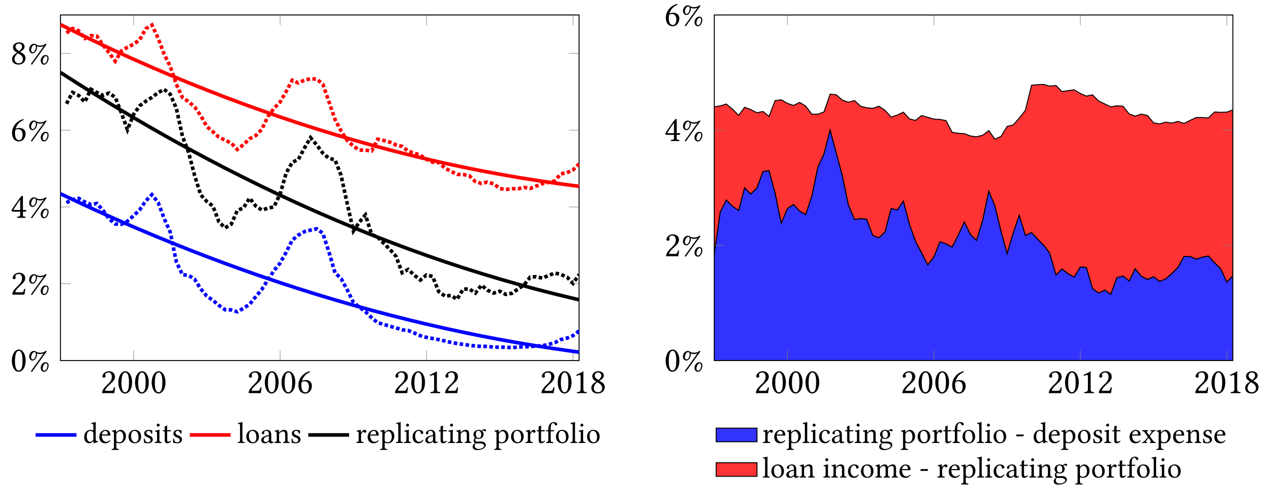
I begin by showing that the steady decline in interest rates over the past 20 years has only been partially passed through to loan rates.[12] The reason is a rise in the maturity-adjusted loan spread between loans and bonds, that mirrors a decline in the deposit spread between bonds and deposits.

Figure 2 decomposes the difference between loan interest income (as a fraction of total loans) and deposit interest expense (as a fraction of total deposits) for U.S. banks between 1997Q2 (when banks started reporting the repricing maturity of their assets in the Call Reports) and 2018Q2. The left panel shows the realized returns on loans and deposits. These measures do not reflect the rates on *new* loans and deposits, which I will look at in Section 2.2. Instead, they are the interest accruing from past loans and deposits, using book value accounting. The corresponding total spread, shown on the right panel as the sum of the red and blue areas, has been remarkably stable in spite of a large decline in interest rates. This is consistent with Drechsler et al. (2018)'s work on the stability of the net interest margin (NIM), which also includes interest income from securities and interest expense on wholesale liabilities. In fact, this "loan-deposit spread" has been *more stable* than the NIM, which has fallen from 4.4% in 1997Q2 to a low of 2.95% in 2015Q1, before slightly rebounding to 3.2% in 2018. The reason is that the return on securities has fallen by more than the return on loans, consistent with my theory in which banks are the marginal pricers of loans but not of bonds.

[11] The construction is detailed in Nakamura and Steinsson (2018). The shock is the first principal component of several futures.

[12] A common view is that the "bank prime loan rate", posted by the Federal Reserve among its selected interest rates in release H.15, is a good indicator of effective lending terms. For instance, Rocheteau, Wright and Zhang (2018) regress the prime rate on the Fed funds rate to measure the short-run pass-through of monetary policy. Figure 16 in Appendix A.1, however, shows that the bank prime loan rate has been mechanically set at a 3% markup over the Fed funds rate since around 1994 (when the Fed started releasing statements about its Federal

Figure 2: Decomposing the difference between the return earned on loans and the return paid on deposits.



Note: "Loan rate" is loan income divided by total loans. "Deposit rate" is deposit expense divided by total deposits. "Replicating portfolio" is described in the text. The left panel shows rates and quadratic trends. The right panel shows spreads by normalizing the deposit rate to zero: the blue area is the difference between the return on the replicating Treasury portfolio and the deposit rate, and the red area is the difference between the loan rate and the return on the Treasury portfolio. Sources: Author's calculations from Call Reports and Federal Reserve data.

Table 2: Loan and deposit spreads, relative to Treasury portfolio.

	1997Q2-2007Q4	2010Q1-2018Q2
Loan spread	1.71%	2.84%
Deposit spread	2.60%	1.57%
Sum	4.31%	4.41%

To correct for the term premia included in loan rates, I construct a Treasury portfolio that replicates the repricing maturity of the loan portfolio, computed from the Call Reports as in English et al. (2018) and displayed in Figure 13 in the Appendix.[13] I then decompose the total loan-deposit spread using the Treasury portfolio's interest income, recorded at book value to match the accounting convention on loans. The "loan spread", in red on the right panel of Figure 2, represents the return on a strategy that borrows Treasuries to invest in loans with the same maturity. The "deposit spread", in blue, represents the return on a maturity mismatched strategy that borrows at the average deposit rate to invest in the Treasury portfolio. The loan spread has widened by around 1%, while the deposit spread has shrunk by the same amount, see Table 2.[14]

Risk premium and operating costs. Figure 18 in Appendix A.1 shows that credit risk and operating costs cannot explain the discrepancy: loss provisions were high during the Great Recession but have reverted (since around 2012) to the same levels as in the 2000-2008 period, while operating costs, measured by non-interest expense, have fallen by around 1% of earning assets. Even holding credit risk constant, part of the higher excess loan spread I find could be due to a higher risk premium. Indeed, the right panel of Figure 2 is reminiscent of the increasing equity risk premium described by Caballero, Farhi and Gourinchas (2017). Figure 15 in Appendix A.1, however, shows the same pattern as in Figure 2 for very short-term, low risk, commercial and industrial loans. This suggests that risk premia are not the full story for bank loans.

Corporate bonds. The increase in credit spreads is specific to bank loans. Figure 17 in Appendix A.1 shows the corporate bond spread constructed by Gilchrist and Zakrajsek (2012), updated through August 2016. Its 2010-2016 average equals 2.3%, almost exactly the same value as the 1997-2007 average of 2.2%.

---

funds rate target) and is of little relevance for the actual loan rates faced by consumers and firms.

[13]For simplicity, I aggregate the four bins in the Call Reports (less than 1 year, 1 to 3 years, 3 to 5 years, 5 to 15 years and more than 15 years) into two bins: "short-term" (1 year) and "long-term" (10 years) bonds. It is important to keep track of both short-term and long-term assets in banks' portfolios. Begenau and Stafford (2018) and Drechsler et al. (2018) approximate the return on bank assets with passive strategies holding only 6-year and 10-year Treasuries, respectively. However, Figure 14 in Appendix A.1 shows that the return on a portfolio with only long-term (10 year) bonds differs significantly from the return on the mixed portfolio I construct with more detailed information. Figure 14 also shows the return on bank securities reported by banks: it is almost identical to the return on my mixed portfolio, even though the mixed portfolio has a shorter average duration than a pure long-term bond and the average repricing maturity of securities is *higher* than that of loans.

[14]Table 5 in Appendix A.1 shows the relationship between bond rates (defined as returns on the Treasury portfolio) and spreads.

2.2 Short-run pass-through of monetary shocks: lower at lower rates

Next, I estimate whether the *short-run* pass-through of monetary shocks to retail rates (at a 1-year horizon) depends on the level of interest rates. I use Jordá (2005)'s local projection method and estimate the following regressions equations at horizons $h = 0, \ldots, 12$ months:

$$y_{b,t+h} - y_{b,t-1} = \alpha_{b,h} + \delta_{1h}\Delta i_t + \delta_{2h} i_{t-1} + \beta_h \Delta i_t \times i_{t-1} + \gamma_h \text{controls}_{b,t-1} + \epsilon_{b,t+h} \tag{1}$$

Dependent variables y_b are the branch-level retail rates on various types of deposits and loans, $\alpha_{b,h}$ are branch fixed effects for each horizon regression, Δi_t is the monetary shock from Nakamura and Steinsson (2018) cumulated at the monthly level, and normalized to have a +100 bps impact on the 1-year Treasury rate at a 12-months horizon. I control for 4 lags of retail and Treasury rates, and in order to isolate the role of the level of interest rates from that of cyclical conditions, I also control for 4 lags of unemployment and GDP growth.[15] Since the economy might have been affected by other changes (for instance demographic changes as in Wong 2018, or higher concentration in the banking sector), I add linear and quadratic time trends. Finally, I add interactions of these controls with the monetary shock.

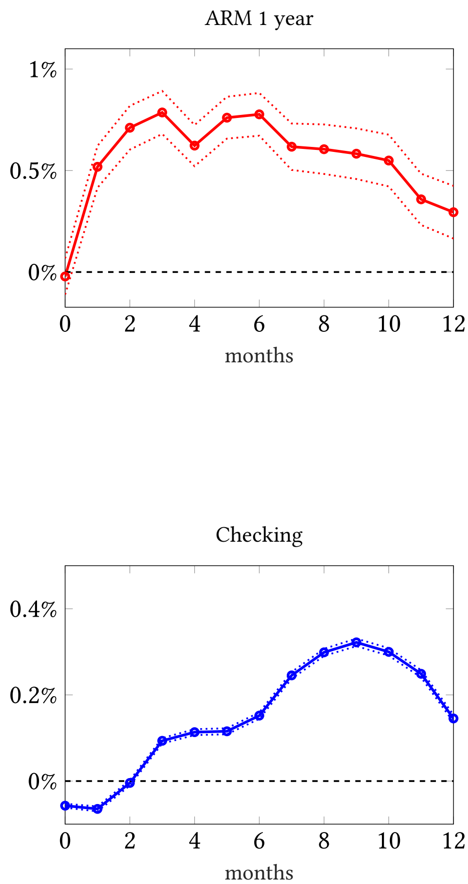
The sequence of estimates $\left\{\hat{\beta}_h\right\}_{h=0,\ldots,12}$ traces out the *relative impulse response* of retail rates $y_{b,t+h}$ to a monetary shock Δi_t when the shock takes place at a 100 bps higher interest rate i_{t-1}. I find that deposit and loan pass-through is lower when the interest rate is lower. Figure 19 displays the results for two types of loans ("ARM 1 year" is the fixed rate for the first year on adjustable-rate mortgages; "3-year auto loans " is the rate on auto loans for new vehicles; and two types of deposits, checking and savings deposits. The estimates $\left\{\hat{\beta}_h\right\}_{h=0,\ldots,12}$ are above zero in all cases, which means that the pass-through of monetary shocks to retail rates is higher at higher rates, for both loans and deposits.[16] These results suggest that the transmission of monetary policy to the rates faced by firms and consumers is weakened in a low rate environment. In the rest of the paper, I will use the model to provide an explanation for this pattern, and to draw implications for the real effects of monetary policy.

Asymmetric pass-through? In Appendix A.2, I show that the pass-through of policy rates to retail rates is asymmetric (and differently for loan and deposit rates, as we would expect) but that results are not driven by asymmetric pass-through, because interest rate hikes are not more likely at low rates in my sample.

[15] Tenreyro and Thwaites (2016) find that US monetary policy is less powerful in recessions.

[16] Clustered standard errors account for autocorrelation in error terms within branches, following Jordá (2005); Appendix A.2 shows the results with two-way (branch-month) clustering.

Figure 3: Additional pass-through (in percentage points) of a monetary shock to retail rates when the 1-year rate is 100 bps higher.



Note: The regression equations are $y_{b,t+h} - y_{b,t-1} = \alpha_{b,h} + \delta_{1h}\Delta i_t + \delta_{2h} i_{t-1} + \beta_h \Delta i_t \times i_{t-1} + \gamma_h \text{controls}_{b,t-1} + \epsilon_{b,t+h}$ for each horizon h. The figures show the sequences $\{\hat{\beta}_h\}_{h=0,\ldots,12}$ with 90% confidence bands. Standard errors are clustered at the branch level. Appendix A.2 shows the results with two-way clustering. Sources: Federal Reserve and RateWatch.

Raw changes in interest rates. In my baseline specification (1), I use a monthly version of Nakamura and Steinsson (2018)'s monetary shocks as independent variable Δi_t to isolate the causal effect of monetary policy on retail rates. It is common in standard pass-through regressions (for instance in the exchange rate pass-through literature, surveyed by Burstein and Gopinath 2014) to use raw changes in policy variables instead of shocks identified using high-frequency asset prices, which can inform about raw correlations instead of "causal effects". In Appendix A.2, I show that the same results hold when using monthly changes in the 1-year rate $\Delta i_t = i_t - i_{t-1}$ instead. This also allows to extend the sample to 2018 (instead of 2014, when the series for monetary shocks stops) and thus include the years 2015-2018 that saw a rise in policy rates with limited pass-through to retail rates.

Pre-2007. The results are not driven by the zero lower bound period. In Appendix A.2, I show that the same results hold when truncating the sample in 2007.

3 A model of banks, credit and liquidity

I now present a model of financial intermediation between heterogeneous agents that can explain the patterns described in Section 2. The key idea of the model is that liquidity premia and credit spreads are entangled through the balance sheets of commercial banks.

Overview of the model. There are three types of agents. Banks intermediate funds between two types of households, "borrowers" and "savers". On the asset side, banks can hold bonds or finance loans of arbitrary maturity to borrowers, while on the liability side they can issue bonds or short-term deposits. Households can save in bonds, deposits or cash. Relative to bonds, cash and deposits provide liquidity services. Monetary policy sets the nominal interest rate on bonds, while the rates on loans and deposits adjust endogenously. Spreads between loans and bonds and bonds and deposits can persist because banks are subject to financial constraints.

3.1 Environment

Time is discrete, starting at $t = 0$.

Firms and technology. Competitive firms produce the final good Y from labor N with a linear technology A:

$$Y_t = A_t N_t.$$

Assets. The available assets are short-term bonds with face value a_t, money m_t, short-term bank deposits d_t, and bank loans $l_{t,t+k}$ with maturity $k \in \{1, 2, \ldots, K\}$. Let R_t, R_t^d and $R_{t,t+k}^l$ be the respective real gross returns on bonds, deposits and loans with maturity k. The corresponding (real) asset prices at date t are denoted $q_t = \frac{1}{R_t}$, $q_t^d = \frac{1}{R_t^d}$ and $q_{t,t+k}^l = \frac{1}{R_{t,t+k}^l}$. For the special case of short-term loans I simply write $R_t^l = R_{t,t+1}^l$. The real return on money R_t^m is the inverse of inflation, $R_t^m = \frac{P_t}{P_{t+1}}$.[17] Net inflation is $\pi_{t+1} = \frac{P_{t+1}}{P_t} - 1$. The net nominal rate on bonds is $i_t = \frac{R_t}{R_t^m} - 1$. For consistency, I express all rates in real terms.[18]

There is no risk, but assets can be imperfect substitutes for two reasons. First, different assets are associated with different borrowing constraints. For instance, bank-dependent borrowers are able to short loans but not bonds; and households can place their savings in bonds and deposits but not in loans directly—only banks have the expertise to manage loans. Second, money and deposits are not only valued for their pecuniary returns, but also for the transaction services they provide.

Households. Households come in two types that differ in their preferences, in the pattern of their labor endowments and in their financial constraints. "Savers" are unconstrained households that can be viewed as also incorporating all the borrowers in the economy who do not depend on banks. "Borrowers" are *defined* as the bank-dependent agents.

Savers. There is a mass $\mu^s = 1$ of infinitely-lived savers. Each saver is endowed with $\overline{n}^s$ unit of labor in each period, and supplies labor inelastically. Savers have a discount factor β, and value consumption and liquidity services from money m and deposits d. They solve

$$\max_{c_t, a_{t+1}, m_{t+1}, d_{t+1}} \sum_{t=0}^{\infty} \beta^t \left[u\left(c_t\right) + v\left(x\left(m_t, d_t\right)\right) \right]$$

$$\text{s.t.} \quad c_t + \frac{a_{t+1}}{R_t} + \frac{m_{t+1}}{R_t^m} + \frac{d_{t+1}}{R_t^d} \leq w_t \overline{n}^s + a_t + m_t + d_t + \text{Div}_t + T_t^s.$$

Every budget constraint will be expressed in real terms. w_t is the real wage, Div_t are aggregate net bank dividends (see below) and T_t^s are lump-sum transfers from the government. Under flexible prices, non-financial firms make no profits; in Section 5, I will describe the distribution of the profits that arise under sticky prices.

A central ingredient of the model is the demand for public liquidity (money m) and private liquidity (deposits d) that arises from the aggregator x.

[17] Buying m_{t+1} units of real money maturing at date $t+1$ at date-t price q_t^m costs a nominal amount $M_t \equiv P_t q_t^m m_{t+1}$ at date t thus $\frac{M_t}{P_t} = q_t^m \frac{M_t}{P_{t+1}}$.

[18] Whether financial assets are real or nominal is irrelevant until Sections 5.2 and 6 where I consider unanticipated monetary shocks.

Assumption 1 (Liquidity). *The aggregator $x(m, d)$ is increasing, homothetic, differentiable, and concave.*

Borrowers. To generate loan demands while limiting the number of state variables, I assume that borrowing from banks is entirely driven by household lifecycle motives.[19] To capture different loan maturities, I use a "preferred habitat" framework: there are overlapping generations of borrowers, who are heterogeneous with respect to the maturity of the loans they need. At each date t, a mass μ_k of borrowers, indexed by their life span $k \in \{1, \ldots, K\}$, is born. Each borrower of type k lives only at two dates, t and $t+k$, and is endowed with $\overline{n}^{y,k}$ units of labor when young, and $\overline{n}^{o,k}$ units when old. Borrowers of type k born at date t have utility

$$u\left(c_t^{y,k}\right) + \beta^k u\left(c_{t+k}^{o,k}\right).$$

The following financial friction gives a role to banks' asset side:

Assumption 2 (Credit frictions). *Borrowers cannot short bonds and must borrow through loans.[20]*

By selling loans l_{t+k} due when old at $t+k$, borrowers receive an amount $\frac{l_{t+k}}{R^l_{t,t+k}} \geq 0$ when young.[21] They solve

$$\begin{aligned}
\max_{c_t^{y,k}, c_{t+k}^{o,k}, l_{t,t+k}} \quad & u\left(c_t^{y,k}\right) + \beta^k u\left(c_{t+1}^{o,k}\right) \\
\text{s.t.} \quad & c_t^{y,k} \leq w_t \overline{n}^{y,k} + \frac{l_{t+k}}{R^l_{t,t+k}} \\
& c_{t+k}^{o,k} \leq w_{t+k} \overline{n}^{o,b} - l_{t+k} \\
& l_{t+k} \geq 0.
\end{aligned}$$

The last constraint states that borrowers can only borrow, and not lend, through loans—only banks can lend. Borrowers could save by buying bonds, but they will not in equilibrium, so in what follows I ignore this possibility to ease notation. I also shut down any fiscal transfer to borrowers to avoid having the government play the role of a financial intermediary able to alleviate constraints on the flow of funds between private agents.

[19] I consider investment in Appendix D.2.

[20] Assumption 2 could be relaxed by allowing bond issuance subject to a borrowing constraint. Once the bond constraint binds, marginal borrowing is in the form of loans so nothing changes.

[21] For clarity I switch sign convention for l between banks and borrowers.

The total endowment of labor is constant equal to 1:

$$\sum_{k=1}^{K} \mu_k \left(\overline{n}^{y,k} + \overline{n}^{o,k}\right) + \mu^s \overline{n}^s = 1.$$

Banks. I follow closely the workhorse model of banks developed by Gertler and Kiyotaki (2010) with two important differences. First, here banks are not only specialists in credit provision, but also play a role through their deposit claims which are valued for their liquidity. Second, I allow for the long-term loans described above, as maturity mismatch is central to understand the impact of monetary policy on banks. Long-term loans will play no particular role until Section 6, but I introduce them now to provide a complete model.

There is a unit mass of banks. Banks are owned by savers but operated by bankers. The vectors $\mathbf{l}^t$ and $\mathbf{q}_t^l$ represent respectively the loan portfolio at the beginning of date-t and the price of loans maturing at $t, t+1, \ldots, t+K$:

$$\mathbf{l}^t \equiv \{l_{t+k}\}_{k=0}^{K-1}, \quad \mathbf{q}_t^l \equiv \left\{q_{t,t+k}^l\right\}_{k=0}^{K}$$

with $q_{t,t}^l \equiv 1$. The equity, or capital, of a bank $i \in [0,1]$ with a portfolio $\left[a_t(i), d_t(i), \mathbf{l}^t(i)\right]$ at the beginning of period t is

$$e_t(i) \equiv \mathbf{q}_t^l \cdot \mathbf{l}^t(i) + a_t(i) - d_t(i) \tag{2}$$

e_t is the *marked-to-market book value* of equity (and not the market value, in the sense that it does not capitalize future profits); note that the relevant asset prices to discount future loan payoffs and hence book equity are loan prices, inclusive of potential future spreads. Aggregate bank equity at the beginning of period t is

$$E_t = \int_0^1 e_t(i)\, di.$$

The following assumption determines the process for bank dividends:

Assumption 3 (Entry and exit). *In each period, a mass ρ of banks exits and a mass of banks ρ enters, each of them with exogenous startup equity $\zeta_t E_t/\rho$. Each exiting bank sells its loan portfolio to remaining banks and then rebates its equity to the representative saver. Net payouts are high enough for banking to be relevant in steady state:*

$$\rho - \zeta > 1 - \beta.$$

Thus aggregate net dividends are

$$\text{Div}_t = (\rho - \zeta_t) E_t.$$

In equilibrium, banks must have long positions in loans ($l_{t+k} \geq 0$) and short positions in deposits ($d_{t+1} \geq 0$). In general, banks could be long or short in bonds; but given the non-satiation of deposit services, in equilibrium R_t^d is strictly lower than R_t, thus bonds are dominated by deposits on the liability side and banks' bond holdings a_{t+1} will always be non-negative. Therefore the leverage ϕ_t, defined as the ratio of liabilities over equity, is

$$\phi_t = \frac{q_t^d d_{t+1}}{e_t}.$$

Active banks take as given the discount factor $q_t = \frac{\beta u'\left(c_{t+1}^s\right)}{u'\left(c_t^s\right)}$ and maximize expected dividends, solving

$$\begin{aligned}
V_t(e_t) \equiv \max_{a_{t+1}, d_{t+1}, \mathbf{l}^{t+1}} \quad & q_t \left\{ \rho e_{t+1} + (1-\rho) V_{t+1}(e_{t+1}) \right\} \\
\text{s.t.} \quad & \mathbf{q}_t^l \cdot \mathbf{l}^{t+1} + q_t a_{t+1} = e_t + q_t^d d_{t+1} \\
& e_{t+1} = \mathbf{q}_{t+1}^l \cdot \mathbf{l}^{t+1} + a_{t+1} - d_{t+1} \\
& \phi_t \leq \overline{\phi}_t \\
& a_{t+1} \geq 0,
\end{aligned}$$

where e_t is given by (2) for incumbent banks and $e_t = \zeta_t E_t / \rho$ for new banks. Unlike households that can neither take short positions in bonds and deposits nor long positions in loans, banks can take any position subject to a cap $\overline{\phi}_t$ on their leverage. I will assume that the maximal leverage ratio $\overline{\phi}_t$ stems from a limited pledgeability constraint, either due to a moral hazard problem or a perceived risk of run:

Assumption 4 (Limited pledgeability)**.** *At date t, banks can only pledge a fraction $\theta \in [0,1]$ of date-$t+1$ assets to cover their liabilities.*[22]

In the banks' program, loans of different maturities are perfectly substitutable, which implies that the expectation hypothesis holds for loan rates: $R_{t,t+k}^l = \prod_{j=0}^{k-1} R_{t+j}^l$.[23] Since, in addition, deposits are the only liability, the pledgeability constraint gives us an expression

[22]The same pledgeability parameter θ applies to all assets; one could assume that bonds are more pledgeable than loans, but this would only introduce an additional wedge between bonds and loans without changing any result below.

[23]Equivalently, asset prices satisfy $\mathbf{q}_t^l = q_{t,t+1}^l \mathbf{q}_{t+1}^l$.

for the maximal leverage ratio where only the *short-term* loan rate appears:

$$\overline{\phi}_t = \frac{\theta R_t^l / R_t^d}{1 - \theta R_t^l / R_t^d}.$$

Thus a higher spread between the short-term loan rate R_t^l and the deposit rate R_t^d relaxes the bank leverage constraint at t by making more interest income pledgeable to cover less interest expense at $t+1$. This captures a positive dependence of bank lending capacity in current profits. I discuss banks' constraints in theory and in practice in Section 3.5.

Banks' excess returns. The dynamics of bank capital are governed by the return on bank equity ROE_t, defined as

$$\text{ROE}_t = \frac{E_{t+1}}{E_t - \text{Div}_t}. \tag{3}$$

If there are excess returns $\text{ROE}_t - R_t \geq 0$, as will be the case in equilibrium, it is optimal for banks to delay dividends until exit. We can reexpress banks' budget constraints using the expectation hypothesis for loan rates to obtain the key equation:

$$\text{ROE}_t - R_t = \underbrace{\overline{\phi}_t \left(R_t - R_t^d\right)}_{\substack{\text{excess return from} \\ \text{liquidity provision}}} + \underbrace{\left(1 + \overline{\phi}_t\right)\left(R_t^l - R_t\right)}_{\substack{\text{excess return from} \\ \text{credit provision}}}. \tag{4}$$

The excess return on equity is the sum of two terms, reflecting the two distinct intermediation activities performed by banks. On the one hand, the excess return from deposit liquidity creation is equal to the spread $R_t - R_t^d$ leveraged by a factor ϕ_t; on the other hand, banks can earn an excess return from the spread $R_t^l - R_t$ on loans, leveraged by a factor $1 + \phi_t$.[24] I define loan and deposit spreads as follows, with the convention that both spreads are non-negative:

Definition 1. *At date t, the loan spread is*

$$\tau_t^l = \frac{R_t^l - R_t}{R_t},$$

and the deposit spread is

$$\tau_t^d = \frac{R_t - R_t^d}{R_t}.$$

If savers could freely arbitrage between bonds and bank equity, they would demand more equity as long as there are excess returns, bringing down ROE_t to R_t in equilibrium. In my

[24] If the spread $R_t^l - R_t$ is positive, then banks hold no bonds on the asset side and $\frac{\mathrm{q}_t^l \cdot \mathrm{l}_{t+1}^{t+1}}{e_t} = 1 + \overline{\phi}_t$; otherwise, if $\frac{\mathrm{q}_t^l \cdot \mathrm{l}_{t+1}^{t+1}}{e_t} < 1 + \overline{\phi}_t$, then the spread is zero and the expression still holds.

baseline model, net dividends are exogenous from Assumption 3, so the return on equity can dominate the real interest rate. Section 4.4 will consider the intermediate case of costly equity issuance.

Monetary and fiscal policy. The central bank implements uniquely (for instance through a Taylor rule) a sequence of nominal rates $\{i_t\}_{t\geq 0}$. I begin with a traditional implementation based on household money demand: given an initial price level $P_0 > 0$ we can back out the implied sequence of money supply $\{M_t\}_{t\geq 0}$.

The seigniorage revenue from outside money creation is rebated lump-sum, in the same period, to savers, who are the ones who pay for it. This ensures that monetary policy does not imply a mechanical redistribution from savers to borrowers. As a result of the transfer rules and the fact that only savers and borrowers hold bonds in equilibrium, Ricardian equivalence holds regarding the timing of transfers $\{T_t^s\}$, and I assume without loss of generality that the government runs a balanced budget

$$T_t^s = (1 + \pi_{t+1})\, m_{t+1} - m_t, \tag{5}$$

where m_t are the equilibrium real money balances and π_{t+1} is the net inflation rate from t to $t+1$. Therefore bonds are in zero net supply.

3.2 Equilibrium

I start with a standard equilibrium concept that assumes flexible prices and full employment. The flexible prices equilibrium is suitable for studying long-run issues; I will introduce nominal rigidities in Section 5 to address short-run issues and highlight where the financial frictions interact with or alter the traditional New Keynesian channel of monetary policy transmission.

Let $\mathbf{L}^t = \{L_{t+k}^t\}_{k=0}^{K-1}$ be the aggregate stock of loans outstanding at the beginning of period t. L_{t+s}^t is the sum of individual loan positions l_{t+k} over banks that are active at $t-1$. The economy has $K+2$ aggregate state variables, summarized in the vector $\mathbf{Z}_t = [a_t^s, D_t, \mathbf{L}^t]$.

Definition 2 (Flexible prices equilibrium). *Given initial conditions $\mathbf{Z}_0$, a path for monetary policy $\{R_t^m\}_{t\geq 0}$ and an initial price level $P_0 > 0$, a flexible prices equilibrium is a sequence of allocations $\left\{\{c_t^i\}_i, \mathbf{Z}_t, m_t\right\}_{t\geq 0}$, real wages $\{w_t\}_{t\geq 0}$ and rates $\left\{R_t, R_t^d, \{R_{t,t+k}^l\}_{k=1}^K\right\}_{t\geq 0}$ such that firms, households and banks optimize, and markets for goods and all assets clear.*

The equilibrium is fully characterized by equations (30) to (44) in Appendix B. From now on, I assume

$$u(c) = \log c$$
$$v(x) = \chi \log x$$

to simplify expressions. All the results can be obtained with more general CRRA preferences that may differ for borrowers and savers, as long as loan demand curves slope down.[25] I now discuss the two key markets in this model: the market for liquidity, and the market for loans.

3.3 Liquidity side: competition between public and private liquidity

In addition to their intertemporal consumption-savings problem, solved by the standard Euler equation $u'\left(c_t^s\right) = \beta R_t u'\left(c_{t+1}^s\right)$, savers face a static optimal liquidity demand problem. The first-order conditions with respect to money and deposits are respectively

$$\frac{v'\left(x\left(m_{t+1}, d_{t+1}\right)\right) \cdot x_d\left(m_{t+1}, d_{t+1}\right)}{u'\left(c_{t+1}^s\right)} = \frac{q_t^d - q_t}{q_t} \equiv s_t^d. \tag{6}$$

$$\frac{v'\left(x\left(m_{t+1}, d_{t+1}\right)\right) \cdot x_m\left(m_{t+1}, d_{t+1}\right)}{u'\left(c_{t+1}^s\right)} = \frac{q_t^m - q_t}{q_t} = i_t. \tag{7}$$

s_t^d is the opportunity cost of deposits, or equivalently the "price of private liquidity"; the nominal interest rate i_t is the price of public liquidity. Given consumption c_{t+1}^s, equations (6) and (7) determine the optimal date-t liquidity holdings d_{t+1}, m_{t+1} as functions of the prices i_t and s_t^d. We can divide (6) by (7) to obtain

$$\frac{x_d\left(m_{t+1}, d_{t+1}\right)}{x_m\left(m_{t+1}, d_{t+1}\right)} = \frac{s_t^d}{i_t},$$

which, by homogeneity of degree 0, yields a function for the money-deposit ratio

$$\frac{m_{t+1}}{d_{t+1}} = f\left(\frac{s_t^d}{i_t}\right)$$

[25] Calling σ_b borrowers' EIS, the condition is $\frac{1/\sigma_b - 1}{G^{k/\sigma}} \leq \frac{\gamma^{y,k}}{\beta_b^k \gamma^{o,k}}$ for all $k \geq 1$. A sufficient condition is $\sigma_b \geq 1$.

f is increasing by the concavity of the aggregator x. We can then replace the money-deposit ratio in (6) to obtain the optimal aggregate deposit demand conditional on c_{t+1}^s:

$$\mathcal{D}_{t+1}\left(s_t^d, i_t, c_{t+1}^s\right) = \frac{x_d\left(f\left(\frac{s_t^d}{i_t}\right), 1\right)}{s_t^d x\left(f\left(\frac{s_t^d}{i_t}\right), 1\right)} \chi c_{t+1}^s \tag{8}$$

In this paper, I follow and generalize Drechsler et al. (2017) in using substitutability between money and deposits to generate incomplete pass-through of monetary policy to deposit rates.[26] From now on, I assume that the two forms of liquidity are sufficiently close substitutes:

Assumption 5 (Gross substitutes)**.** *Money and deposits are gross substitutes, in the sense that*

$$\frac{x_d(f,1)}{x(f,1)} \text{ is non-increasing in } f.$$

Under Assumption 5, deposit demand $\mathcal{D}$ is not only decreasing in the price of deposit liquidity s_t^d, but also increasing in the price of the competing liquidity, which is the nominal interest rate i_t.[27] Holding deposit supply fixed, a lower nominal rate thus leads to a decline in s_t^d and hence a decline in the deposit spread τ_t^d that governs banks' excess return in (4). In Section 4, I will show the implications of this competition once we endogenize the supply of private liquidity (through the dynamics of bank capital) and combine it with equilibrium in loan markets.

Functional forms. The most common specification for x is a CES aggregator:

$$x(m,d) = \left[\alpha m^{\frac{\epsilon-1}{\epsilon}} + (1-\alpha) d^{\frac{\epsilon-1}{\epsilon}}\right]^{\frac{\epsilon}{\epsilon-1}}. \tag{9}$$

Following Chetty (1969), who first introduced the CES aggregator (9), a large literature has estimated high elasticities of substitution between money and "near-monies" such as deposits. For instance, Poterba and Rotemberg (1987) estimate an elasticity of substitution between currency and savings deposits that lies between 1.2 and 100, depending on the set of instru-

[26] Appendix D.1 contains a detailed comparison between a model with bank market power over deposits, a la Drechsler et al. (2017), and a model with financially constrained banks, as in this paper and Di Tella and Kurlat (2017).

[27] I have expressed the condition when $v(x) = \chi \log x$. In the more general case $v(x) = \chi \frac{x^{1-\frac{1}{\omega}}}{1-\frac{1}{\omega}}$, money and deposits are gross substitutes if $\frac{x_d(f,1)}{x(f,1)^{1/\omega}}$ is non-increasing in f.

ments.[28]

Another appealing specification that I will consider is the *hierarchical CES* aggregator:

$$x(m,d) = \left[\alpha m^{\frac{\epsilon-1}{\epsilon}} + (1-\alpha)(m+d)^{\frac{\epsilon-1}{\epsilon}}\right]^{\frac{\epsilon}{\epsilon-1}}. \tag{10}$$

(10) ensures that nominal deposit rates are always non-negative, because money m is a strictly better form of liquidity than deposits d. The interpretation of (10) is that some transactions can only be made in cash (for instance to evade taxation) while for others, cash and deposits are perfect substitutes.[29]

Lemma 1. *If x is CES as in (9), then Assumption 5 is equivalent to*

$$\epsilon \geq 1. \tag{11}$$

If x is hierarchical CES as in (10), then a sufficient condition for Assumption 5 to hold is

$$\epsilon \geq \alpha. \tag{12}$$

Condition (12) can be much weaker than (11), because the share α of pure cash transactions can be very small. Intuitively, the gross substitutes assumption becomes very natural once we account for the fact that money and deposits are not just two arbitrary goods.

3.4 Credit side: two regimes of bank lending

Fixing future policy, the equilibrium is a function of i_t and state variables $\mathbf{Z}_t$. There are two regimes, depending on whether equity is high enough to satisfy loan demand. Aggregate equity, after date-t dividends have been paid out and new banks have entered, is

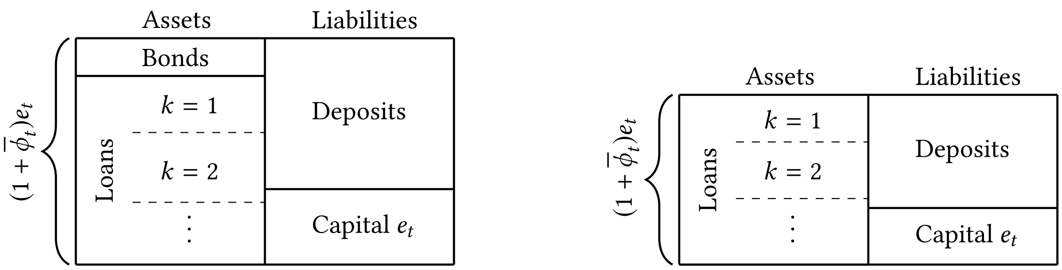
$$L_t^t - D_t - \text{Div}_t + \sum_{k=1}^{K-1} q_{t,t+k}^l L_{t+k}^t.$$

$L_t^t - D_t - \text{Div}_t$ is banks' cash-on-hand. It is predetermined, and if all loans are short-term ($K = 1$) then this is the only component of capital. If $K > 1$ then capital also comprises the market value of the loan book $\sum_{k=1}^{K-1} q_{t,t+k}^l L_{t+k}^t$, which is a forward-looking term. Combining banks' budget and leverage constraints, we have that new lending at t is bounded above by

[28] See Ball (2012), Lucas and Nicolini (2015), and Ireland (2015) for more recent discussions. Cysne and Turchick (2010) provide a survey of post-Volcker estimates of ϵ, with most of them well above 1. One central question in this literature, since at least Barnett (1980), is actually whether money and deposits should be treated as perfect substitutes.

[29] In Appendix B.1, I show how (10) can arise from cash-in-advance constraints à la Lucas and Stokey (1987).

Figure 4: Bank balance sheet in the two regimes



(a) Unconstrained lending **(b)** Constrained lending

banks' lending capacity Λ_t

$$\underbrace{\sum_{k=1}^{K} q^l_{t,t+k}\left(L^{t+1}_{t+k} - L^{t}_{t+k}\right)}_{\text{new loans}} \leq \Lambda_t \equiv \left(1+\overline{\phi}_t\right)\left(L^t_t - D_t - \mathrm{Div}_t\right) + \overline{\phi}_t \sum_{k=1}^{K-1} q^l_{t,t+k} L^t_{t+k}. \tag{13}$$

At any date t the economy can be in one of two regimes. In the *unconstrained lending* regime, the demand for new loans is lower than banks' lending capacity at date t. Bonds and loans must then be perfect substitutes from the banks' viewpoint, and there is no credit spread, i.e., $R^l_t = R_t$. In the *constrained lending* regime, the inequality binds. Date-t credit demand would exceed banks' lending capacity if the ongoing loan rate were R_t, thus a spread has to open up to clear the credit market, i.e., $R^l_t > R_t$. Bank balance sheets in the two regimes are depicted in Figure 4. All else equal, lower equity will shrink the size of banks' balance sheets and push the economy into the constrained lending regime, which is key to give a macroeconomic role to banks. Indeed, absent binding credit frictions in equilibrium, a classical dichotomy holds:

Proposition 1 (Classical dichotomy)**.** *Suppose that in equilibrium bank lending is unconstrained at all times. Then equilibrium consumptions are the same as in a model without liquidity frictions.*

When bonds and loans are perfect substitutes, the imperfect substitutability between bonds and deposits is irrelevant for consumption allocations. This dichotomy results from two facts: first, liquidity services $v(x)$ are separable from consumption utility $u(c)$. Second, both public and private seigniorage revenues are rebated lump-sum in proportion to their usage. The private seigniorage from deposits ends up being rebated to savers through bank dividends. It is well known that non-separable liquidity services or redistribution of seigniorage can generate non-neutrality in the long run. I show below that in the constrained lending regime, non-(super)neutrality arises endogenously through the banking sector's dual role as credit and liquidity provider.

3.5 Discussion of the main assumptions

Household demand for liquidity. Directly putting liquidity services in the utility is the simplest way to generate a demand for assets with dominated return; another route would be to explicitly model transaction frictions, but many such models (e.g., cash-in-advance constraints) are isomorphic to assuming liquidity in the utility (Feenstra 1986). I follow Chetty (1969), Poterba and Rotemberg (1987), and more recently Nagel (2016) and Di Tella and Kurlat (2017) who all incorporate two substitutable forms of liquidity through an aggregator x. In the background, x might stand for inattentive depositors or search costs (see Driscoll and Judson 2013, Yankov 2014, Drechsler et al. 2017). In Appendix B.1, I provide a cash-credit microfoundation for x as in Lucas and Stokey (1987).

Bank leverage constraint. Limited pledgeability à la Holmström and Tirole (1997) is just one of many possible microfoundations for banks' leverage constraint; but different microfoundations will only differ in the exact form of the cap $\overline{\phi}_t$ on leverage. For instance, a limited commitment constraint as in Gertler and Kiyotaki (2010) would impose that bankers not wish to run away with a fraction θ of the value of assets at t

$$V_t \geq \theta \left[q_t a_{t+1} + \mathbf{q}_t^l \cdot \mathbf{l}^{t+1} \right].$$

This constraint would lead to $\overline{\phi}_t = \frac{v_t}{\theta} - 1$ where $v_t \equiv \frac{V_t}{e_t}$ is the market-to-book ratio, that capitalizes not only current profits, as in the constraint I use, but also the whole stream of future profits. At the other extreme, we could also assume a fixed regulatory leverage ratio $\overline{\phi}$ that does not depend on profitability.

There are two interpretations of banks' leverage constraint. It can be viewed as directly imposed by informed depositors, for instance through the threat of runs, as in Diamond and Rajan (2002), or, in the "representation hypothesis" of Dewatripont and Tirole (1994), by bank regulators representing depositors unable to exert the contingent control rights of debtholders in non-financial firms. In practice, banks face a wide range of regulatory capital requirements: Greenwood, Stein, Hanson and Sunderam (2017) list at least ten constraints, with different constraints binding for different banks. Some of those constraints, most clearly the capital charges implied by the Comprehensive Capital Analysis and Review (CCAR) and Dodd-Frank Act stress tests, are directly relaxed by a higher net interest margin through higher "pre-provision net revenue". Finally, capital requirements are based on a mix of market equity values and book equity values (both historical cost book equity and mark-to-market book equity), which means that bank lending capacity indeed depends on the value of long-term assets; see Fuster and Vickery (2018) for a recent discussion.

Firm borrowing. Financial flows only take place between households. Another possibility would be to have constrained firms borrowing from households in order to invest. I take the household route for several reasons. It allows me to abstract from the dynamics of *physical* capital, in line with the basic New Keynesian model (Woodford, 2003), and maintain an exogenous natural output. Moreover, there is a growing literature on the importance of household borrowing (Mian, Rao and Sufi, 2013); and while small firms are mostly bank-dependent borrowers, large firms have access to liquid corporate bond markets, but no household can issue securities. In Appendix D.2, I consider a variant of the model where bank loans finance firm investment.

Bank balance sheets. Banks' balance sheets are highly simplified relative to reality: on the liability side, all funding (in particular the marginal funding) is through deposits and there is no wholesale (bond) funding; on the asset side, banks hold no bonds in the constrained lending regime. From the perspective of bank profitability, the balance sheets displayed in Figure 4 can be viewed as netting out securities held on the asset side and wholesale funding on the liability side, as both would pay the same bond rate and earn no excess return. These simplifications allow me to zoom in on the synergies between the lending and deposit-taking businesses. In Appendix D.3 I add excess reserves to banks' asset side, which allows me to study the interaction between conventional and unconventional monetary policy.

4 Long run: permanent interest rate changes

In this section, I solve for stationary flexible price equilibria and derive comparative statics with respect to long-run trends in interest rates. The main results are that a decline in the equilibrium real rate r^* compresses deposit spreads but widens loan spreads, while an increase in the inflation target has the opposite effect. Holding inflation fixed, a lower r^* can thus explain the incomplete long-run pass-through I documented in Section 2.1.

4.1 Steady state

I suppose that productivity (hence output) grows at a constant gross rate G and consider steady states (also known as balanced growth paths):

Definition 3. *A steady state is an equilibrium where real quantities divided by output Y_t and asset prices are constant.*

Quantities without time subscripts are normalized by Y_t, i.e., $x \equiv \frac{x_t}{Y_t}$. I only consider steady states with positive bank equity because there is always an (unstable) "financial autarky" steady state with $E = 0$, just like the zero physical capital steady state of the Solow growth model. Appendix B.3 contains all the steady state equations, and I describe here the most important ones. Savers' Euler equation pins down the steady state interest rate

$$R^* = \frac{G}{\beta}.$$

The main steady state equation is then the stationary version of (4), which we can rewrite as

$$\left(1 + \overline{\phi}\right) \tau^l = \frac{\text{ROE}}{R^*} - 1 - \overline{\phi}\tau^d. \tag{14}$$

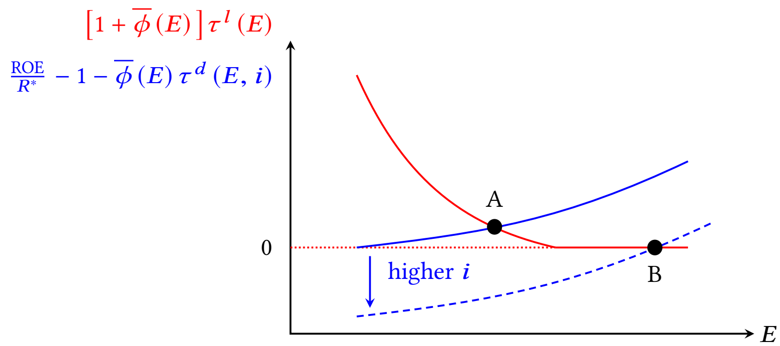
Equation (14) states that the steady state excess return on equity must be sustained by a combination of leverage, loan spreads, and deposit spreads. The terms $\overline{\phi}$, τ^l and τ^d are all endogenous to the steady state level of bank capital E. In an unconstrained lending steady state, the deposit spread is high enough that no credit spread is needed to attain the required return on equity, so the left-hand side reduces to zero. In a constrained steady state, a positive credit spread must open up.[30] From (3), the steady state return on bank equity is ROE $= \frac{G}{1-\rho+\zeta}$ and therefore

$$\frac{\text{ROE}}{R^*} = \frac{\beta}{1 - \rho + \zeta}.$$

The left-hand side of (14) can be interpreted as a *demand for loanable funds*, coming from the credit side, while the right-hand side is the steady state *supply of loanable funds*, that combines deposit market clearing and banks' long-run break-even condition. All else equal, an increase in E increases the supply of loans. Market clearing for loans then implies a lower the credit spread τ^l. Similarly, an increase in E increases the supply of deposits and and deposit market clearing impies that the equilibrium price of deposits, i.e., the deposit spread τ^d, must be lower as E increases. As a result, the leverage ratio $\overline{\phi} = \frac{\theta\frac{1+\tau^l}{1-\tau^d}}{1-\theta\frac{1+\tau^l}{1-\tau^d}}$ is also a decreasing function of E. Hence the left-hand side of (14) is decreasing in E, and the right-hand side is increasing in E, as depicted in Figure 5. Demand and supply intersect at point A, which thus gives the steady state level of aggregate bank capital E.

[30] Equation (14) is merely an accounting identity that will hold across a wide range of models. Here, there is no risk premium, and the ROE is entirely pinned down by the exogenous exit rate of banks ρ. But in any model with financial constraints binding in steady state, a similar equation must hold.

Figure 5: Demand and supply of loanable funds.



Note: Left-hand side of (14) in red, right-hand side in blue. The flat part of the red curve corresponds to the unconstrained lending regime. All else equal, a higher nominal rate i (due for instance to a higher inflation target) allows banks to earn higher deposit spreads τ^d and hence increases steady state capital, reduces the loan rate. A large enough increase in i can push the economy from a constrained lending steady state (A) to an unconstrained steady state (B).

4.2 The nominal interest rate and the composition of bank income

I now investigate how the steady state level of the nominal bond rate i affects intermediation spreads. The nominal rate i depends on both the real interest rate R^* and the inflation target. Changes in the real interest rate affect many parts of the economy, so I start with the simpler case of inflation.

Inflation. Monetary policy is *superneutral* if steady state consumptions are independent of inflation π. As in any monetary model, monetary policy can directly influence liquidity and its cost in steady state. The question is whether consumption allocations are also affected. In the equilibrium condition (14) that determines E, the right-hand side is decreasing in π (because the deposit spread τ^d is increasing in the nominal rate i, as explained below). Therefore a higher steady state inflation π shifts the supply of credit up in Figure (5), which results in higher equilibrium bank equity E. Whether this has an impact on consumptions depends on whether lending is constrained. From the dichotomy in Proposition 1, we already know that monetary policy is superneutral in an unconstrained lending steady state. In the general case, we have:

Proposition 2 (Non-superneutrality). *Fixing other parameters and policy π, there exists a threshold $\overline{\rho}$ such that lending is constrained in steady state if and only if $\rho > \overline{\rho}$.*

- *In an unconstrained lending steady state, monetary policy is superneutral: a local increase in π increases bank capital E but leaves R^*, R^d and R^l unchanged.*

- *In a constrained lending steady state, monetary policy is superneutral if and only if $x(m,d)$ is a Cobb-Douglas aggregator.*[31]

- *If $x(m,d)$ is* not *Cobb-Douglas (i.e., m and d are strict gross substitutes), then there exists a threshold $\bar{i}(\rho)$ such that lending is constrained if and only if the nominal rate i is strictly lower than $\bar{i}(\rho)$. When $i < \bar{i}(\rho)$, a local increase in the inflation target π increases E and $\bar{\phi}$, lowers R^d, decreases R^l and leaves R^* unchanged.*

To understand Proposition 2, we can view the effect of inflation through the lens of public-private competition in liquidity provision. There are two providers: commercial banks and the government. A higher nominal interest rate increases the price of public liquidity, which reduces the competition faced by private liquidity issuers. The steady state private seigniorage earned by banks is

$$s^d D = \frac{x_d\left(f\left(\frac{s^d}{i}\right),1\right)}{x\left(f\left(\frac{s^d}{i}\right),1\right)}\chi c^s, \tag{15}$$

which (for a given c^s) increases with the nominal interest rate i. While savers' consumption c^s goes down (because, as owners of the banks, they extract smaller intermediation rents from borrowers), the total effect on private seigniorage $s^d D$ is still positive. The higher private seigniorage relaxes banks' constraint by boosting retained earnings and fueling a higher bank capital stock, which ends up benefiting bank-dependent borrowers by lowering the loan spread they face. Thus *expected* inflation ends up redistributing from savers to borrowers. Figure 5 represents an increase in inflation that is large enough to shift the steady state from point A in the constrained region to point B in the unconstrained region.[32]

The effect of inflation relates to the literature on public liquidity in the form of government bonds. Krishnamurthy and Vissing-Jorgensen (2012; 2015) show that a higher public supply of liquidity in the form of U.S. Treasuries reduces the liquidity premium and crowds out private liquidity creation. Nagel (2016) constructs a measure of the liquidity premium of several near-money assets and shows that it is highly correlated with the nominal interest rate, which indicates a high elasticity of substitution between money and near-money. Greenwood, Hanson and Stein (2015) study the social benefits of this crowding-out when private liquidity creation entails negative pecuniary externalities. Here, public liquidity (currrency) also crowds out private liquidity (deposits), but I point out that crowding-out might

[31] That is, there exists $\omega \in [0,1]$ such that $x(m,d) = m^{\omega} d^{1-\omega}$.

[32] Superneutrality holds in the knife-edge Cobb-Douglas case because a change in the price of money i does not affect the total cost of deposits $s^d D$ paid by savers, so bank balance sheets remain unchanged when the inflation target changes.

have a social cost as well, if tighter financial constraints on private liquidity issuers spill over to higher credit spreads.

Decline in R^*. Steady state deposit demand (15) is shifted by the nominal interest rate i, which depends on both the inflation target and the real interest rate. In practice, central banks target a stable level of inflation while the real interest rate R^* adjusts to changes in the fundamentals of the economy, such as productivity growth and savings rates. Unlike a change in inflation, a change in R^* also affects ROE$/R^*$ directly in equation (14), as well as other parts of the economy such as credit demand. If the fall in R^* is due to lower growth G, ROE and R^* adjust proportionately, hence ROE$/R^*$ remains constant. In this case, the shock to R^* only has an effect on the composition of bank excess returns, exactly as when we varied the inflation target. The response to a change in savings rates, however, depends on banks' payout policy. If the net payout rate $\rho-\zeta$ adjusts so as to keep $\frac{\text{ROE}}{R^*}$ constant, a discount factor shock has the same effect as a growth shock.

Proposition 3 (Lower real rate)**.** *Suppose that*

- *either productivity growth G falls permanently,*
- *or the discount factor β and the rate of retained earnings $1-\rho+\zeta$ increase permanently in proportion.*

Then R^* falls, and ROE$/R^*$ is unchanged. In the unconstrained lending regime, τ^d and τ^l are unchanged. In the constrained lending regime, τ^d falls and τ^l rises.

If we assume that the net payout rate $\rho-\zeta$ is constant instead, a fall in R^* due to higher β will have an additional effect on spreads by mechanically raising banks' required excess return on equity ROE$/R^*$. τ^l will increase further, and τ^d will fall by less. The next section shows that when the net payout rate is endogenous, ROE indeed falls in response to a lower R^*.

4.3 Isolating the interactions between loans and deposits

To see why my results depend crucially on the assumption that banks are *both* credit and liquidity providers, suppose that instead of having a single institution performing the two functions within the same balance sheet, there are two kinds of intermediaries: specialists in lending, with superscript l, and specialists in liquidity provision, with superscript d. The first kind of intermediaries could be viewed as non-bank lenders such as mortgage companies, while the second kind are a form of narrow banks or money market funds. Loan providers finance themselves with bonds at the real interest rate R, while deposit-taking intermediaries invest their deposits in bonds but not loans. I maintain Section 3's assumptions on banks' limited pledgeability θ and payout rate ρ; all intermediaries are still owned by savers. Decoupling the two functions of banks leads back to the classical dichotomy of Proposition 1:

Proposition 4. *In the model with two types of banks,*

- *both spreads τ^l and τ^d are higher than with a single bank;*
- *monetary policy is superneutral: steady state spreads $\left(\tau^l, \tau^d\right)$ and consumption allocations are independent of the inflation target.*

The model with two banks leads to two versions of equation (14):

$$\frac{\mathrm{ROE}}{R^*} = 1 + \left(1 + \overline{\phi}^l\right)\tau^l$$
$$\frac{\mathrm{ROE}}{R^*} = 1 + \overline{\phi}^d \tau^d,$$

where $\overline{\phi}^l = \frac{\theta\left(1+\tau^l\right)}{1-\theta\left(1+\tau^l\right)}$ and $\overline{\phi}^d = \frac{\theta/\left(1-\tau^d\right)}{1-\theta/\left(1-\tau^d\right)}$. Lenders l are insulated from monetary policy and the loan spread only reflects their required excess return on equity $\frac{\mathrm{ROE}}{R^*} = \frac{\beta}{1-\rho}$. Since they cannot rely on any private seigniorage earnings to fuel equity accumulation, their equity is lower than the equity of a two-sided bank. A change in the inflation target also leaves the deposit spread τ^d unchanged, because the equity of "narrow banks" d adjusts to completely offset the shift in deposit demand, exactly as in the unconstrained lending regime of Proposition 2.

4.4 Equity issuance and endogenous ROE

I have so far assumed an exogenous net dividend policy governed by the entry and exit dynamics in Assumption 3. I now show that endogenizing banks' equity issuance and required return on equity ROE reinforces the results.

Bank profitability. Let us first take ROE as given, as in the baseline setup. Monetary policy affects banks' book equity and thus market value, even when it is superneutral in terms of consumption. The total market capitalization V of the banking sector is proportional to the book value of equity E. More precisely, let $v_t = \frac{V_t}{E_t - \mathrm{Div}_t}$ be the market-to-book ratio, defined as the net present value of dividends over book equity. The law of motion of v is $v_t = \frac{\mathrm{ROE}_t}{R_t}\left[\rho + (1-\rho)\, v_{t+1}\right]$, thus in steady state

$$v = \frac{\rho \mathrm{ROE}/R^*}{1 - G\,(1-\rho)\,\mathrm{ROE}/R^*} \geq 1. \tag{16}$$

As depicted in Figure 5, in both the constrained and unconstrained lending regimes, book equity E increases in reaction to a higher inflation target. Therefore, the market value of banks V also increases. The intuition is that when inflation is higher, banks earn more private seigniorage for a given level of capital, which increases τ^d and $\overline{\phi}$. But since the required return on equity is invariant to inflation, $\left(1+\overline{\phi}\right)\tau^l + \overline{\phi}\tau^d$ must remain constant. In an unconstrained lending steady state, this can only be achieved by keeping $\overline{\phi}\tau^d$ constant through higher capital. Part of the extra earnings from private seigniorage is consumed as dividends, and the rest is kept as retained earnings, which ends up increasing E, until the point where the levered excess return from liquidity provision falls back to its initial level. In a constrained lending steady state, a similar process takes place, but part of the adjustment falls on the (lower) credit spread.

Equity issuance. Suppose next that in each period entering banks can issue equity e at a convex cost $C(e)$ instead of having an exogenous startup equity. Without loss of generality, only entering banks issue equity. Given the equilibrium market-to-book ratio v, entrants issue e to maximize $ve - C(e)$ which yields an endogenous rate ζ of equity issuance

$$\zeta = \rho \frac{(C')^{-1}(v)}{E}.$$

The only difference with the previous section is that the long-run required return on equity is now endogenous. From (16), ROE solves

$$\frac{\text{ROE}}{R^*} = \frac{\beta}{1-\rho\left[1-\frac{(C)^{-1}\left(\frac{\rho\text{ROE}/R^*}{1-G(1-\rho)\text{ROE}/R^*}\right)}{E}\right]}. \tag{17}$$

Equation (17) generalizes the Modigliani-Miller case where ROE/R^* is constant at 1.

Proposition 5. *Suppose that entrants can issue equity subject to a positive, increasing and convex cost C. Then, in the constrained lending regime, a fall in the real interest rate R^* due to either lower productivity growth G or higher discount factor β lowers the deposit spread τ^d and bank capital E, and raises the loan spread τ^l.*

The required return on equity ROE adjusts with R^* because when R^* falls, new banks issue more equity in response to the higher market-to-book ratio v, bringing down ROE closer to R^*.

4.5 Operating costs

To ease exposition, I have focused on banks' net interest income. To map the model to the data, we need to take into account other components of bank profits, in particular operating costs as measured by "non-interest expense" in the Call Reports. In this section, I argue that accounting for these costs only strengthens my results.

The before-tax return on assets ROA is defined as

$$\text{ROA} = \text{NIM} + \text{non-interest income} - \text{non-interest expense} - \text{loan loss provisions}$$

where all terms are defined as percent of earning assets. Equation (4) only considers the net interest margin, setting the last three terms to zero. Figure 18 in Appendix A.1 shows that loan loss provisions soared during the financial crisis, but then reverted to pre-2008 levels in 2012 and have been stable since. Meanwhile, mostly thanks to lower data-processing costs, non-interest expense has declined from 4% to 3% of earning assets. I can accomodate non-interest expense by introducing a cost κ of operating loans in the model. The return net of operating costs on a short-term loan is then $\frac{R_t^l}{1+\kappa_t}$. There is some ambiguity regarding how to attribute non-interest expense to the deposit and lending businesses, because a large part of operating costs simply comes from the costs of maintaing branches that are useful both for making loans and deposits. One common choice in the literature is an equal split that attributes 50% of reported non-interest expenses to the loan side (e.g., Greenwood et al. 2017 or Begenau and Stafford 2018). For exposition, I attribute all the non-interest expense to loans, but it is straightforward to attribute part of the costs to the deposit side with additional notation (κ^l, κ^d). My argument is unchanged as long as the non-interest expense attributed to loans is increasing in total non-interest expense.

In the presence of operating costs, the accounting identity (4) becomes $\text{ROE}_t = \left(1+\overline{\phi}_t\right) \times \text{ROA}_t$ where

$$\text{ROA}_t = \underbrace{R_t^l - \frac{\overline{\phi}_t}{\left(1+\overline{\phi}_t\right)} R_t^d}_{\text{NIM}} \quad - \quad \underbrace{\frac{R_t^l \kappa_t}{1+\kappa_t}}_{\text{non-interest expense}}.$$

In the unconstrained lending regime, banks' no-arbitrage condition between bonds and loans requires $R_t^l = R_t(1+\kappa_t)$. We then have, in steady state:

Proposition 6. *A permanent decrease in κ is fully passed through to lower loan rates in the unconstrained lending regime:*

$$\frac{d \log R^l}{d \log (1+\kappa)} = 1,$$

and partly passed through in the constrained lending regime:

$$\frac{d \log R^l}{d \log (1+\kappa)} > 0.$$

The two components of banks' "marginal cost" of lending, the real interest rate and non-interest expense, have fallen, yet the loan spread $\frac{R^l_t}{R_t}$ has been stable or increasing. My model can explain this pattern through the combination of a lower real rate R^* that pushes up the loan spread, and an offsetting decrease in operating costs.

5 Short run: monetary shocks

The previous section considers how long-run trends in bond rates are passed through to retail rates. In this section, I show how credit and liquidity frictions alter the transmission of monetary *shocks* in the short run, in the presence of nominal rigidities.[33]

There are similarities and differences between long run and short run, illustrated in Figure 6. I show the impact of a contractionary monetary shock on real rates. On the left, the nominal interest rate increases permanently, through a higher inflation target. We see Proposition 1 at work: the real deposit and loan rates fall permanently, while the real interest rate remains constant at R^*. In the short run, when nominal rigidities prevail, the real bond rate moves with the nominal interest rate shock. Spreads behave as in the long run: the loan spread contracts, and the deposit spread widens. What matters for the aggregate output response, however, is not the spreads, but the bond and loan rates, which both go up. The increase in the loan spread means that the loan rate does not rise as much as the bond rate, and I will now show that this dampens the real effects of monetary policy.

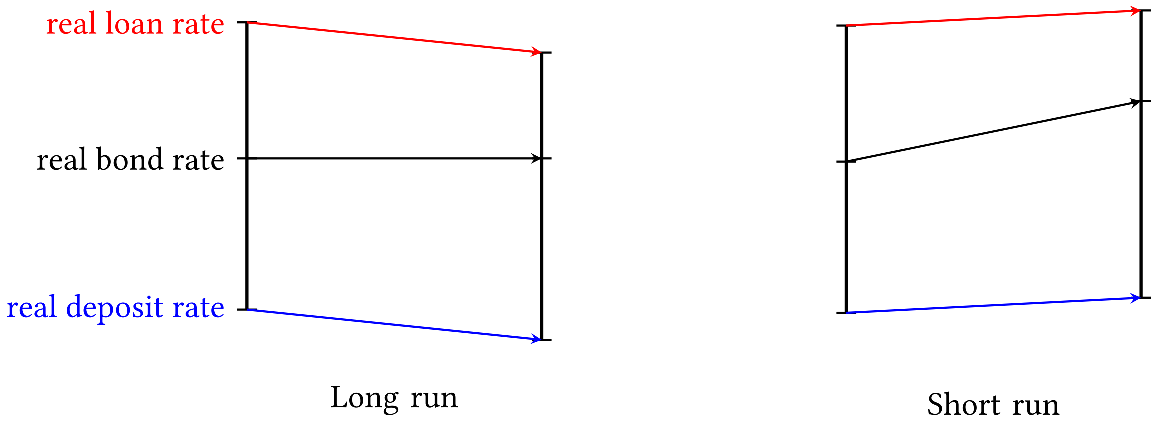
5.1 Nominal rigidities

I model nominal rigidities in a tractable way by fixing inflation π_t at some level $\bar{\pi}$.[34] With fixed inflation, monetary policy effectively controls the sequence of real bond rates $\{R_t\}$ by setting the nominal bond rate. Output Y_t can deviate from its natural level $Y^*_t = A_t$, which would prevail if prices were flexible in all periods. Since inflation is sticky, financial assets

[33]The same forces that yield long-run real effects of monetary policy in Section 4 give rise to short run non-neutrality in spite of flexible prices. However, output is always equal to productivity under flexible prices, so this non-neutrality is just redistributive.

[34]This can be justified by binding downward nominal rigidity as in Schmitt-Grohé and Uribe (2016); alternatively, as explained in Werning (2015), this equilibrium concept can be viewed as characterizing the "demand-side" of the economy.

Figure 6: Reaction to a long run vs. short run increase in the nominal interest rate.



can be considered as real without loss of generality.[35]

With sticky prices or wages, firms make non-zero profits. The distribution of those profits matters for the consumption behavior of households, and thus for aggregate consumption given the heterogeneity and market incompleteness. Following Werning (2015), instead of separating labor income and profits, I define directly the share γ_t^i of aggregate income Y_t that accrues to agent i at date t. Relative to the baseline model, this amounts to replacing labor endowments $\overline{n}_t^i$ with γ_t^i. The equilibrium with nominal rigidities is defined as follows:

Definition 4 (Sticky prices equilibrium)**.** *Given initial conditions $\mathbf{Z}_0$, inflation $\overline{\pi}$, a path for monetary policy $\{R_t\}_{t\geq 0}$ and an initial price level $P_0 > 0$, a sticky prices equilibrium*[36] *is a sequence of allocations $\left\{\left\{c_t^i\right\}, \mathbf{Z}_t, m_t\right\}_{t\geq 0}$ and rates $\left\{R_t^d, \left\{R_{t,t+k}^l\right\}_{k=1}^K\right\}_{t\geq 0}$ such that households and banks optimize, and markets for goods and all assets clear.*

Around a steady state with high interest rates, in the sense that $i \geq \bar{i}(\rho)$ as defined in Proposition 2, we have the sticky prices counterpart of the classical dichotomy result in Proposition 1:

Corollary 1. *Around a steady state with unconstrained lending ($i \geq \bar{i}(\rho)$), the response of consumption allocations and hence aggregate output to a sequence of monetary shocks $\{dR_t\}_{t\geq 0}$ is the same as in a Modigliani-Miller model that features no banks, no credit frictions (i.e., borrowers can freely issue bonds) and no liquidity frictions (i.e., $v = 0$).*

The standard New Keynesian model used to analyze monetary policy abstracts away from

[35] Doepke and Schneider (2006) and Auclert (2017) study in depth the redistribution from surprise inflation in the presence of nominal assets.

[36] I use the more common term "sticky prices" equilibrium instead of "sticky inflation" even though I allow for trend inflation, which plays a role in deposit pricing.

the financial sector and features only a single interest rate, the policy rate R_t controlled by the central bank. Corollary 1 shows that this simplification is justified in a world of high steady state interest rates: when banks earn a sufficiently high private seigniorage, loan rates are equal to bond rates and there is no loss in ignoring what happens within banks to understand the transmission of monetary policy. If, however, the steady state nominal rate falls below the threshold $\bar{i}(\rho)$, we must take into account what happens to banks and spreads.

In my model, the gross substitutability in Assumption 1 remains crucial to understand how relevant banks are. Even when lending is constrained, the following benchmark provides conditions ensuring that all spreads are independent of monetary policy and thus that there is full pass-through to deposit and loan rates. If, in addition, there are no initial revaluation effects, the aggregate output response is then exactly the same as in a representative agent New Keynesian model (RANK):

Proposition 7 (Neutrality)**.** *Suppose that*

(i) the liquidity aggregator $x(m, d)$ is Cobb-Douglas,

(ii) $u(c)$ and $v(x)$ are logarithmic,

(iii) there are no assets maturing at $t = 0$, i.e., $L_0^0 = D_0 = a_0^s = 0$,

(iv) net bank dividends are proportional to output Y_t.

Then the sequences of spreads $\left\{\tau_t^d, \tau_t^l\right\}_{t\geq 0}$ are invariant to monetary policy, and the aggregate output response is the same as in the RANK model, i.e., for all t

$$\frac{d \log Y_t}{d \log R_t} = -1. \tag{18}$$

Proposition 7 builds on the aggregation results developed in Werning (2015) for heterogeneous agents New Keynesian (HANK) models. These models focus on idiosyncratic risk, household borrowing constraints and precautionary savings, while my setting highlights a different form of market incompleteness—the role of banks in intermediation. In my context, the key departure from Proposition 7 is to discard condition (i) on Cobb-Douglas liquidity in order to match the incomplete pass-through of policy rates to deposit rates that I estimate in Section 2.2.

5.2 Incomplete pass-through in a two-period example

I now illustrate the consequences of incomplete pass-through in a two-period version of the model to illustrate analytically the main forces driving the response of output, spreads, bank lending and profits to monetary policy shocks. I will then turn to a quantitative analysis in Section 6.

Setup. There are two dates $t = 0$ and $t = 1$, and prices are sticky only at $t = 0$. Thus Y_1 is fixed at its natural level Y_1^*, but Y_0 depends on monetary policy R_0 or, equivalently, i_0. I consider monetary shocks around an arbitrary baseline allocation determined by the policy rate R_0 (not necessarily the natural rate that implements $Y_0 = Y_0^*$). All households have log-utility ($u = v = \log$) and the same discount factor β. Savers have no income at date 1, so I use γ^s and $\gamma^b = 1 - \gamma^s$ without time subscripts to denote the shares of income Y_0 accruing to savers and borrowers, respectively.

Proposition 8. *There exists a decreasing function $R^*(E_0)$ such that lending is constrained if and only if $R_0 < R^*(E_0)$. Output is given by*

$$Y_0 = \begin{cases} \dfrac{Y_1^*}{\beta R_0} & \text{if } R_0 \geq R^*, & (19)\\[2ex] \dfrac{Y_1^*}{\beta R_0} \times \Gamma\left(\tau_0^l\right) & \text{if } R_0 < R^*, & (20)\end{cases}$$

where $\Gamma \leq 1$ is decreasing in τ_0^l. Fixing R_0, a negative shock to E_0 or θ increases τ_0^l if $R_0 < R_0^*$.

The proof in Appendix C gives the closed form for Γ. In the unconstrained lending regime, output is given by the standard dynamic IS curve (19), that exactly matches the RANK model, as in Proposition 7. Thus liquidity frictions alone are irrelevant for aggregate output, which reflects Proposition 1's dichotomy between consumption and liquidity allocations in the unconstrained lending regime. In the constrained lending regime, the new term Γ captures credit frictions. The loan spread τ_0^l appearing in (20) is an equilibrium object, but a negative shock to bank capital E_0 or pledgeability θ increases τ_0^l all else equal. A higher loan spread τ_0^l hurts borrowers but benefits savers, who receive higher dividends from banks at $t = 1$. The proof of Proposition 8 shows that the total effect on output is negative, and thus an increase in τ_0^l acts as a negative aggregate demand shifter.

The pass-through of monetary shocks. Equation (20) makes apparent that a monetary shock (a change in R_0) has two effects in the constrained regime: the standard RANK effect on the first term $\frac{Y_1^*}{\beta R_0}$, and a new effect on the second term $\Gamma\left(\tau_0^l\right)$ if the shock affects the spread between loan and bond rates. I now describe how loan rates are, in turn, related to deposit rates. Let

$$\eta^Y \equiv -\frac{d \log Y_0}{d \log R_0}, \quad \eta^l \equiv \frac{d \log R_0^l}{d \log R_0}, \quad \eta^d \equiv \frac{d \log R_0^d}{d \log R_0}$$

denote the interest-elasticities, or *pass-through*, of monetary policy to output (with a minus sign to have positive numbers), loan rates and deposit rates, respectively. These three elasticities are tied together by the fact that banks are both the issuers of deposits and the providers of loans:

Proposition 9. *In the constrained lending regime, loan pass-through η^l and the interest-elasticity of output η^Y are both increasing in deposit pass-through η^d. η^Y is positive but smaller than in the unconstrained lending regime (where $\eta^Y = 1$); more precisely, in the constrained lending regime η^Y is bounded by*

$$1 - \frac{\left(1+\overline{\phi}\right)\Gamma}{1+\overline{\phi}+\frac{(1-\epsilon^{\Gamma})\beta\gamma^b Y_0}{Y_1^*/R_0^l-\beta\gamma^b Y_0}} \leq \eta^Y \leq 1 - \frac{\Gamma}{1+\overline{\phi}+\frac{(1-\epsilon^{\Gamma})\beta\gamma^b Y_0}{Y_1^*/R_0^l-\beta\gamma^b Y_0}},$$

where $\epsilon^{\Gamma} = \frac{d\log\left(1+\tau_0^l\left[1-\gamma^s\beta\left(\frac{1+\tau_0^l}{1+\beta+\tau_0^l}\right)\right]\right)}{d\log\left(1+\tau_0^l\right)} \in (0,1)$.

The proof, in Appendix C, contains closed-form formulas mapping η^d to η^l and η^Y. The intuition behind Proposition 9 is as follows. In this two-period setting, equity E_0 is fixed and loan supply only depends on the leverage ratio $\overline{\phi}_0$, which increases in the loan-deposit spread R_0^l/R_0^d because higher profits relax banks' pledgeability constraint.[37] The size of banks' balance sheet $\left(1+\overline{\phi}_0\right)E_0$ ties together the deposit and loan markets. All else equal, a higher deposit pass-through η^d means that an interest rate hike contracts credit supply by more, because the profits banks make on deposits fall by more. The loan rate must then increase by more to clear the credit market, that is, loan pass-through η^l must be higher. Since the fall in credit supply acts as a deleveraging shock, the more credit supply contracts, the more output falls in response to the rate hike. Thus a higher deposit pass-through leads to a stronger effect of monetary shocks on output η^Y through a "bank lending channel" whose strength depends on deposit pass-through.[38] Next, I show how low nominal interest rates can dampen deposit pass-through.

[37] Recall that $\overline{\phi}_0 = \frac{\theta R_0^l/R_0^d}{1-\theta R_0^l/R_0^d}$.

[38] The interest-elasticity of output depends on the deposit rate only indirectly, through the loan rate. Although savers do hold deposits that pay R_0^d and thus earn an *average* return between R_0^d and the bond rate R_0 (and in fact exactly R_0^d in the constrained regime), their relevant *marginal* rate of intertemporal substitution remains R_0, hence there is always full pass-through of monetary policy to savers' consumption. Assuming that savers can only save in deposits would mechanically imply incomplete pass-through even absent credit frictions. However, in my model, allowing for savers' portfolio choice between bonds, deposits and money is crucial to generate incomplete deposit pass-through.

Low interest rates and low pass-through. Proposition 9 takes deposit pass-through η^d as given, but it is itself an equilibrium object, that depends in particular on the level of interest rates. The following result shows how low rates can dampen deposit pass-through:

Lemma 2. *If*

$$\frac{d^2}{df^2}\left[\frac{x_d(f,1)}{x(f,1)}\right] \leq 0, \tag{21}$$

then the deposit pass-through η^d is increasing in the nominal rate i_0. If x is CES as in (9), then a sufficient condition for (21) to hold is

$$\epsilon > \frac{2-2\alpha}{1-2\alpha}.$$

In Section 2.2, I show that deposit pass-through is lower at lower rates in U.S. data. This suggests that condition (21) is satisfied for U.S. deposit rates. Gross substitutability implies that the money-deposit ratio is a decreasing function of the nominal rate i. The intuition behind (21) is that there is more scope for substitution between money and deposits when the money-deposit ratio $f = \frac{m}{d}$ is higher, which happens at lower i. Within my framework, this assumption on x is the simplest way to capture the state-dependence of deposit pass-through, but there could be other reasons for deposit pass-through to depend on the level of interest rates, for instance in models that feature depositor search (Yankov 2014).

Together with Proposition 9, condition (21) implies that the loan pass-through η^l and the interest-elasticity of output η^Y are also lower when the nominal interest rate i_0 is lower. In my model, the fact that loan pass-through is lower at lower rates, as we also see in Section 2.2, is therefore *a consequence of the lower deposit pass-through* at lower rates. Deposit and loan markets are connected because the same intermediaries are providing loans and deposits within the same balance sheet. In a narrow banking system where different intermediaries specialize in lending or liquidity provision, low interest rates would still decrease deposit pass-through, but would have no reason to affect loan pass-through.

5.3 Bank profitability and lending

So far I have studied how banks affect the transmission of monetary policy to households and aggregate output. I conclude this section by focusing on the impact of monetary policy on banks themselves, looking at bank lending and bank profitability. I show in particular how the interest-elasticity of bank profits depends on the level of interest rates, and provide evidence from the response of U.S. bank stock prices to monetary shocks.

One measure of banks' profits is the return on equity

$$\text{ROE} = \frac{E_1}{E_0} = R_0^l \times \frac{1-\theta}{1-\theta\frac{R_0^l}{R_0^d}}. \tag{22}$$

ROE has two components. One is the loan rate R_0^l; the other is the loan volume, which is increasing in leverage hence in the spread R_0^l/R_0^d. In the unconstrained lending regime, bank profits always increase with R_0 as both the loan rate and the loan volume increase. In the constrained lending regime, an increase in R_0 is partially passed through to R_0^l, but it can also decrease the loan-deposit spread and hence the volume of lending.

Proposition 10 (Bank profits and lending)**.** *Suppose that the liquidity aggregator x satisfies condition (21). Then there exist $\underline{\underline{R}}, \underline{R}$ such that $\underline{\underline{R}} \leq \underline{R} \leq R^*$ and*

- *for $R_0 \in \left(\underline{R}, R^*\right)$, bank lending and profits fall with a monetary tightening $dR_0 > 0$;*
- *for $\underline{\underline{R}} \leq R_0 \leq \underline{R}$, bank lending falls but bank profits rise with a monetary tightening;*
- *for $R_0 \leq \underline{\underline{R}}$, both bank profits and lending rise with a monetary tightening.*

The intuition behind Proposition 10 is that the strength (but not the sign) of the output response depends on the loan rate pass-through η^l, while the response of bank loans depends on the *difference between loan and deposit pass-through $\eta^l - \eta^d$*. Holding deposit pass-through η^d fixed, a higher loan pass-through η^l would indeed be reflected in a positive correlation between the output response and the loan response, as in Bernanke and Blinder (1992). But crucially, in general equilibrium we cannot hold η^d fixed, as η^l and η^d are jointly determined through banks' balance sheets.

The model sheds light on the alignment of incentives between banks and the rest of the economy regarding the conduct of monetary policy. When bank lending is unconstrained, banks always benefit from rate hikes, in the sense that an increase in the policy rate boosts their return on equity. When bank lending is constrained, banks might instead benefit from accommodative monetary policy: this "central bank's put" was widely discussed in the early stages of the financial crisis (Farhi and Tirole 2012, Diamond and Rajan 2012). However, if either banks get recapitalized back to the unconstrained regime, or if rates reach ultra-low levels and bank lending remains constrained, then banks may start benefiting from rate hikes, which would nevertheless hurt output, albeit not necessarily bank lending.

Evidence on bank profitability. I conclude this section by examining empirically the prediction of the model regarding the impact of monetary shocks on bank profitability. For

Table 3: Response of Fama-French's bank industry portfolio to a 100 bps monetary policy shock.

	(1)	(2)
Δi	-4.38	11.73
	(5.83)	(8.66)
$\Delta i \times$ FFR		-4.85**
		(1.93)
R^2	0.011	0.051
Observations	106	106

Note: The regression equations are (1) $\text{Return}_t = \alpha + \beta\Delta i_t + \epsilon_t$ and (2) $\text{Return}_t = \alpha + \beta_1\Delta i_t + \beta_2 FFR_{t-1} + \gamma\Delta i_t \times FFR_{t-1} + \epsilon_t$. The dependent variable Return_t is the daily return of "Banks" in the 49 Fama-French industry portfolios, taken from Kenneth French's website. Δi_t is the high-frequency Fed funds rate shock from Nakamura and Steinsson (2018). The sample is all regularly scheduled FOMC meetings from 01/01/2000 to 3/19/2014, excluding July 2008 to June 2009. FFR_{t-1} is the Fed funds Rate on the previous day. Robust standard errors in parentheses.

the estimation, I depart slightly from the theory and consider bank stock prices as a measure of profitability because reported measures of the flow of profits, like ROA or ROE, are of low frequency and are smoothed by book-value accounting conventions. The standard regression to estimate the impact of monetary policy on asset prices is

$$\text{Return}_t = \alpha + \beta\Delta i_t + \epsilon_t, \tag{23}$$

where Return_t is the intra-day stock return, and Δi_t is a measure of the monetary shock. Bernanke and Kuttner (2005), Gürkaynak et al. (2005), Gertler and Karadi (2015), and Nakamura and Steinsson (2018) (among others) all estimate equation (23) for various assets (e.g., the market portfolio or long-term bonds), while English et al. (2018) focus on bank stock prices. I estimate instead how this relation changes with the level of the Fed funds rate FFR_{t-1} right before the shock, in the regression[39]

$$\text{Return}_t = \alpha + \beta_1\Delta i_t + \beta_2 FFR_{t-1} + \gamma\Delta i_t \times FFR_{t-1} + \epsilon_t. \tag{24}$$

Table 3 displays my results. For Δi_t I use the unexpected change in the Fed funds rate in a 30-minute window around the FOMC announcement. In the standard specification (23), I find

[39] Ampudia and den Heuvel (2018) perform a similar exercise for European banks, comparing the periods 1999-2008 ("high rates") and 2012-2017 ("low rates").

a negative (but not significant) effect $\beta = -4.38$ on the Fama-French bank industry portfolio's return on FOMC announcement days. The point estimate is consistent with English et al. (2018), who find that bank stocks fall by 8% around FOMC announcements in reaction to an unexpected +100 bps shock to the short rate. In my specification (24), I find a negative and significant coefficient $\gamma = -4.85$ on the interaction term: a Fed tightening hurts bank stocks less (and might even increase them) when interest rates are lower. This pattern is consistent with Proposition 10.

In Appendix A.3, I show that low interest rates have a specific effect on banks, and not just on all industries. I plot the reactions of all 49 Fama-French industries to the same monetary shock, in low and high interest rate subsamples. Table 8, also in Appendix A.3, makes the same point: while there is also a significant negative interaction term $\gamma = -2.32$ for the market portfolio, it is weaker than that for banks. My model also predicts that monetary shocks have a state-dependent effect on the economy as a whole, albeit weaker than on bank values.

6 Short run: quantitative illustration

In this section, I illustrate the quantitative relevance of the mechanisms by calibrating the dynamic model to the U.S. banking sector and the evidence from Section 2. Relative to the two-period model of Section 5.2, the dynamic model adds two ingredients, a stock effect and a flow effect. The stock effect is that monetary policy affects bank lending capacity by revaluating long-term assets. The flow effect is that current spreads not only affect the current leverage ratio, but also, through the return on bank equity, the accumulation of bank capital.[40]

6.1 Calibration

The model period is 1 year. This choice allows for a higher effective duration of deposits and short-term loans, to capture the fact that those rates adjust not only partially but also sluggishly, see Figure 19. My baseline calibration fits the period 2000-2008 and starts with an annual steady state real rate of $r^* = 3\%$, that is approximately the Laubach and Williams (2015) estimate for the period. Table 4 summarizes the calibration.

Liquidity services from money and deposits are combined using the CES aggregator (9). $\epsilon > 1$ is the elasticity of substitution between money and deposits. I impose $u = v = \log$ to remain as close as possible to the neutrality conditions of Proposition 7. I allow for two loan maturities: 60% of short-term loans with maturity $k = 1$ year and 40% of long-term

[40]This flow effect is akin to what Brunnermeier and Sannikov (2016) call the "stealth recapitalization".

loans with maturity $K = 10$ years. This approximates the loan repricing maturity structure from the Call Reports, shown in Figure 13 in Appendix A.1. Short-term credit is meant to capture both small business lending and consumer credit, e.g., adjustable-rate mortgages or unsecured credit in response to health shocks as in Chatterjee et al. (2007). Long-term loans are closer to mortgages.

I set the average deposit spread at 2.5% and the short-term loan spread at 1%. Average "core deposits" defined as checking plus saving deposits at commercial banks, represent 30% of GDP. The stock of loans to GDP averages at 38%. There are multiple ways to achieve these quantities; I set $\gamma^{y,1} = \gamma^{y,K} = 0$, which implies $\gamma^s = 52\%$, $\gamma^{o,1} = 45\%$, and $\gamma^{o,K} = 3\%$. The corresponding target leverage ratio $\phi = 3.8$ is achieved with a pledgeability parameter $\theta = 0.79$. This target leverage is lower than typical numbers because I do not count securities on the asset side or wholesale funding on the liability side. The payout rate $\rho - \iota$ is then set at 17% to ensure the implied bank equity/annual GDP ratio E of 8%.

The remaining parameters are related to liquidity. Over a grid for ϵ, I set $\alpha(\epsilon)$ and $\chi(\epsilon)$ to target the deposits/GDP ratio already mentioned and a 12% money-deposit ratio, which corresponds to half of currency component of M1 (approximately the share held domestically, see, e.g., Schmitt-Grohé and Uribe 2012) divided by core deposits. I then solve for the steady state and compute the reaction of the deposit rate to a 100 bps monetary shock in sticky prices equilibrium. The elasticity of substitution ϵ is set to match my estimates of the 1-year deposit pass-through estimated over 2000-2008 as in Section 2.2. I aggregate over checking deposits, saving deposits and money market deposit accounts according to the weights in the 2004 Survey of Consumer Finances, which yields a 70% pass-through and thus a value $\epsilon = 8$. This yields values $\alpha = 0.97$ and $\chi = 0.02$, close to those in Di Tella and Kurlat (2017).

6.2 Baseline results

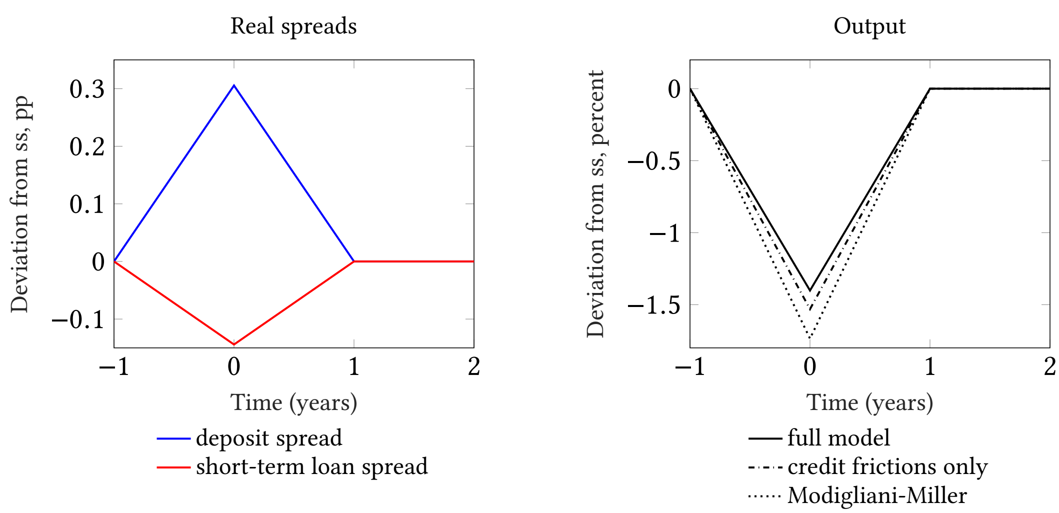
Figure 7 displays the impact response of spreads and output to a +100 bps monetary policy shock that lasts for only one period ($t = 0$). The deposit spread increases by 30 bps, or equivalently the deposit pass-through η_0^d is 0.7. The loan spread falls, but only by around 15 bps, which means a relatively high loan pass-through $\eta_0^l = 0.8$.

The interest-elasticity of output $\frac{d \log Y_0}{d \log R_0} = -\eta_0^Y$ is equal to -1.41. This is higher than the RANK value of 1 (the elasticity of intertemporal substitution) due to the revaluation effects studied by Auclert (2017). Borrowers have higher marginal propensities to consume (MPC) than savers, and they have negative "unhedged interest rate exposures". The negative co-variance between MPC and unhedged rate exposures means that redistribution amplifies the output effect of monetary policy.

Table 4: Calibration. Sources: see text.

Parameter	Description	Value	Target
β	Discount factor	0.97	$r^* = 3\%$
π	Inflation target	2%	
Savers			
χ	Liquidity services	0.02	2.5% deposit spread
α	Weight on money	0.97	money-deposit ratio
ϵ	Elasticity m, d	8	70% deposit pass-through
γ^s	Income share	52%	38% loans/GDP
Borrowers			
$\gamma^{o,1}$	Income share	45%	1% loan spread
$\gamma^{o,K}$	"	3%	40% long-term loans
Banks			
ρ	Net payout rate	17%	30% deposits/GDP
θ	Pledgeability	0.79	bank leverage

Figure 7: Response to a +100 bps monetary shock at $r^* = 3\%$.



Two benchmarks help to evaluate the specific role of banks. The first benchmark is a Modigliani-Miller economy where bonds, loans and deposits are perfect substitutes: both deposit and loan spreads are zero and thus there is full pass-through. The Modigliani-Miller case still features heterogeneity and redistribution, hence it allows us to isolate the role of financial frictions. In fact, by Proposition 1, the output response in an economy without loan spreads (but potentially non-zero deposit spreads) is the same as in the Modigliani-Miller benchmark.

The second benchmark is a "credit frictions only" economy, that goes further by allowing for banks and constrained lending, but shutting down liquidity frictions (i.e., setting $v = 0$). Comparing the full model to the "credit frictions only" benchmark highlights the interaction between deposit and loan markets.

The interest-elasticity of output is equal to -1.53 in the "credit frictions only" economy and -1.74 in the Modigliani-Miller economy. Thus credit frictions alone dampen the New Keynesian transmission mechanism by 12%, while in the full model the output response is muted by 20% relative to the Modigliani-Miller benchmark.

6.3 Low r^* and monetary policy transmission

I now turn to my key counterfactual exercise, which examines how a lower r^* affects monetary policy transmission. Figure 8 varies the steady state real interest rate r^* (by changing β) from 3% to -1% while keeping the rest of the calibration unchanged. At each steady state r^*, I compute the response to a date-0 unanticipated +100 bps monetary shock that lasts for only one period. The figures display the impact responses of retail rates and output.

Both pass-through and output sensitivity are lower at lower r^*. Comparing the full model to the "credit frictions only" and the "Modigliani-Miller" benchmarks shows that this is entirely due to the interaction between deposit and loan rates, as r^* has no effect on monetary policy transmission to output if there are no frictions, only liquidity frictions, or only credit frictions. When r^* varies from 3% to -1%, the deposit pass-through falls from 0.7 to 0.4. If there were no credit frictions, this would have no bite on the output effects of monetary policy; but in the full model, loan pass-through drops from 0.8 to 0.5 and, as a result, the output effect falls by 13% (in absolute value), from -1.41 to -1.23. We saw that financial frictions mute the output response by 20% relative to the Modigliani-Miller benchmark in the baseline $r^* = 3\%$ calibration; the dampening reaches 32% at $r^* = -1\%$.

Over the whole range of r^*, bank lending falls when monetary policy tightens, as shown in Figure 5. In theory, the opposite could happen at very low rates as in Section 5, but under the calibrated maturity structure and elasticity of substitution between money and deposits,

Figure 8: Date-0 effects of monetary policy on rates and output as a function of r^*.

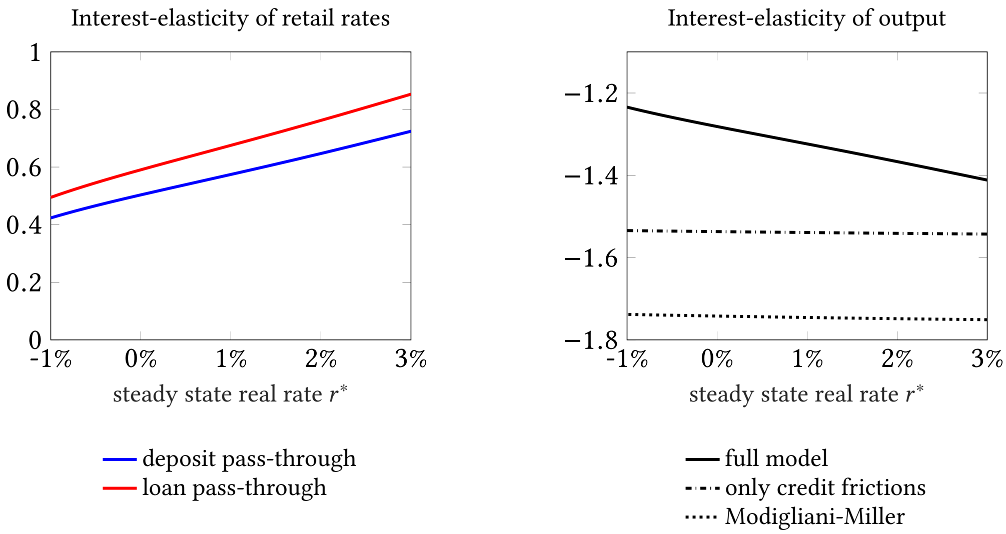
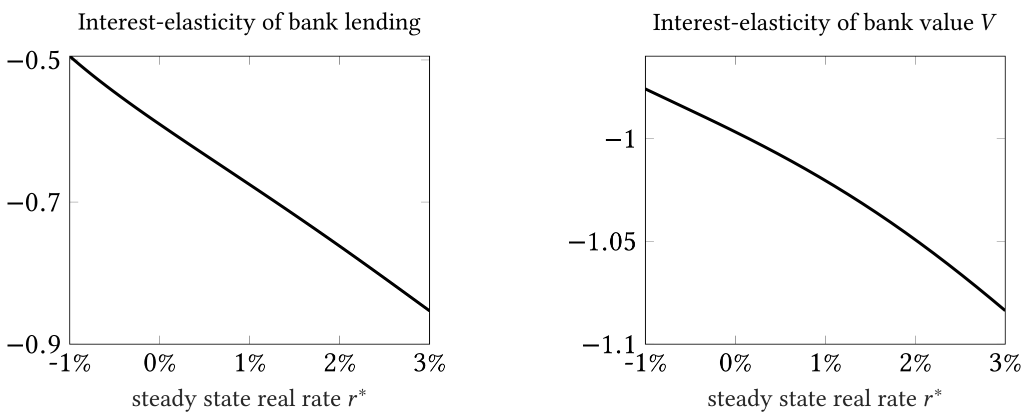


Figure 9: Date-0 effects of monetary policy on bank lending and value as a function of r^*.



deposit pass-through remains high enough and revaluation effects strong enough to prevent any "reversal" of bank lending. I investigate the role of maturity mismatch in the next section. Finally, we can see in Figure 9 that bank value V is negatively affected by contractionary monetary shocks, but less so at lower r^*. Quantitatively, however, the effect of low interest rates is modest, relative to my empirical results in Table 3. One reason is that my model features no risk, and Bernanke and Kuttner (2005) find a significant effect of monetary policy on risk premia. An interesting extension would thus be to incorporate aggregate risk and study how it interacts with deposit and loan spreads.

6.4 Inflation target and capital requirements

My model can be used to analyze how monetary policy transmission depends on two steady state policies: the inflation target, previously set at $\pi = 2\%$, and bank regulation.

Inflation target. Recall from equation (8) that deposit demand $\mathcal{D}$ depends on the nominal interest rate i. So far I have kept the inflation target fixed and let the steady state real interest rate r^* vary. We saw in Section 4 that monetary policy is not superneutral in the constrained lending regime because inflation can affect deposit and loan spreads. Similarly, the inflation target can affect monetary policy transmission in the short run by changing the steady state nominal interest rate. Figure 10 varies the inflation target while keeping r^* fixed at 3%. A higher inflation target improves deposit pass-through, hence loan pass-through. The intuition is exactly as in the previous case where I varied r^*.

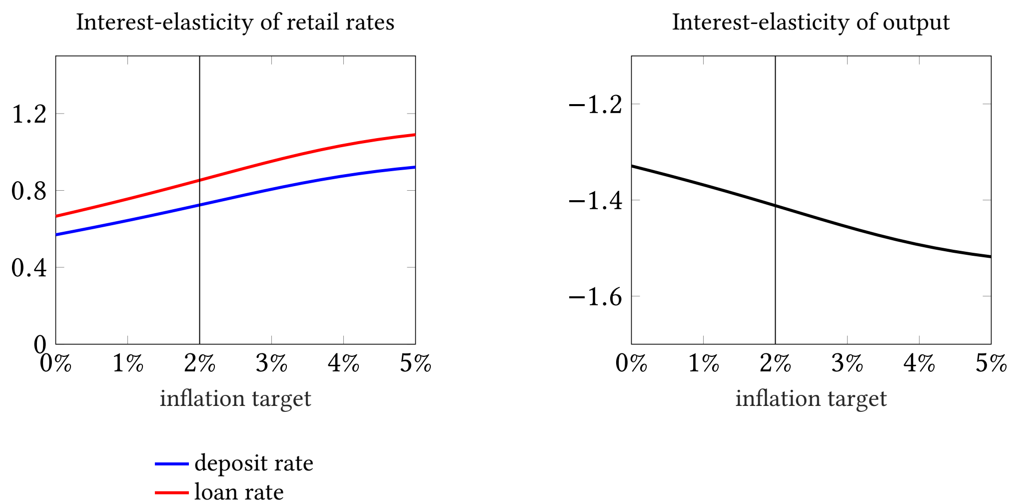
Capital requirements. The level of the nominal interest rate has a direct effect on monetary policy transmission through the substitutability between money and deposits, but also an indirect effect working through the endogenous steady state level of bank capital E: recall from Section 4 that all else equal, a higher steady state nominal interest rate will boost the liquidity premia earned by banks and thus increase E, with a potential side effect of reducing steady state loan spreads. We can isolate the link between bank capital and monetary policy transmission by holding interest rates fixed and varying either net payouts $\rho - \zeta$ or the pledgeability parameter θ.

I now depict the impact of varying θ, as this has an interpretation in terms of capital requirements (see the discussion in Section 3.5 about the banks' financial constraints). Figure 11 shows how bank regulation affects the pass-through of monetary policy to retail rates and output. As θ varies, I plot interest-elasticities against the steady state leverage ratio of deposits over equity

$$\overline{\phi} = \frac{\theta R^l / R^d}{1 - \theta R^l / R^d},$$

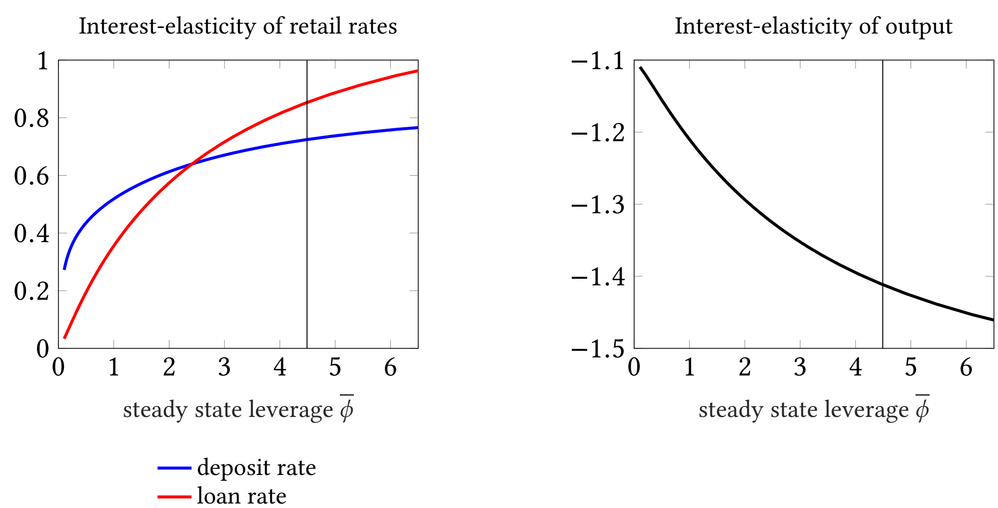
taking into account all the general equilibrium responses of capital E and spreads τ^l and τ^d. The result is that tighter bank regulation hampers monetary policy transmission. Several effects are at play. First, when banks are more constrained (lower θ), monetary shocks relax banks' lending capacity by less. Second, steady state deposit spreads are higher hence deposit rates are lower, and thus closer to the low nominal rates region where money and deposits are better substitutes. This dampens deposit pass-through, which again translates into lower loan pass-through and weaker output effects of monetary policy.

Figure 10: Date-0 effects of monetary policy on rates and output as a function of the steady state inflation target.



Note: Vertical lines denote the baseline calibration.

Figure 11: Date-0 effects of monetary policy on rates and output as a function of steady state bank leverage $\overline{\phi}$.



Note: Vertical lines denote the baseline calibration.

6.5 Maturity mismatch and hedging

Revaluation effects are thought to be an important channel through which monetary policy affects bank lending (Van den Heuvel 2002, Brunnermeier and Koby 2018). I now investigate how banks' maturity mismatch affect the transmission of monetary policy to rates and output in general equilibrium. Recall that bank lending capacity is given by (13):

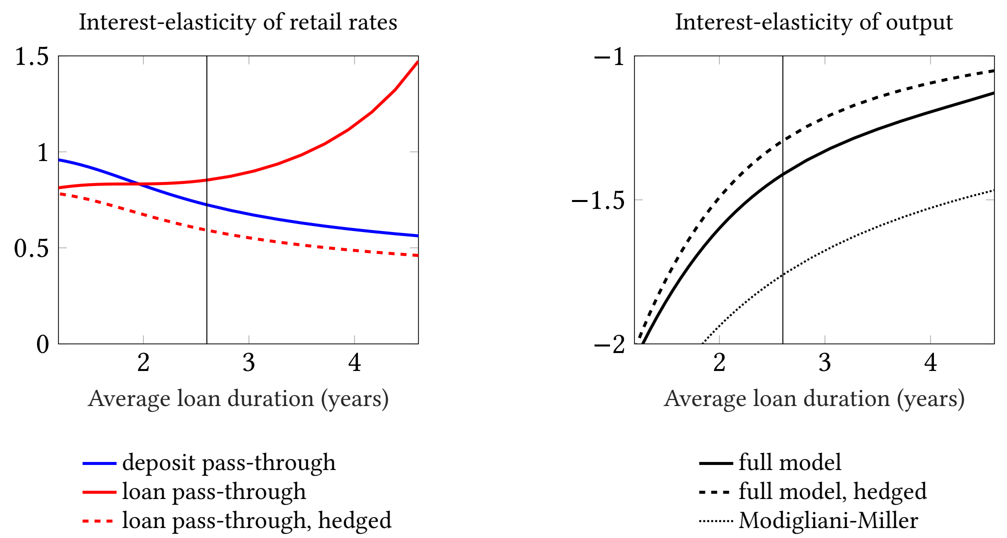
$$\underbrace{\sum_{k=1}^{K} q^l_{t,t+k}\left(L^{t+1}_{t+k} - L^t_{t+k}\right)}_{\text{new loans issued at } t} \leq \left(1 + \overline{\phi}_t\right)\left(L^t_t - D_t - \mathrm{Div}_t\right) + \overline{\phi}_t \sum_{k=1}^{K-1} q^l_{t,t+k} L^t_{t+k}.$$

Monetary policy can affect credit supply in two ways: by changing bank leverage $\overline{\phi}_t$ or by changing the value of long-term assets $\sum_{k=1}^{K-1} q^l_{t,t+k} L^t_{t+k}$. In the two-period model, book equity E_0 is fixed and banks' lending capacity Λ_0 can only vary with the leverage ratio $\overline{\phi}_0$. In the dynamic model, capital gains or losses on long-term assets also affect Λ_0, and the difference between η^l_0 and η^d_0 is no longer sufficient to tell how bank lending reacts to monetary policy.

To isolate the role of maturity mismatch, I introduce another benchmark, the "hedged banks" model: bank equity E_t is fully hedged and remains equal to its steady state value after monetary policy shocks. Interest rate risk is shifted to savers. Figure 12 shows how maturity mismatch affects pass-through and the output effect of monetary policy. Total loans and deposits are kept fixed (at respectively 38% and 30% of GDP), and the model is recalibrated to vary only the share of short-term loans (equal to 60% in the baseline calibration). The left panel of Figure 12 shows how pass-through depends on maturity mismatch. As average loan duration grows, so does the difference between the full model and the "hedged" economy. Loan pass-through increases with duration due to a stronger revaluation effect on banks, and it can even exceed 1 when duration is high enough. By contrast, loan pass-through actually declines with maturity in the "hedged" benchmark.

The right panel of Figure 12 shows that output sensitivity decreases with loan duration. This is due to the force described in Auclert (2017), as illustrated by the "Modigliani-Miller" benchmark in the right panel. Even absent banking frictions, longer loan durations imply smaller unhedged rate exposures for borrowers, and thus a smaller redistribution channel of monetary policy. The "hedged" model controls for the Auclert (2017) redistribution channel, so that comparing the full model to the hedged economy isolates the pure effect of the revaluation of bank equity. Output effects of monetary policy are higher when banks do *not* hedge: when the central bank tightens, bank lending capacity Λ falls, which amplifies the contractionary effect of the rate hike. Quantitatively, the absence of bank hedging increases output sensitivity from by 9% at the baseline duration of 2.6 years.

Figure 12: Date-0 effects of monetary policy on rates and output as a function of average bank loan duration.



Note: Vertical lines denote the baseline calibration. The "hedged" benchmark is an economy where bank equity remains at its steady state value and interest rate risk is fully shifted to savers.

7 Conclusion

In this paper, I argue that the dual role of commercial banks as credit and liquidity providers has implications for interest rate pass-through in both the long and short runs. Because privately-issued deposits compete with publicly-issued money, the level of the nominal interest rate affects real loan and deposit spreads. A lower nominal rate compresses deposit spreads and widens loan spreads. In the long run, the reaction of spreads explains why the persistent decline in bond rates has not been fully transmitted to loan rates. In the short run, the response of spreads dampens the transmission of monetary shocks to output. Moreover, pass-through is lower at lower rates in the model and in the data, thus monetary policy becomes less potent in a world of low interest rates. Policy can trade-off liquidity frictions against credit frictions: raising the inflation target decreases loan spreads and improves monetary policy transmission, but only at the cost of higher deposit spreads.

References

Aiyagari, S. Rao, R. Anton Braun, and Zvi Eckstein (1998), "Transaction Services, Inflation, and Welfare", *Journal of Political Economy*, Vol. 106, pp. 1274–1301.

Altavilla, Carlo, Miguel Boucinha, and Jose-Luis Peydro (2017), "Monetary Policy and Bank Profitability in a Low Interest Rate Environment", working paper 2105.

Ampudia, Miguel and Skander Van den Heuvel (2018), "Monetary policy and bank equity values in a time of low interest rates", working paper.

Auclert, Adrien (2017), "Monetary Policy and the Redistribution Channel", working paper.

Ball, Laurence (2012), "Short-run money demand", *Journal of Monetary Economics*, Vol. 59, pp. 622 – 633.

Barnett, William A. (1980), "Economic monetary aggregates an application of index number and aggregation theory", *Journal of Econometrics*, Vol. 14, pp. 11 – 48.

Basten, Christoph and Mike Mariathasan (2018), "How Banks Respond to Negative Interest Rates: Evidence from the Swiss Exemption Threshold", working paper.

Begenau, Juliane, Monika Piazzesi, and Martin Schneider (2015), "Banks' Risk Exposures", working paper.

Begenau, Juliane and Erik Stafford (2018), "Do Banks have an Edge?", working paper.

Berger, Allen N. and Gregory F. Udell (1992), "Some Evidence on the Empirical Significance of Credit Rationing", *Journal of Political Economy*, Vol. 100, pp. 1047–1077.

Bernanke, Ben S. and Alan S. Blinder (1988), "Credit, Money, and Aggregate Demand", *The American Economic Review*, Vol. 78, pp. 435–439, Papers and Proceedings of the One-Hundredth Annual Meeting of the American Economic Association.

——— (1992), "The Federal Funds Rate and the Channels of Monetary Transmission", *American Economic Review*, Vol. 82, pp. 901–921.

Bernanke, Ben S. and Kenneth N. Kuttner (2005), "What Explains the Stock Market's Reaction to Federal Reserve Policy?", *The Journal of Finance*, Vol. 60, pp. 1221–1257.

Bianchi, Javier and Saki Bigio (2017), "Banks, Liquidity Management and Monetary Policy", working paper.

Boivin, Jean and Marc P Giannoni (2006), "Has Monetary Policy Become More Effective?", *The Review of Economics and Statistics*, Vol. 88, pp. 445–462.

Boivin, Jean, Michael T. Kiley, and Frederic S. Mishkin (2010), "Chapter 8 - How Has the Monetary Transmission Mechanism Evolved Over Time?", Vol. 3 of Handbook of Monetary Economics: Elsevier, pp. 369 – 422.

Brunnermeier, Markus K. and Yann Koby (2018), "The Reversal Interest Rate: The Effective Lower Bound of Monetary Policy", working paper.

Brunnermeier, Markus K. and Yuliy Sannikov (2016), "The I Theory of Money", working paper 22533, National Bureau of Economic Research.

Burstein, Ariel and Gita Gopinath (2014), "International Prices and Exchange Rates", *Handbook of International Economics*, Vol. 4, pp. 391–451.

Caballero, Ricardo J., Emmanuel Farhi, and Pierre-Olivier Gourinchas (2017), "The Safe Assets Shortage Conundrum", *Journal of Economic Perspectives*, Vol. 31, pp. 29–46.

Chatterjee, Satyajit, Dean Corbae, Makoto Nakajima, and José-Víctor Ríos-Rull (2007), "A Quantitative Theory of Unsecured Consumer Credit with Risk of Default", *Econometrica*, Vol. 75, pp. 1525–1589.

Chetty, V. Karuppan (1969), "On Measuring the Nearness of the Near-Moneys", *American Economic Review*, Vol. 59, pp. 270–81.

Claessens, Stijn, Nicholas Coleman, and Michael Donnelly (2018), "Low-For-Long interest rates and banks' interest margins and profitability: Cross-country evidence", *Journal of Financial Intermediation*, Vol. 35, pp. 1–16.

Committee on the Global Financial System (2018), "Financial stability implications of a prolonged period of low interest rates",Technical Report 61.

Cúrdia, Vasco and Michael Woodford (2016), "Credit frictions and optimal monetary policy", *Journal of Monetary Economics*, Vol. 84, pp. 30–65.

Cysne, Rubens Penha and David Turchick (2010), "Welfare costs of inflation when interest-bearing deposits are disregarded: A calculation of the bias", *Journal of Economic Dynamics and Control*, Vol. 34, pp. 1015–1030.

De Bondt, Gabe (2002), "Retail Bank Interest Rate Pass-Through: New Evidence at the Euro Area Level", ECB Working Paper 136.

Del Negro, Marco, Domenico Giannone, Marc P. Giannoni, and Andrea Tambalotti (2018), "Global Trends in Interest Rates", working paper.

Dewatripont, Mathias and Jean Tirole (1994), *The Prudential Regulation of Banks*: M.I.T. Press.

Di Maggio, Marco, Amir Kermani, and Christopher Palmer (2016), "How Quantitative Easing Works: Evidence on the Refinancing Channel", working paper.

Di Tella, Sebastian and Pablo Kurlat (2017), "Why are Banks Exposed to Monetary Policy?", working paper.

Diamond, Douglas W. and Raghuram G. Rajan (2002), "A Theory of Bank Capital", *The Journal of Finance*, Vol. 55, pp. 2431–2465.

——— (2012), "Illiquid Banks, Financial Stability, and Interest Rate Policy", *Journal of Political Economy*, Vol. 120, pp. 552–591.

Doepke, Matthias and Martin Schneider (2006), "Inflation and the Redistribution of Nominal Wealth", *Journal of Political Economy*, Vol. 114, pp. 1069–1097.

Drechsler, Itamar (2018), "Discussion of "Do Banks Have an Edge" by Juliane Begenau and Erik Stafford", NBER Summer Institute Corporate Finance.

Drechsler, Itamar, Alexi Savov, and Philipp Schnabl (2017), "The Deposits Channel of Monetary Policy", *The Quarterly Journal of Economics*, Vol. 132, pp. 1819–1876.

——— (2018), "Banking on Deposits: Maturity Transformation without Interest Rate Risk", working paper.

Driscoll, John C. and Ruth A. Judson (2013), "Sticky Deposit Rates", working paper.

Egan, Mark, Stefan Lewellen, and Aditya Sunderam (2017), "The Cross Section of Bank Value", working paper.

Eggertsson, Gauti B., Ragnar E. Juelsrud, and Ella Getz Wold (2017), "Are Negative Nominal Interest Rates Expansionary?", Working paper 24039, National Bureau of Economic Research.

Elenev, Vadim, Tim Landvoigt, and Stijn Van Nieuwerburgh (2018), "A Macroeconomic Model with Financially Constrained Producers and Intermediaries", working paper.

English, William B. (1999), "Inflation and financial sector size", *Journal of Monetary Economics*, Vol. 44, pp. 379 – 400.

English, William B., Skander J. Van den Heuvel, and Egon Zakrajsek (2018), "Interest rate risk and bank equity valuations", *Journal of Monetary Economics*.

Farhi, Emmanuel and Jean Tirole (2012), "Collective Moral Hazard, Maturity Mismatch, and Systemic Bailouts", *American Economic Review*, Vol. 102, pp. 60–93.

Feenstra, Robert C. (1986), "Functional equivalence between liquidity costs and the utility of money", *Journal of Monetary Economics*, Vol. 17, pp. 271 – 291.

Fuster, Andreas and James Vickery (2018), "Regulation and Risk Shuffling in Bank Securities Portfolios", Federal Reserve Bank of New York Staff Reports 851.

Galí, Jordi and Luca Gambetti (2009), "On the Sources of the Great Moderation", *American Economic Journal: Macroeconomics*, Vol. 1, pp. 26–57.

Gerali, Andrea, Stefano Neri, Luca Sessa, and Federico M. Signoretti (2010), "Credit and Banking in a DSGE Model of the Euro Area", *Journal of Money, Credit and Banking*, Vol. 42, pp. 107–141.

Gertler, Mark and Peter Karadi (2011), "A model of unconventional monetary policy", *Journal of Monetary Economics*, Vol. 58, pp. 17 – 34, Carnegie-Rochester Conference Series on Public Policy: The Future of Central Banking April 16-17, 2010.

——— (2015), "Monetary Policy Surprises, Credit Costs, and Economic Activity", *American Economic Journal: Macroeconomics*, Vol. 7, pp. 44–76.

Gertler, Mark and Nobuhiro Kiyotaki (2010), "Financial Intermediation and Credit Policy in Business Cycle Analysis", in *Handbook of Monetary Economics*, Vol. 3, Chap. 11.

Gilchrist, Simon and Egon Zakrajsek (2012), "Credit Spreads and Business Cycle Fluctuations", *American Economic Review*, Vol. 102, pp. 1692–1720.

Gomez, Matthieu, Augustin Landier, David Sraer, and David Thesmar (2016), "Banks' Exposure to Interest Rate Risk and the Transmission of Monetary Policy", working paper.

Goodfriend, Marvin and Bennett T. McCallum (2007), "Banking and interest rates in monetary policy analysis: A quantitative exploration", *Journal of Monetary Economics*, Vol. 54, pp. 1480–1507.

Gourinchas, Pierre-Olivier and Hélène Rey (2018), "Global Real Rates: A Secular Approach", working paper.

Greenwood, Jeremy, Zvi Hercowitz, and Gregory W. Huffman (1988), "Investment, Capacity Utilization, and the Real Business Cycle", *American Economic Review*, Vol. 78, pp. 402–417.

Greenwood, Robin, Samuel G. Hanson, and Jeremy C. Stein (2015), "A Comparative-Advantage Approach to Government Debt Maturity", *The Journal of Finance*, Vol. 70, pp. 1683–1722.

Greenwood, Robin, Jeremy C Stein, Samuel G Hanson, and Adi Sunderam (2017), "Strengthening and streamlining bank capital regulation", *Brookings Papers on Economic Activity*, Vol. 2017, pp. 479–565.

Gürkaynak, Refet S., Brian Sack, and Eric Swanson (2005), "The Sensitivity of Long-Term Interest Rates to Economic News: Evidence and Implications for Macroeconomic Models", *American Economic Review*, Vol. 95, pp. 425–436.

Hannan, Timothy H. and Allen N. Berger (1991), "The Rigidity of Prices: Evidence from the Banking Industry", *The American Economic Review*, Vol. 81, pp. 938–945.

Heider, Florian, Farzard Saidi, and Glenn Schepens (2018), "Life Below Zero: Bank Lending Under Negative Policy Rates", working paper.

Holmström, Bengt and Jean Tirole (1997), "Financial Intermediation, Loanable Funds, and The Real Sector", *The Quarterly Journal of Economics*, Vol. 112, pp. 663–691.

——— (1998), "Private and Public Supply of Liquidity", *Journal of Political Economy*, Vol. 106, pp. 1–40.

Hong, Gee Hee and John Kandarac (2018), "Pushed Past the Limit? How Japanese Banks Reacted to Negative Interest Rates", working paper.

Ireland, Peter N. (2015), "Comment on: "On the stability of money demand" by Robert E. Lucas Jr. and Juan Pablo Nicolini", *Journal of Monetary Economics*, Vol. 73, pp. 66 – 69, Carnegie-Rochester-NYU Conference Series on Public Policy.

Jordá, Óscar (2005), "Estimation and Inference of Impulse Responses by Local Projections", *American Economic Review*, Vol. 95, pp. 161–182.

Kaplan, Greg, Benjamin Moll, and Giovanni L. Violante (2018), "Monetary Policy According to HANK", *American Economic Review*, Vol. 108, pp. 697–743.

Karagiannis, Stelios, Yannis Panagopoulos, and Prodromos Vlamis (2010), "Interest rate pass-through in Europe and the US: Monetary policy after the financial crisis", *Journal of Policy Modeling*, Vol. 32, pp. 323 – 338.

Kashyap, Anil K. and Jeremy C. Stein (1995), "The impact of monetary policy on bank balance sheets", *Carnegie-Rochester Conference Series on Public Policy*, Vol. 42, pp. 151 – 195.

Krishnamurthy, Arvind and Annette Vissing-Jorgensen (2012), "The Aggregate Demand for Treasury Debt", *Journal of Political Economy*, Vol. 120, pp. 233–267.

——— (2015), "The impact of Treasury supply on financial sector lending and stability", *Journal of Financial Economics*, Vol. 118, pp. 571 – 600, NBER Symposium on New perspectives on corporate capital structures.

Laubach, Thomas and John C. Williams (2015), "Measuring the Natural Rate of Interest Redux", working paper.

Lucas, Robert E. and Juan Pablo Nicolini (2015), "On the stability of money demand", *Journal of Monetary Economics*, Vol. 73, pp. 48 – 65, Carnegie-Rochester-NYU Conference Series on Public Policy.

Lucas, Robert E. and Nancy L. Stokey (1987), "Money and Interest in a Cash-in-Advance Economy", *Econometrica*, Vol. 55, pp. 491–513.

Mian, Atif, Kamalesh Rao, and Amir Sufi (2013), "Household Balance Sheets, Consumption, and the Economic Slump*", *The Quarterly Journal of Economics*, Vol. 128, pp. 1687–1726.

Mojon, Benoit (2000), "Financial Structure and the Interest Rate Channel of ECB Monetary Policy", ECB Working Paper 40.

Nagel, Stefan (2016), "The Liquidity Premium of Near-Money Assets", *The Quarterly Journal of Economics*, Vol. 131, pp. 1927–1971.

Nakamura, Emi and Jón Steinsson (2018), "High-Frequency Identification of Monetary Non-Neutrality: The Information Effect*", *The Quarterly Journal of Economics*, Vol. 133, pp. 1283–1330.

Neumark, David and Steven A. Sharpe (1992), "Market Structure and the Nature of Price Rigidity: Evidence from the Market for Consumer Deposits*", *The Quarterly Journal of Economics*, Vol. 107, pp. 657–680.

Piazzesi, Monika and Martin Schneider (2018), "Payments, Credit and Asset Prices", working paper.

Poterba, James M. and Julio J. Rotemberg (1987), "Money in the Utility Function: An Empirical Implementation", in W. Barnett and K. Singleton eds. *New Approaches to Monetary Economics*, pp. 219–240: Cambridge University Press.

Ramey, Valerie A. (2016), "Chapter 2 - Macroeconomic Shocks and Their Propagation", Vol. 2 of Handbook of Macroeconomics: Elsevier, pp. 71 – 162.

Rocheteau, Guillaume, Randall Wright, and Cathy Zhang (2018), "Corporate Finance and Monetary Policy", *American Economic Review*, Vol. 108, pp. 1147–1186.

Rognlie, Matthew (2016), "What Lower Bound? Monetary Policy with Negative Interest Rates", working paper.

Schmitt-Grohé, Stephanie and Martín Uribe (2012), "Foreign Demand for Domestic Currency and the Optimal Rate of Inflation", *Journal of Money, Credit and Banking*, Vol. 44, pp. 1207–1224.

Schmitt-Grohé, Stephanie and Martín Uribe (2016), "Downward nominal wage rigidity, currency pegs, and involuntary unemployment", *Journal of Political Economy*, Vol. 124, pp. 1466–1514.

Schwert, Michael (2018), "Does Borrowing from Banks Cost More than Borrowing from the Market?", working paper.

Tenreyro, Silvana and Gregory Thwaites (2016), "Pushing on a String: US Monetary Policy Is Less Powerful in Recessions", *American Economic Journal: Macroeconomics*, Vol. 8, pp. 43–74.

Van den Heuvel, Skander J. (2002), "The Bank Capital Channel of Monetary Policy", working paper.

Werning, Iván (2015), "Incomplete Markets and Aggregate Demand", working paper 21448, National Bureau of Economic Research.

Wong, Arlene (2018), "Transmission of Monetary Policy to Consumption and Population Aging", working paper.

Woodford, Michael (2003), *Interest and Prices: Foundations of a Theory of Monetary Policy*: Princeton University Press.

Yankov, Vladimir (2014), "In Search of a Risk-Free Asset", working paper.

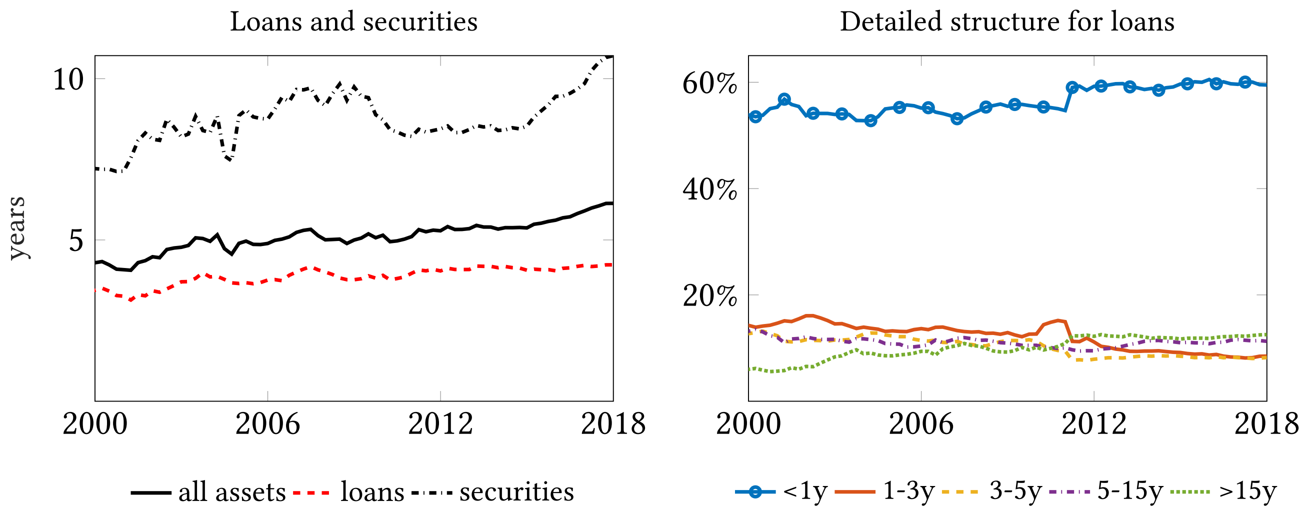
Zentefis, Alexander K. (2018), "Bank Net Worth and Frustrated Monetary Policy", working paper.

Appendix

A Figures and tables

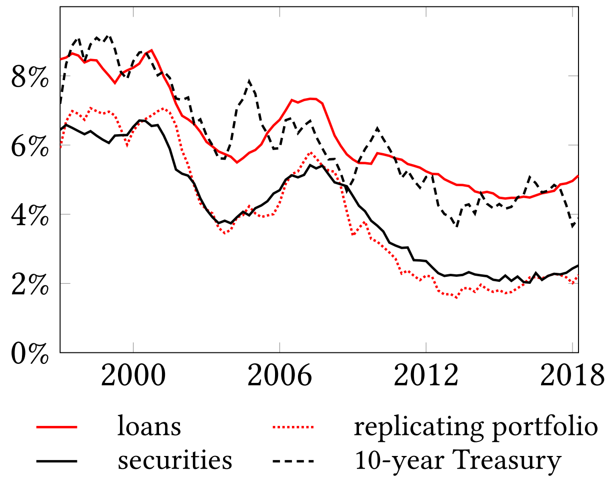
A.1 Long run

Figure 13: Repricing maturity structure of U.S. banks' assets.



Note: This figure shows the maturity structure of U.S. banks' assets between 1997Q3 and 2018Q1. Left: Average repricing maturity period of all assets, loans, and securities. Right: Loan repricing maturities corresponding to the buckets in the Call Reports, expressed as shares of total loans. Source: Call Reports.

Figure 14: Realized returns on loans, securities, and Treasury portfolios.



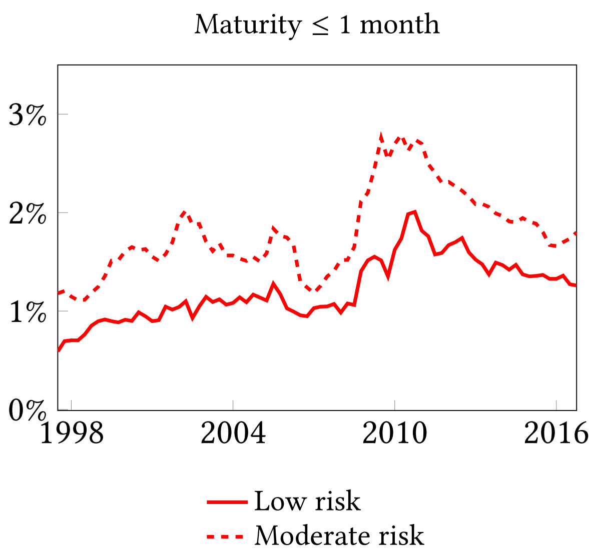
Note: This figure shows the average (asset-weighted) realized returns on different assets earned by the U.S. banking sector between 1997Q3 and 2018Q1. All rates of return are computed at book value. The solid lines are the return on loans and securities from the FDIC. The black dashed line is the realized return on a 10-year Treasury bond. The dotted red line is the realized return on the replicating portfolio described in Section 2.1. Source: Call Reports, Federal Reserve and author's calculations.

Table 5: Spreads against rates.

	(1) Loan spread	(2) Deposit spread
Treasury	-0.31***	0.29***
	(0.02)	(0.02)
R^2	0.729	0.753
Observations	85	85

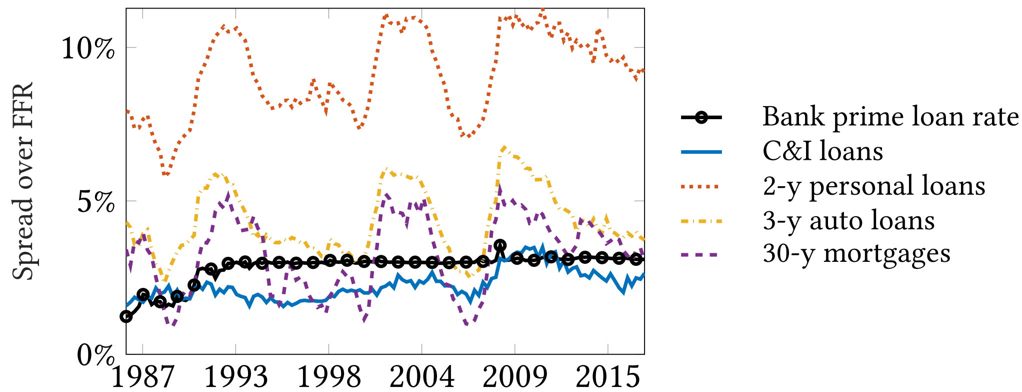
Note: This table shows the effect of interest rates on loan and deposit spreads. The regression equations are $\text{Spread}_t = \alpha + \gamma \text{Treasury}_t + \epsilon_t$ where Spread_t is the loan spread or the deposit spread constructed in the main text, and Treasury_t is the realized return on the Treasury portfolio constructed in the main text. Robust standard errors in parentheses.

Figure 15: Spread between rate on commercial and industrial loans and Fed funds rate. Source: Federal Reserve Release E.2.



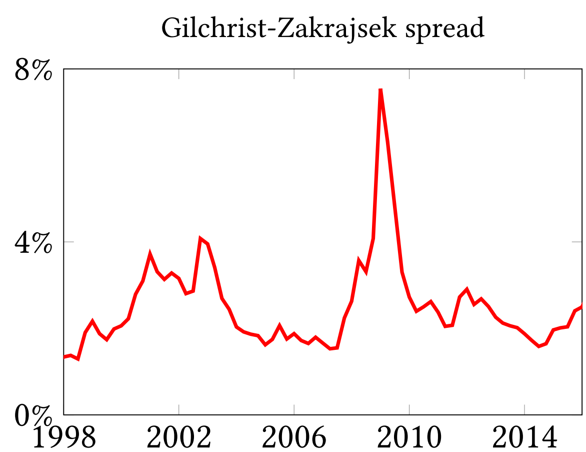
Note: This figure shows the spread between commercial and industrial loans with maturity less than 1 month and the Fed funds rate between 1998 and 2016, for loans classified as "low risk" and "moderate risk". The spread on both types of loans has increased over time by approximately the same amount, which suggests that higher spreads do not reflect higher credit risk premia.

Figure 16: Spreads over Fed funds rate. Source: Federal Reserve Board.



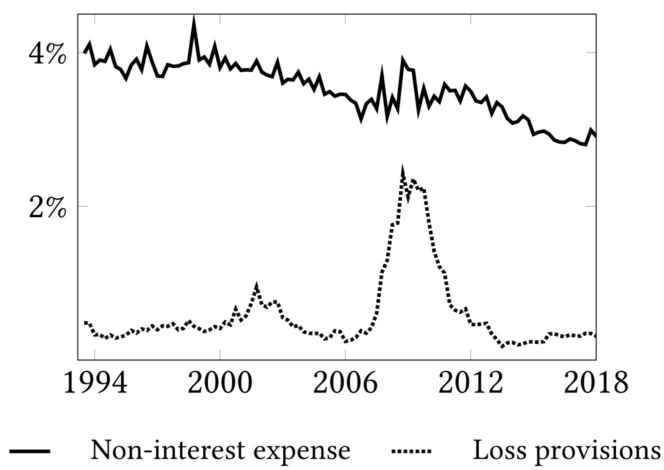
Note: This figure shows that the "bank prime loan rate", which is the main lending rate reported by commercial banks and the Federal Reserve, does not reflect the actual loan rates paid by firms and consumers.

Figure 17: Credit spread on corporate bonds from Gilchrist and Zakrajsek (2012), updated through August 2016.



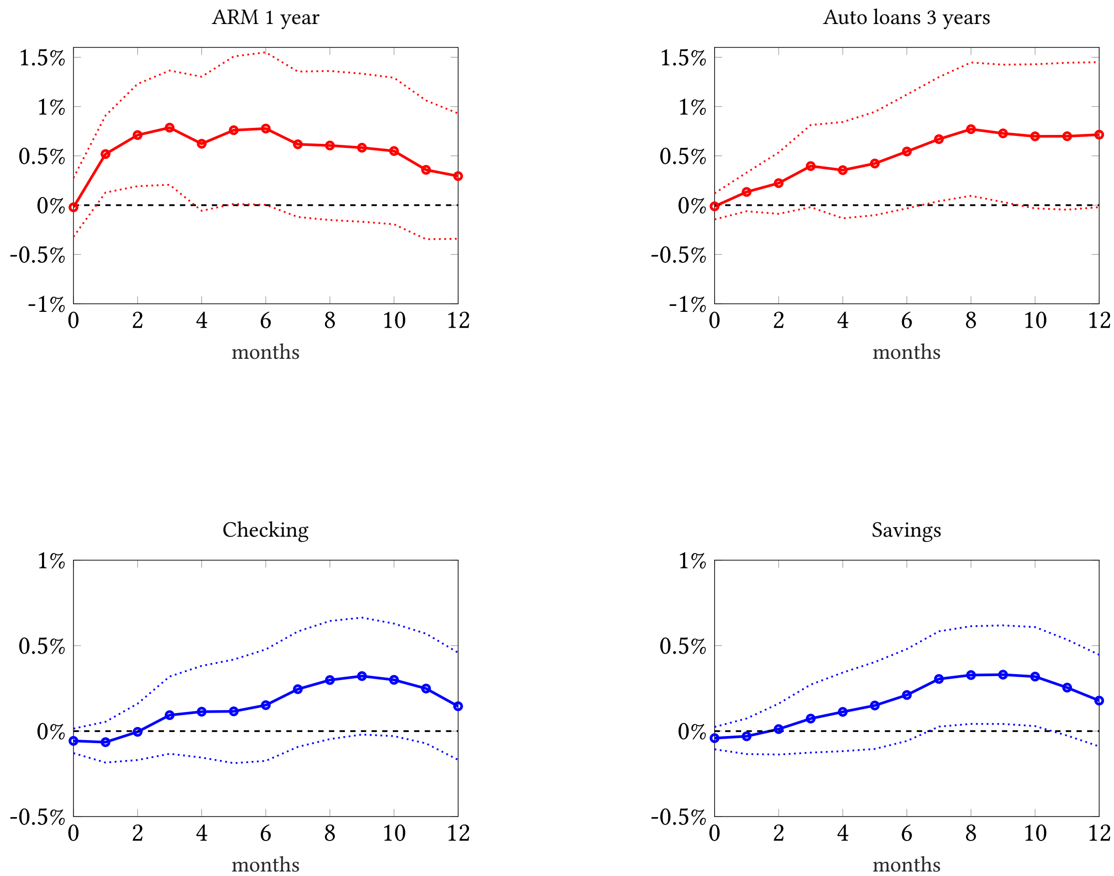
Note: This figure shows that unlike the loan spread, the spread on corporate bonds (computed over Treasuries with the same maturity) has not increased with the decline in interest rates. The spread averages at 2.2% between 1997 and 2007, and 2.3% between 2010 and 2016.

Figure 18: Components of bank profits, as percentage of earning assets. Source: FDIC.



Note: This figure shows two components of bank profits: non-interest expense and loan loss provisions. Non-interest expense has declined steadily, driven by a decline in data processing costs. After 2012, loan loss provisions have reverted to levels similar to or lower than pre-2008.

Figure 19: Additional pass-through (in percentage points) of a monetary shock to retail rates when the 1-year rate is 100 bps higher with two-way clustered standard errors.



Note: The regression equations are $y_{b,t+h} - y_{b,t-1} = \alpha_{b,h} + \delta_{1h}\Delta i_t + \delta_{2h} i_{t-1} + \beta_h \Delta i_t \times i_{t-1} + \gamma_h \text{controls}_{b,t-1} + \epsilon_{b,t+h}$ for each horizon h. The figures show the sequences $\{\hat{\beta}_h\}_{h=0,\ldots,12}$ with 90% confidence bands. Standard errors are two-way clustered at the branch-month level. Sources: Federal Reserve and RateWatch.

A.2 Short run

Asymmetric pass-through. One potential explanation for the state-dependence I find is asymmetric pass-through, under two conditions: (i) deposit rates adjust more downward than upward, while loan rates adjust more upward than downward; (ii) the policy rate i_t tends to go up when low, and down when high. To examine part (i), I estimate

$$\begin{aligned} y_{i,t+h} - y_{i,t-1} = {} & \mathbb{I}(\Delta i_t < 0)\left[\alpha^-_{i,h} + \beta^-_h \Delta i_t + \gamma^-_h \text{controls}_{i,t-1}\right] \\ & + \mathbb{I}(\Delta i_t > 0)\left[\alpha^+_{i,h} + \beta^+_h \Delta i_t + \gamma^-_h \text{controls}_{i,t-1}\right] + \epsilon_{i,t+h}. \end{aligned} \tag{25}$$

Table 6: Asymmetric pass-through.

	$\Delta i_t > 0$	$\Delta i_t < 0$
Checking deposits	0.17	0.59
	(0.004)	(0.006)
Saving deposits	0.47	0.53
	(0.004)	(0.003)
Money market deposits	0.91	1.11
	(0.009)	(0.008)
Adjustable rate mortgages, 1 year	2.67	0.72
	(0.09)	(0.06)
Personal loans, 2 years	0.73	0.16
	(0.37)	(0.20)
Auto loans, 3 years	2.46	0.25
	(0.027)	(0.016)

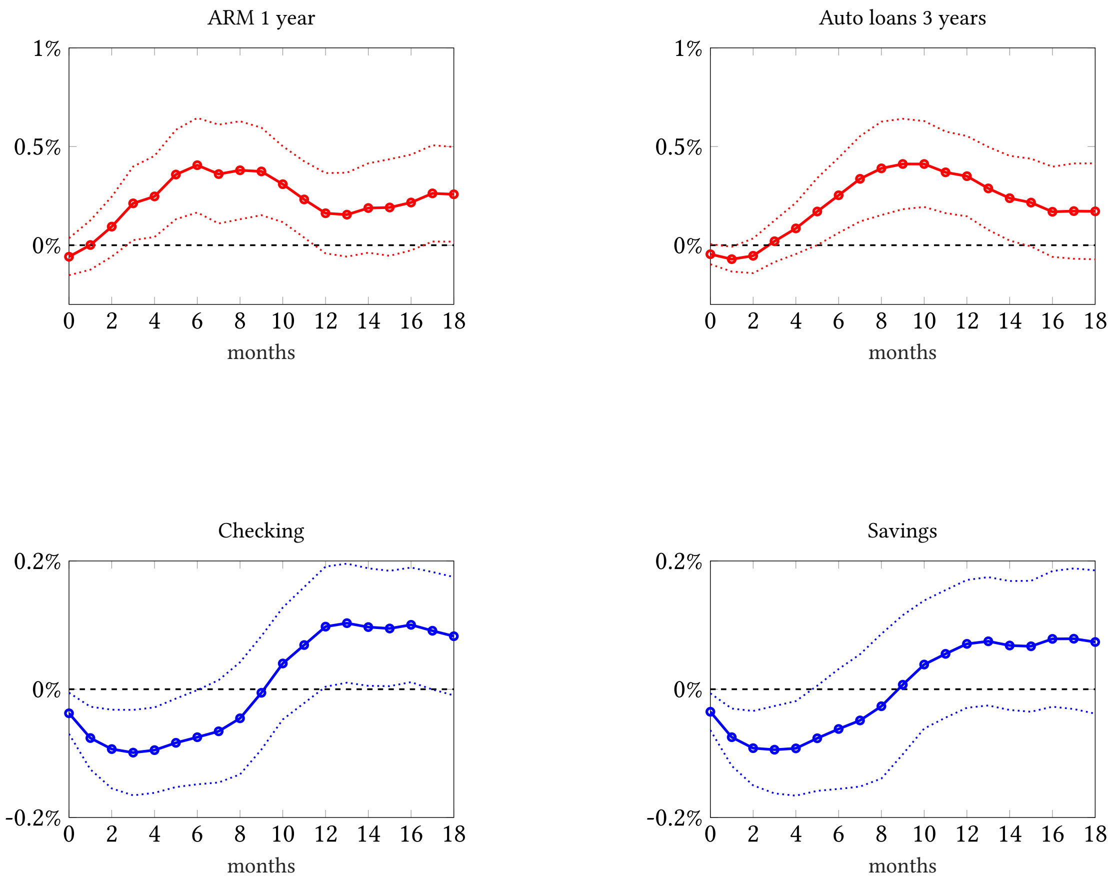
Note: Estimates $\hat{\beta}^{+}$ (for positive shocks $\Delta i_t > 0$) and $\hat{\beta}^{-}$ (for negative shocks $\Delta i_t < 0$) at horizon $h = 12$ in equation (25). Robust standard errors in parentheses, clustered at the branch level.

Table 6 shows the results. I find a significant asymmetry: the pass-through to loan and deposit rates is higher for positive monetary shocks than for negative shocks. However, estimating

$$\Delta i_t = \alpha + \gamma i_{t-1} + \epsilon_t$$

shows that condition (ii) is not satisfied: in my sample, rate hikes are *not* more likely when rates are low. This is no surprise given that rate shocks are supposed to be unexpected (were the level of rates a predictor, it would be incorporated in futures). Thus asymmetric pass-through does not drive the results of Section 2.2.

Figure 20: Additional pass-through of a raw +100 bps change in the 1-year Treasury rate i_t to deposit and loan rates when i_{t-1} is 100 bps higher.



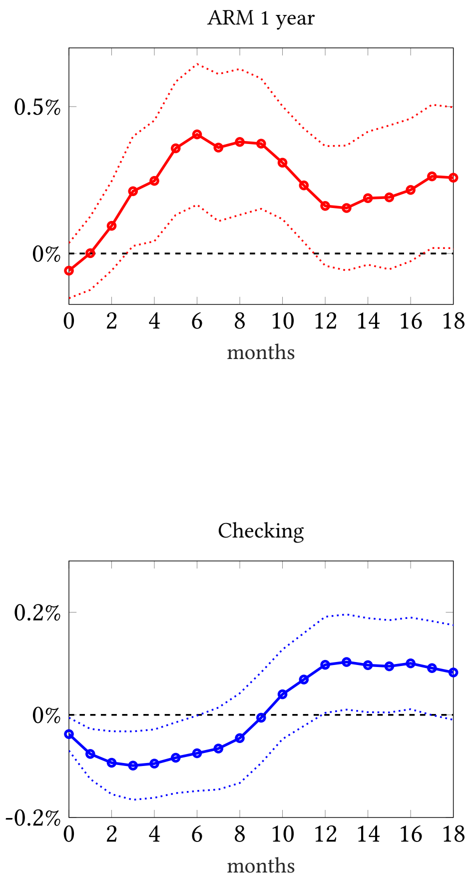
Note: The regression equations are $y_{b,t+h} - y_{b,t-1} = \alpha_{b,h} + \delta_{1h}\Delta i_t + \delta_{2h} i_{t-1} + \beta_h \Delta i_t \times i_{t-1} + \gamma_h \text{controls}_{b,t-1} + \epsilon_{b,t+h}$ for each horizon h. The figures show the sequences $\{\hat{\beta}_h\}_{h=0,\ldots,12}$ with 90% confidence bands. Standard errors are two-way clustered at the branch-month level. Sources: Federal Reserve and RateWatch.

Table 7: No asymmetry in monetary shocks.

	Δi_t
i_{t-1}	-0.00
	(0.00)
Constant	0.00
	(0.01)
R^2	0.057
Observations	155

Note: The regression equation is $\Delta i_t = \alpha + \gamma i_{t-1} + \epsilon_t$. Robust standard errors in parentheses.

Figure 21: Additional pass-through of the monetary shock from Nakamura and Steinsson (2018) to deposit and loan rates when i_{t-1} is 100 bps higher. Only using the 1998-2007 sample.



Note: The regression equations are $y_{b,t+h} - y_{b,t-1} = \alpha_{b,h} + \delta_{1h}\Delta i_t + \delta_{2h} i_{t-1} + \beta_h \Delta i_t \times i_{t-1} + \gamma_h \text{controls}_{b,t-1} + \epsilon_{b,t+h}$ for each horizon h. The figures show the sequences $\{\hat{\beta}_h\}_{h=0,\ldots,12}$ with 90% confidence bands. Standard errors are two-way clustered at the branch-month level. Sources: Federal Reserve and RateWatch.

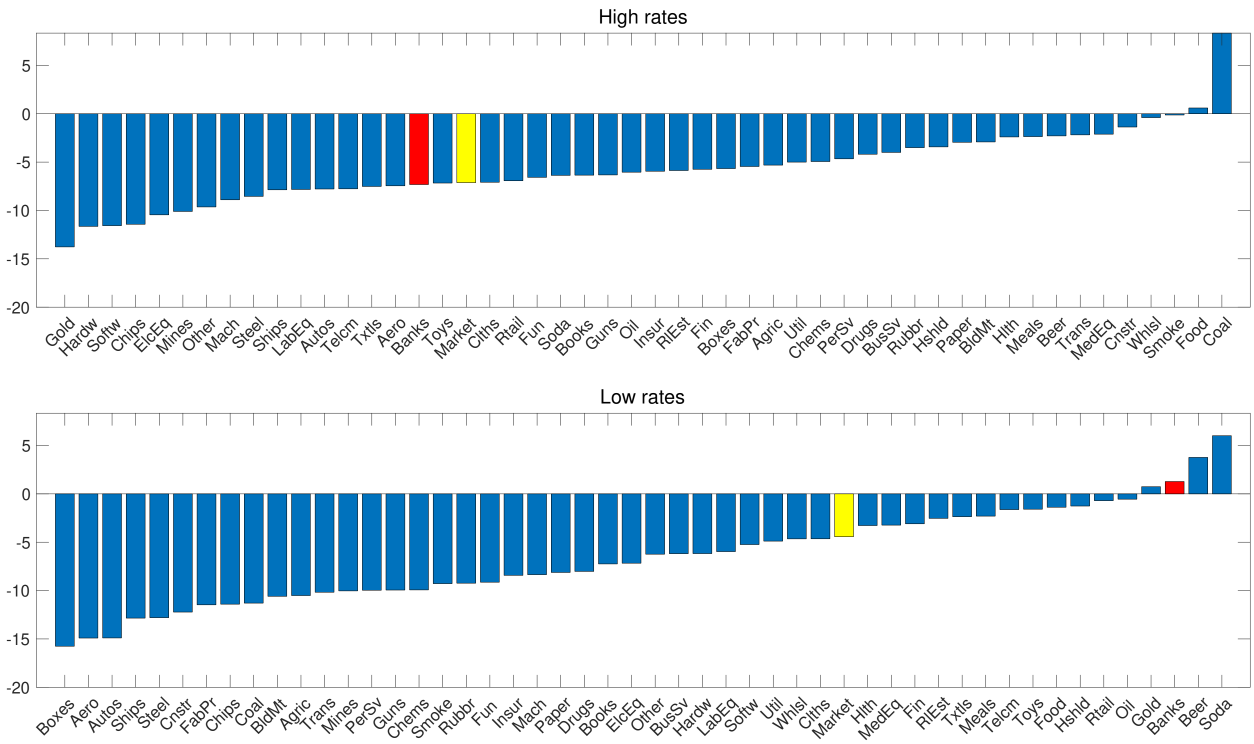
A.3 Banks vs. market portfolio

Table 8: Effect of a 100 bps monetary policy shock on Fama-French's bank industry portfolio vs. market portfolio.

	(1) Banks	(2) Banks	(3) Market	(4) Market
Δi	-4.38	11.73	-6.10*	1.61
	(5.83)	(8.66)	(3.27)	(5.31)
$\Delta i \times$ FFR		-4.85**		-2.32**
		(1.93)		(1.17)
R^2	0.011	0.051	0.048	0.069
Observations	106	106	106	106

Note: The regression equations are $y_t = \alpha + \beta\Delta i_t + \epsilon_t$ and $y_t = \alpha + \beta_1\Delta i_t + \beta_2 FFR_t + \gamma\Delta i_t \times FFR_t + \epsilon_t$. The dependent variables y are the daily return of "Banks" in the 49 Fama-French industry portfolios, and the daily return on the market portfolio, taken from Kenneth French's website. Δi_t is the high-frequency Fed funds rate shock from Nakamura and Steinsson (2018). The sample is all regularly-scheduled FOMC meetings from 01/01/2000 to 3/19/2014, excluding July 2008 to June 2009. FFR_t is the Fed funds Rate on the previous day. Robust standard errors in parentheses.

Figure 22: Response of Fama-French's 49 industry portfolios and market portfolio to a 100 bps shock to the Fed funds rate.



Note: Each bar shows the coefficient β_j in the regression $y_{j,t} = \alpha_j + \beta_j \Delta i_t + \epsilon_{j,t}$ where $y_{j,t}$ is the daily return for industry j. The bank industry portfolio is highlighted in red and the market portfolio in yellow. The top (resp. bottom) panel shows results when the Fed funds rate is above (resp. below) its median level in the sample of 1.64%.

B Appendix to Sections 3 and 4

Proof of Lemma 1. Suppose x is CES as in (9):

$$x(m,d) = \left[\alpha m^{\frac{\epsilon-1}{\epsilon}} + (1-\alpha) d^{\frac{\epsilon-1}{\epsilon}}\right]^{\frac{\epsilon}{\epsilon-1}}.$$

Then

$$\frac{x_d(f,1)}{x(f,1)} = \frac{1}{\frac{\alpha}{1-\alpha} f^{\frac{\epsilon-1}{\epsilon}} + 1},$$

which is non-increasing in f if and only if $\epsilon \geq 1$.

Suppose x is hierarchical CES as in (10):

$$x(m,d) = \left[\alpha m^{\frac{\epsilon-1}{\epsilon}} + (1-\alpha)(m+d)^{\frac{\epsilon-1}{\epsilon}}\right]^{\frac{\epsilon}{\epsilon-1}}.$$

Then

$$\frac{x_d(f,1)}{x(f,1)} = \frac{1}{\frac{\alpha}{1-\alpha} f \left(\frac{1+f}{f}\right)^{1/\epsilon} + 1 + f},$$

and the denominator is non-decreasing in f if

$$\alpha \left[\frac{1+f}{f}\right]^{1/\epsilon} \left[1 - \frac{\alpha}{\epsilon(1+f)}\right] + 1 - \alpha > 0$$

hence a sufficient condition (given $f \geq 0$) is

$$\epsilon \geq \alpha.$$

B.1 Cash-in-advance foundations for liquidity demand

In this section, I describe one potential microfoundation for the liquidity-in-utility model. Following Lucas and Stokey (1987), suppose that there are three types of consumption goods: a "credit good" c_1, a "cash good" c_2 that can only be purchased with cash and a "deposit good" c_3 that can be purchased only with deposits. Given an initial portfolio, savers maximize

$$\sum_{t\geq 0} \beta^t U(c_{1t}, c_{2t}, c_{3t})$$

subject to flow budget constraints

$$c_{1t} + c_{2t} + c_{3t} + q_t \left[\Omega_{t+1} + i_t m_{t+1} + s_t^d d_{t+1}\right] \leq \Omega_t + \gamma_t^s Y_t \tag{26}$$

where $\Omega_t = a_t + m_t + d_t$ denotes total wealth, and cash-in-advance constraints

$$c_{2t} \leq m_t \tag{27}$$

$$c_{3t} \leq d_t \tag{28}$$

As Lucas and Stokey (1987), each period should be thought of as split in two subperiods, but I use a different timing to be consistent with the convention in Section 3: in the first subperiod, the household separates into workers and shoppers subject to the cash-in-advance constraints, while in the second subperiod there is a centralized securities market for bonds, money and deposits.

The first-order conditions imply

$$\frac{U_{c_2,t+1} - U_{c_1,t+1}}{U_{c_1,t+1}} = i_t$$

$$\frac{U_{c_3,t+1} - U_{c_1,t+1}}{U_{c_1,t+1}} = s_t^d$$

Comparing to equations (6) and (7), since (27) and (27) must bind in any equilibrium with positive liquidity premia i_t and s_t^d, we can define total consumption as

$$c_t \equiv c_{1t} + c_{2t} + c_{3t}$$

and rewrite utility as[41]

$$U\left(c_t - c_{2t} - c_{3t}, c_{2t}, c_{3t}\right) = U\left(c_t - m_t - d_t, m_t, d_t\right)$$

Note that the hierarchical CES model described in 3.3 can also be nested if the third good c_3 can be purchased with money or deposits, by replacing (28) with $c_{3t} \leq m_t + d_t$.

B.2 Dynamic model

Denote the vector of state variables $\mathbf{Z}_t = \left[a_t^s, D_t, L_t^t, \ldots, L_{t+K-1}^t\right]'$ (setting $L_{t+K}^t = 0$). L_{t+k}^t captures the amount maturing at $t + k$ at the beginning of period t. Given monetary policy, at each date t we need to solve for $11 + 4K$ variables: q_t (in flexible prices equilibrium) or Y_t (in sticky prices equilibrium), together with

$$q_t^d, \left\{q_{t,t+k}^l\right\}_{k=1}^K, a_{t+1}^s, a_{t+1}^B, D_{t+1}, \left\{L_{t+1+k}^{t+1}\right\}_{k=0}^{K-1}, C_t^s, C_{t+1}^s, \left\{c_t^{y,k}, c_{t+k}^{o,k}\right\}_{k=1}^K, \lambda_t, \nu_t, \zeta_t \tag{29}$$

[41] The separable specification $u\left(c_t\right) + v\left(x\left(m_t, d_t\right)\right)$ also requires $U_{c_1c_2} = U_{c_1c_1}$ and $U_{c_1c_3} = U_{c_1c_1}$.

where L_{t+k}^{t+1} is the amount maturing at $t+k$ outstanding at the beginning of period $t+1$ and $\lambda_t, \nu_t, \zeta_t$ are Lagrange multipliers defined below. I separate the full equilibrium conditions into "blocks":

- Savers block: The three equations of "savers' block" are

$$D_{t+1} = \frac{1}{x\left(f\left(\frac{s_t^d}{i_t}\right), 1\right)} \left(v'\right)^{-1} \left(\frac{s_t^d u'\left(C_{t+1}^s\right)}{x_d\left(f\left(\frac{s_t^d}{i_t}\right), 1\right)}\right) \tag{30}$$

$$q_t^d D_{t+1} + q_t a_{t+1}^s + C_t^s = \gamma_t^s Y_t + D_t + a_t^s + (\rho - \zeta_t) E_t \tag{31}$$

$$q_t u'\left(C_t^s\right) = \beta u'\left(C_{t+1}^s\right) \tag{32}$$

where $E_t = a_t^B - D_t + \sum_{k=0}^{K-1} q_{t,t+k}^l L_{t+k}^t$. (30) is optimal deposit demand combined with deposit market clearing, (31) is savers' binding budget constraint, and (32) is the standard Euler equation. Note that even when savers only hold deposits in equilibrium ($a_{t+1}^s = 0$), their marginal rate of intertemporal substitution is the bond rate R_t, not the deposit rate R_t^d, because they are not constrained to save in deposits. Thus in the short-run analysis, incomplete deposit pass-through does not mechanically translate into incomplete pass-through of monetary policy to savers' consumption. In fact, allowing for unconstrained substitution between bonds and deposits is crucial to *generate* the incomplete deposit pass-through.

- Borrowers block: for each type k we have

$$q_{t,t+k}^l u_b'\left(c_t^{y,k}\right) = \beta^k u_b'\left(c_{t+k}^{o,k}\right) \tag{33}$$

$$c_{t+k}^{o,k} = \gamma_{t+k}^{o,k} Y_{t+k} - \frac{c_t^{y,k} - \gamma_t^{y,k} Y_t}{q_{t,t+k}^l} \tag{34}$$

(33) is type k borrowers' Euler equation and (34) is their binding budget constraint.

- Bank block:

$$q_t^d D_{t+1} = \overline{\phi}_t \left(1 - \rho + \zeta_t\right) E_t \tag{35}$$

$$q_t a_{t+1}^B + \sum_{k=1}^{K} q_{t,t+k}^l L_{t+k}^{t+1} = \left(1 - \rho + \zeta_t\right) E_t + q_t^d D_{t+1} \tag{36}$$

(35) is banks' binding limited pledgeability constraint and (36) is their binding budget constraint. Calling λ_t, ν_t, and ξ_t the Lagrange multiplier on, respectively, the budget

constraint, the leverage constraint, and the $a^B_{t+1} \geq 0$ constraint, the optimality conditions of banks are

$$\lambda_t q^d_t = \nu_t q^d_t + q_t \left\{\rho + (1-\rho)\left(\lambda_{t+1} + \overline{\phi}_{t+1}\nu_{t+1}\right)\right\} \tag{37}$$

$$\lambda_t q^l_{t,t+k} = q^l_{t+1,t+k} q_t \left\{\rho + (1-\rho)\left(\lambda_{t+1} + \overline{\phi}_{t+1}\nu_{t+1}\right)\right\} \quad \forall k \geq 1 \tag{38}$$

$$q_t - q^l_{t,t+1} = \frac{\xi_t}{\lambda_t} \tag{39}$$

$$a^B_{t+1} \geq 0 \tag{40}$$

$$a^B_{t+1}\left(q_t - q^l_{t,t+1}\right) = 0 \tag{41}$$

Lending is unconstrained at t if $\xi_t = 0$.

- Bond market clearing

$$a^s_{t+1} + a^B_{t+1} = 0 \tag{42}$$

- Loan market clearing

$$\forall k = 1, \ldots, K \quad L^{t+1}_{t+k} - L^t_{t+k} = \mu_k\left(\gamma^{o,k}_{t+k} Y_{t+k} - c^{o,k}_{t+k}\right) \tag{43}$$

- Goods market clearing

$$Y_t = C^s_t + C^B_t + \sum_{k=1}^{K} \mu_k c^{y,k}_t + \left[\sum_{k=1}^{K} \mu_k \gamma^{o,k}_t Y_t - L^t_t\right] \tag{44}$$

Given a guess for λ_t, there are $4K+11$ equations, hence we can solve for all variables (including λ_{t+1}) in terms of λ_t. Given a sequence $\left\{q^m_t\right\}$ that converges to some q^m_∞, we can solve for the steady state and then search for λ_0 such that iterating forward leads λ_t to converge to its steady state value.

Special case: only short-term loans. With $K > 1$ the solution depends on the path of monetary policy $\{i_{t+s}\}_{s\geq 0}$ in its entirety, so there is no finite recursive structure (i.e., one state variable is the infinite-dimensional object $\{i_{t+s}\}_{s\geq 0}$). However, in the special case of only short-term loans $K = 1$, we can remove the $K+1$ equations (37)-(38) and the variables (λ_t, ν_t) to obtain a system of $10 + 3K$ equations in $9 + 4K$ unknowns. Thus with $K = 1$ the system is recursive as follows.

Summing savers' and banks' budget constraints

$$q_t^d D_{t+1} + q_t a_{t+1}^s + C_t^s = \gamma_t^s Y_t + D_t + a_t^s + \rho \left(E_t - \iota_t\right)$$
$$q_t a_{t+1}^B + q_t^l L_{t+1}^{t+1} = \left(1-\rho\right) E_t + q_t^d D_{t+1}$$

we get

$$C_t^s = \gamma_t^s Y_t + L_t^t - q_t^l L_{t+1}^{t+1}.$$

The key simplification when all loans are short-term is that (36) is sufficient to pin down q_t^l:

$$q_t a_{t+1}^B + q_t^l L_{t+1}^{t+1}\left(q_t^l, Y_{t+1}\right) = \left(1-\rho\right)\left(1+\overline{\phi}_t\right) E_t.$$

If lending is constrained, $a_{t+1}^B = 0$ and this gives us q_t^l. If lending is unconstrained, $q_t^l = q_t$ and this gives us a_{t+1}^B. Future output matters because loan demand L_{t+1}^{t+1}, determined by (43)-(34), is forward-looking. Note that conditional on Y_{t+1}, only current monetary policy i_t matters, but in equilibrium the whole path of monetary policy $\{i_{t+s}\}_{s\geq 0}$ matters to determine the sequence $\{Y_{t+s}\}_{s\geq 0}$, just as in the standard New Keynesian model. $\overline{\phi}_t$ depends not only on R_t^l but also on R_t^d, which can be obtained from deposit market clearing. Then equilibrium simply requires that we find policy functions

$$q_t^l\left(R_t, Y_{t+1}, \mathbf{Z}_t\right), q_t^d\left(R_t, Y_{t+1}, \mathbf{Z}_t\right)$$

that jointly solve

$$u'\left(\gamma_t^s Y_t + L_t^t - q_t^l L_{t+1}^{t+1}\right) = \beta R_t u'\left(\gamma_{t+1}^s Y_{t+1} + L_{t+1}^{t+1} - q_{t+1}^l L_{t+2}^{t+2}\right),$$
$$\overline{\phi}_t E_t = \frac{q_t^d}{x\left(f\left(\frac{s_t^d}{i_t}\right),1\right)}\left(v'\right)^{-1}\left(\frac{s_t^d u'\left(\gamma_{t+1}^s Y_{t+1} + L_{t+1}^{t+1} - q_{t+1}^l L_{t+2}^{t+2}\right)}{x_d\left(f\left(\frac{s_t^d}{i_t}\right),1\right)}\right).$$

Therefore we obtain a modified dynamic IS curve

$$Y_t = F\left(Y_{t+1}, R_t, \mathbf{Z}_t\right)$$

that can be solved either for Y_t given R_t (in sticky prices equilibrium) or for the natural rate R_t^n given $Y_t = Y_t^*$ (in flexible prices equilibrium), together with a law of motion for the state variable

$$\mathbf{Z}_{t+1}\left(Y_{t+1}, R_t, \mathbf{Z}_t\right).$$

Proof of Proposition 1. Combining banks' leverage and budget constraints (35)-(36), we have

$$q_t a^B_{t+1} + \sum_{k=1}^{K} q^l_{t,t+k} L^{t+1}_{t+k} = \left(1 + \overline{\phi}_t\right)(1 - \rho + \zeta_t) E_t. \tag{45}$$

Combining banks' budget constraint (36) and savers' budget constraint (31), we have, after simplifying with bond-market clearing (42),

$$\begin{aligned}
\sum_{k=1}^{K} q^l_{t,t+k} L^{t+1}_{t+k} + C^s_t &= \gamma^s_t Y_t + a^s_t + D_t + E_t \\
&= \gamma^s_t Y_t + a^s_t + a^B_t + \sum_{k=0}^{K-1} q^l_{t,t+k} L^t_{t+k} \\
\sum_{k=1}^{K} q^l_{t,t+k} L^{t+1}_{t+k} + C^s_t &= \gamma^s_t Y_t + \sum_{k=0}^{K-1} q^l_{t,t+k} L^t_{t+k},
\end{aligned} \tag{46}$$

where the second line follows from the definition of E_t and the third line from the previous date's bond-market clearing.

Consider now an equilibrium such that $q^l_t = q_t$ for all t (for instance around a steady state with unconstrained lending). Then the system is block-recursive, because we do not need to solve for bank capital E_t. We can replace (36) with the consolidated budget constraint (46) and solve for

$$q_t, a^s_{t+1}, a^B_{t+1}, \left\{L^{t+1}_{t+1+k}\right\}_{k=0}^{K-1}, C^s_t, C^s_{t+1}, \left\{c^{y,k}_t, c^{o,k}_{t+k}\right\}_{k=1}^{K}$$

without any reference to the deposit market, and then back out deposit prices and quantities from (30) and (35). In an equilibrium with constrained lending, we must instead keep track separately of capital E_t and (36) to solve for loan prices q^l_t.

B.3 Steady state

Let $\mathcal{L}\left(R^l\right)$ be the value of the end-of-period stock of loans across all maturities for a given steady state short-term loan rate R^l

$$\mathcal{L}\left(R^l\right) = \sum_{k=1}^{K} \left(R^l\right)^{-k} \left[\sum_{j=k}^{K-k} \frac{\mu_j x_j}{G^{j-k}}\right],$$

where each x_k solves type k borrowers' Euler equation

$$u'\left(\gamma^{y,k} + \left(R^l\right)^{-k} x_k\right) = \left(\beta R^l\right)^k u'\left(\gamma^{o,k} G^k - x_k\right).$$

$\mathcal{L}$ is a decreasing function of R^l if $\sigma \geq 1$, as assumed in the text.

Unconstrained steady state. All variables without time subscripts are normalized by GDP, i.e., $x = \frac{x_t}{Y_t}$ (thus for assets they denote holdings in the beginning of a period). In the unconstrained regime there are 7 equations in 7 unknowns $c^s, D, E, \overline{\phi}, a^s, q, q^d$

$$
\begin{aligned}
c^s + \left(\frac{G}{R^d} - 1\right) D &= \left(1 - \frac{G}{R^*}\right) a^s + \gamma^s + (\rho - \zeta) E \\
D &= \frac{x_d\left(f\left(\frac{s^d}{i}\right), 1\right)}{s^d x\left(f\left(\frac{s^d}{i}\right), 1\right)} \chi c^s \\
R^* &= G/\beta \\
\frac{GD}{R^d} &= \overline{\phi}\,(1 - \rho + \zeta)\, E \\
\mathcal{L}\left(R^*\right) &= \left(1 + \overline{\phi}\right)(1 - \rho + \zeta)\, E + \frac{Ga^s}{R^*} \\
\overline{\phi} &= \frac{\theta R^*/R^d}{1 - \theta R^*/R^d}
\end{aligned}
$$

and the key ROE equation

$$
\frac{G}{1 - \rho + \zeta} = \left(1 + \overline{\phi}\right) R^* - \overline{\phi} R^d.
$$

Constrained steady state. In the constrained regime there are 7 equations in 7 unknowns $c^s, D, E, \overline{\phi}, q^l, q, q^d$. First, in the loan-market clearing condition

$$
\frac{\mathcal{L}\left(R^l\right)}{1 + \overline{\phi}} = (1 - \rho + \zeta)\, E,
$$

the left-hand side decreases with R^l so this defines a decreasing function $R^l(E)$. The remaining equations are

$$
\begin{aligned}
c^s + \left(\frac{G}{R^d} - 1\right) D &= y^s + (\rho - \zeta) E \\
D &= \frac{x_d\left(f\left(\frac{s^d}{i}\right), 1\right)}{s^d x\left(f\left(\frac{s^d}{i}\right), 1\right)} \chi c^s \\
R &= G/\beta \\
D &= \theta\left(1 + \overline{\phi}\right)(1 - \rho + \zeta) E \\
\overline{\phi} &= \frac{\theta R^l/R^d}{1 - \theta R^l/R^d}.
\end{aligned}
$$

Aggregating yields

$$
c^s = y^s - \mathcal{L}\left(R^l\right).
$$

From deposit market clearing we have

$$
\frac{x_d\left(f\left(\frac{s^d}{i}\right), 1\right)}{x\left(f\left(\frac{s^d}{i}\right), 1\right)} = s^d \times \frac{E}{y^s - \mathcal{L}\left(R^l(E)\right)} \times \frac{\theta\left(1 + \overline{\phi}\right)(1 - \rho + \zeta)}{\chi} \tag{47}
$$

With those equations in hand we can prove the propositions in Section 4.

Proof of Propositions 2 and 3. (47) gives us a solution

$$
s^d(i, E)
$$

The right-hand side of (47) is increasing in s^d and in E: as E increases, $R^l(E)$ decreases so $\mathcal{L}$ increases. $s^d(i, E)$ is non-decreasing in i. To see this, recall that we always have, from the strict concavity of x, that f is increasing. Then, from Assumption 5, the left-hand side of (47) is a decreasing function of s^d, and a higher nominal rate i decreases $f\left(\frac{s^d}{i}\right)$, which shifts the left-hand side of (47) up. The solution $s^d(i, E)$ must then increase. As a result, the deposit spread τ^d also increases. Going back to (14), we then have two possibilities:

- suppose that in equilibrium $\overline{\phi}$ increases with i: then $\overline{\phi}\tau^d$ increases with i, and $\left(1 + \overline{\phi}\right)\tau^l$ must decrease, hence τ^l decreases;
- suppose that in equilibrium $\overline{\phi}$ decreases with i. This means that $\frac{1+\tau^l}{1-\tau^d}$ must fall, hence $d\tau^l + d\tau^d < d\tau^l + \frac{1+\tau^l}{1-\tau^d} d\tau^d < 0$. Since $\overline{\phi}\left(\tau^l + \tau^d\right) + \tau^l$ must be constant, τ^l must increase with i, which contradicts the fact that $\overline{\phi}$ decreases.

Thus it must be that following an increase in the steady state nominal rate i, $\overline{\phi}$ increases and τ^l increases. The same proof applies for Proposition 3, as $\frac{\text{ROE}}{R^*} = \frac{\beta}{1-\rho+\zeta}$ is constant and aggregate loan demand $\mathcal{L}$ scales with $1/R^*$ holding τ^l fixed.

A knife-edge case happens when $f \mapsto \frac{x_d(f,1)}{x(f,1)}$ is constant: the nominal rate i is then irrelevant for s^d conditional on E. I now show that this only arises in the case of Cobb-Douglas liquidity x. For superneutrality in both regimes, we need $\frac{x_d\left(f\left(\frac{s^d}{i}\right),1\right)}{x\left(f\left(\frac{i^d}{i}\right),1\right)}$ to be constant in the inflation target π. Rewriting $u = f\left(\frac{s^d}{i}\right)$ we have

$$x_d\left(1, \frac{1}{u}\right) = \text{constant} \times ux\left(1, \frac{1}{u}\right),$$

or, denoting $z = 1/u$ and $\varphi(z) = x(1, z)$,

$$z\varphi'(z) = \text{constant} \times \varphi(z).$$

The solution of this differential equation is

$$\log \varphi(z) = A + B \times \log z,$$

for some constants A and B, which implies Cobb-Douglas liquidity

$$x(1, z) = e^A z^B,$$

as x is homogeneous of degree 1.

C Appendix to Section 5

Unconstrained lending. In the unconstrained lending regime, the equilibrium is given by 9 equations in 9 unknowns $\left\{C_0^s, C_0^b, Y_0, C_1^s, C_1^b, q_0^d, D_1, L_1, A_1\right\}$:

$$
\begin{aligned}
Y_0 &= C_0^s + C_0^b \\
Y_1^* &= C_1^s + C_1^b \\
C_0^s + \left(q_0^d - q_0\right) D_1 + q_0 C_1^s &= \gamma^s Y_0 - E_0 + q_0 \left(A_1 + L_1 - D_1\right) \\
\beta C_0^s &= q_0 C_1^s \\
D_1 &= \theta \left(A_1 + L_1\right) \\
C_0^b + q_0 C_1^b &= \gamma^b Y_0 + q_0 Y_1^* \\
\beta C_0^b &= q_0 C_1^b \\
\frac{D_1}{C_1^s} &= \frac{\chi x_d \left(f\left(\frac{s_0^d}{i_0}\right), 1\right)}{s_0^d x \left(f\left(\frac{s_0^d}{i_0}\right), 1\right)} \\
-q_0 A_1 &= \gamma^s Y_0 - E_0 - C_0^s - q_0^d D_1.
\end{aligned}
$$

Constrained lending. In the constrained lending regime, the equilibrium is given by 9 equations in 9 unknowns $\left\{C_0^s, C_0^b, Y_0, C_1^s, C_1^b, q_0^d, q_0^l, D_1, L_1\right\}$ (relative to the unconstrained lending regime, q_0^l replaces A_1):

$$
\begin{aligned}
Y_0 &= C_0^s + C_0^b \\
Y_1^* &= C_1^s + C_1^b \\
C_0^s + \left(q_0^d - q_0\right) D_1 + q_0 C_1^s &= \gamma^s Y_0 - E_0 + q_0 \left(L_1 - D_1\right) \\
C_0^s + E_0 + q_0^d D_1 &= \gamma^s Y_0 \\
\beta C_0^s &= q_0 C_1^s \\
D_1 &= \theta L_1 \\
C_0^b + q_0^l C_1^b &= \gamma^b Y_0 + q_0^l Y_1^* \\
\beta C_0^b &= q_0^l C_1^b \\
\frac{D_1}{C_1^s} &= \frac{\chi x_d \left(f\left(\frac{s_0^d}{i_0}\right), 1\right)}{s_0^d x \left(f\left(\frac{s_0^d}{i_0}\right), 1\right)}.
\end{aligned}
$$

We also have

$$C_0^s + q_0 C_1^s = \gamma^s Y_0 + \left(q_0 - q_0^l\right) L_1 = \gamma^s Y_0 + \left(1 - \frac{q_0^l}{q_0}\right) q_0 C_1^s$$

$$C_0^s = \frac{\gamma^s Y_0}{1 + \beta - \beta\left(1 - \frac{R_0}{R_0^l}\right)} = \frac{\gamma^s Y_0}{1 + \beta - \frac{\beta \tau_0^l}{1+\tau_0^l}}.$$

Since $R_0^l > R_0$, savers' share of consumption in the constrained lending regime is higher because they earn the credit spread paid by borrowers.

Proof of Propositions 8 and 9. Borrowers' consumption is

$$C_0^b = \frac{\gamma^b Y_0 + Y_1^* / R_0^l}{1 + \beta}$$

therefore

$$\begin{aligned} Y_0 &= \frac{\gamma^b Y_0 + Y_1^* / R_0^l}{1 + \beta} + \frac{\gamma^s Y_0}{1 + \beta - \frac{\beta \tau_0^l}{1+\tau_0^l}} \\ &= \frac{Y_1 / R_0^l}{1 + \beta - \gamma^b - \gamma^s \frac{1+\beta}{1+\beta - \frac{\beta \tau_0^l}{1+\tau_0^l}}} \\ Y_0 &= \frac{Y_1 / R_0^l}{\beta - \gamma^s \left[\frac{\frac{\beta \tau_0^l}{1+\tau_0^l}}{1+\beta - \frac{\beta \tau_0^l}{1+\tau_0^l}}\right]}. \end{aligned}$$

Therefore, Γ in Proposition 8 is given by:

$$\Gamma = \frac{1}{1 + \tau_0^l \left[1 - \gamma^s \beta \left(\frac{1+\tau_0^l}{1+\beta+\tau_0^l}\right)\right]}.$$

Γ is decreasing in τ_0^l because $\frac{d \log\left(1+\tau_0^l\left[1-\gamma^s\beta\left(\frac{1+\tau_0^l}{1+\beta+\tau_0^l}\right)\right]\right)}{d \log\left(1+\tau_0^l\right)}$ equals $1 - \gamma^s$ at $\tau_0^l = 0$, goes to 1 as $\tau_0^l \to \infty$ and has a global minimum at $1 + \tau_0^l = \sqrt{\frac{\beta(\beta+\gamma^s)}{1-\gamma^s}}$ with value $\frac{2\left[\sqrt{\beta(1-\gamma^s)(\beta+\gamma^s)} - \beta(1-\gamma^s)\right]}{(1+\beta)\gamma^s} > 0$.

Therefore the elasticity of the denominator to $1+\tau_0^l$

$$\epsilon^{\Gamma} = \frac{d\log\left(1+\tau_0^l\left[1-\gamma^s\beta\left(\frac{1+\tau_0^l}{1+\beta+\tau_0^l}\right)\right]\right)}{d\log\left(1+\tau_0^l\right)}$$

lies between 0 and 1.

Through the loan-market clearing condition

$$\frac{E_0}{1-\theta\frac{R_0^l}{R_0^d}} = \frac{Y_1^*/R_0^l - \beta\gamma^b Y_0}{1+\beta} \tag{48}$$

we can reexpress loan pass-through as a function of deposit pass-through

$$\eta^l = A \times \left[B\eta^d + (1-B)\right] \tag{49}$$

where A and B are between 0 and 1, given by

$$A = \frac{\frac{(1-\epsilon^{\Gamma})\beta\gamma^b Y_0}{Y_1^*/R_0^l-\beta\gamma^b Y_0} + \overline{\phi}}{1 + \frac{(1-\epsilon^{\Gamma})\beta\gamma^b Y_0}{Y_1^*/R_0^l-\beta\gamma^b Y_0} + \overline{\phi}}$$

$$B = \frac{\overline{\phi}}{\frac{(1-\epsilon^{\Gamma})\beta\gamma^b Y_0}{Y_1^*/R_0^l-\beta\gamma^b Y_0} + \overline{\phi}}.$$

Two forces determine loan pass-through in this model. In equation (49), the first term $AB\eta^d$ reflects *credit supply.* It is proportional to deposit pass-through, and works through banks' balance sheets that tie deposit and loan markets together. A higher deposit pass-through implies that following a rate hike, banks' leverage constraint tightens by more, so credit supply falls by more. The second term $A(1-B)$ reflects a "consumption smoothing" channel of *credit demand*, present even with zero deposit pass-through—for instance in the limit of perfect substitutability between money and deposits. It is proportional to γ^b, the share of income accruing to borrowers. Even if the loan rate doesn't move, so that no substitution effect is triggered, borrowers are exposed to monetary policy because a monetary shock is transmitted through unconstrained savers to output and thus to borrowers' income. For a given loan rate, credit demand increases when borrower's income falls because they are trying to smooth consumption. Thus, fixing credit supply, an increase in the loan rate is required to clear the loan market, which gives a second source of loan pass-through. The same consumption smoothing channel is present in any heterogeneous agents New Keynesian model: unconstrained borrowers will demand more credit if their current income falls.

Remark 1 (Credit cyclicality)**.** *The consumption smoothing channel of credit demand does not make credit (counterfactually) countercyclical. First, my reasoning concerns credit demand* conditional on the loan rate*: in equilibrium, total borrowing might still be procyclical if the loan rate is countercyclical. And indeed, in the Modigliani-Miller benchmark, equilibrium credit is proportional to output Y_0, but credit demand is decreasing in Y_0 conditional on $R_0^l = R_0$. Second, I only consider exogenous monetary policy shocks here, instead of a monetary policy rule that seeks to stimulate the economy when credit demand plummets due to non-monetary factors.*

Differentiating output in the constrained lending regime (20) yields the local output sensitivity

$$\eta^Y = 1 - \epsilon^\Gamma \times \left(1 - \eta^l\right). \tag{50}$$

The output effect of a monetary shock departs from the RANK benchmark if and only if the pass-through to loan rates is not 1. If the pass-through is higher than 1, then the output effect is amplified. If the pass-through is incomplete, i.e., $\eta^l < 1$, then output sensitivity is dampened. I show below that the latter case holds.[42] Expression (50) makes apparent that a monetary shock in the presence of credit frictions can be viewed as the superposition of a pure monetary shock in a frictionless model, giving sensitivity 1 (or more generally sensitivity equal to the elasticity of intertemporal substitution), and a pure financial "deleveraging" shock. With $\eta^l < 1$, the higher the elasticity θ^Γ, the more reactive the economy is to financial shocks not caused by monetary policy, and the more dampening there will be in reaction to monetary shocks.

Combining (50) and (49) yields output sensitivity as a function of deposit pass-through

$$\eta^Y = 1 - \epsilon^\Gamma \times \left(1 - A + AB\left(1 - \eta^d\right)\right). \tag{51}$$

In particular, η^Y is bounded by

$$0 \leq 1 - \frac{\left(1 + \overline{\phi}\right)\Gamma}{1 + \overline{\phi} + \frac{\left(1-\eta^\Gamma\right)\beta\gamma^b Y_0}{Y_1^*/R_0^l - \beta\gamma^b Y_0}} \leq \eta^Y \leq 1 - \frac{\Gamma}{1 + \overline{\phi} + \frac{\left(1-\eta^\Gamma\right)\beta\gamma^b Y_0}{Y_1^*/R_0^l - \beta\gamma^b Y_0}} \leq 1.$$

[42] In the context of firm borrowing, the evidence is mixed. Gertler and Karadi (2015) find, using a structural VAR with high frequency monetary policy shocks as external instruments, that the excess bond premium computed in Gilchrist and Zakrajsek (2012) increases in reaction to a contractionary monetary shock, which suggests amplification. But Ramey (2016), using the same data, finds the opposite effect when using Jordá local projections instead of a fully specified VAR.

Since η^d is between 0 and 1 (see below), the loan pass-through η^l is always lower than 1, which shows that monetary policy shocks are dampened relative to a frictionless benchmark.

Proof of Lemma 2. The deposit market clearing condition

$$D_1 = \theta L_1 = \theta C_1^s$$

rewrites

$$\frac{x_d\left(f\left(\frac{s_0^d}{i_0}\right),1\right)}{x\left(f\left(\frac{s_0^d}{i_0}\right),1\right)} = \frac{\theta}{\chi}s_0^d. \tag{52}$$

Thus we can write the price of deposit liquidity s_0^d as a function $g(i_0)$ of the nominal interest rate with derivative

$$g' = \frac{g\left[\frac{x_m}{x} - \frac{x_{md}}{x_d}\right]}{\left[\frac{x_m}{x} - \frac{x_{md}}{x_d}\right] i_0 + \frac{(i_0)^2}{gf'}}$$

which is non-negative by Assumption 5. Since $s_0^d \geq 0$, the pass-through to deposit rates is positive but lower than 1, as

$$\eta^d = 1 - \left(1 + s_0^d\right) g'$$

The cost of deposit liquidity s_0^d is then monotonically related to the deposit spread τ_0^d as $\tau_0^d = 1 - \frac{1}{1+s_0^d}$. Denote

$$G(z) = \frac{x_d(f(z),1)}{x(f(z),1)}$$

where $f(z)$ is the solution to $\frac{x_d(f,1)}{x_m(f,1)} = z$. Applying the implicit function theorem twice to (52), we have

$$\frac{d^2 s_0^d}{di_0^2} = \frac{s_0^d G''\left(\frac{s_0^d}{i_0}\right)\left[\frac{1}{i_0}\frac{ds_0^d}{di_0} - \frac{s_0^d}{i_0}\right]^2 + 2s_0^d G'\left(\frac{s_0^d}{i_0}\right)\left[\frac{s_0^d}{i_0^3} - \frac{1}{i_0^2}\frac{ds_0^d}{di_0}\right]}{G\left(\frac{s_0^d}{i_0}\right) - \frac{s_0^d}{i_0}G'\left(\frac{s_0^d}{i_0}\right)}$$

From Assumption 5 we have $G' \leq 0$ and $\frac{ds_0^d}{di_0} \leq \frac{s_0^d}{i_0}$ hence condition (21) is sufficient to ensure $\frac{d^2 s_0^d}{di_0^2} \leq 0$.

Proof of Proposition 10. In response to a contractionary monetary shock $\Delta i_0 > 0$, ROE falls if and only if

$$\frac{\overline{\phi}}{1+\overline{\phi}}\eta^d < \eta^l = \frac{\overline{\phi}_0}{1 + \frac{(1-a)\beta\gamma^b Y_0}{Y_1^*/R_0^l - \beta\gamma^b Y_0} + \overline{\phi}_0}\eta^d + \frac{\frac{(1-a)\beta\gamma^b Y_0}{Y_1^*/R_0^l - \beta\gamma^b Y_0}}{1 + \frac{(1-a)\beta\gamma^b Y_0}{Y_1^*/R_0^l - \beta\gamma^b Y_0} + \overline{\phi}_0}$$

or

$$\eta^d < \frac{\left(1-\eta^\Gamma\right)\beta\gamma^b Y_0}{Y_1^*/R_0^l - \beta\gamma^b Y_0} \times \frac{1}{\overline{\phi}\left[\frac{1}{1+\overline{\phi}}\left(1 + \frac{(1-\eta^\Gamma)\beta\gamma^b Y_0}{Y_1^*/R_0^l - \beta\gamma^b Y_0} + \overline{\phi}\right) - 1\right]}$$

$$\overline{\phi}\eta^d < \eta^l\left(1+\overline{\phi}\right) = A \times \left[B\eta^d + (1-B)\right]\left(1+\overline{\phi}\right)$$

$$\eta^d > \frac{A\left(1+\overline{\phi}\right) - AB\left(1+\overline{\phi}\right)}{\overline{\phi} - AB\left(1+\overline{\phi}\right)}, \tag{53}$$

while bank lending falls if and only if the loan-deposit spread R_0^l/R_0^d does, i.e., if and only if

$$\eta^l - \eta^d < 0.$$

The pass-through to the loan-deposit spread is

$$\eta^l - \eta^d = \frac{\frac{(1-a)\beta\gamma^b Y_0}{Y_1^*/R_0^l - \beta\gamma^b Y_0} - \left(1 + \frac{(1-a)\beta\gamma^b Y_0}{Y_1^*/R_0^l - \beta\gamma^b Y_0}\right)\eta^d}{1 + \frac{(1-a)\beta\gamma^b Y_0}{Y_1^*/R_0^l - \beta\gamma^b Y_0} + \overline{\phi}}.$$

Hence lending falls when the pass-through to deposit rates is high enough, that is when

$$\eta^d > \frac{A - AB}{1 - AB}. \tag{54}$$

Finally, the right-hand sides of (53) and (54) are themselves decreasing in in R_0, which implies Proposition 10.

D Extensions

D.1 Market power vs. financial constraints

In this section, I compare the predictions of a model of bank market power over deposits and those of the model with financially constrained banks. Both views are complementary, and heterogeneity in bank market power is undeniably important to explain cross-sectional patterns in deposit spreads and their reactions to monetary policy, as demonstrated by Drechsler et al. (2017). The goal of this section is to understand theoretically which features of deposit and loan supply deliver the incomplete pass-through we observe in the data.

I use time subscripts just to keep notation consistent with the rest of the paper, but this section only concerns the static portfolio choice of savers. Consider a more general version of Section 3.3. A saver maximizes at date t

$$U\left(c_{t+1}, x\left(m_{t+1}, d_{t+1}\right)\right)$$

subject to

$$\begin{aligned} q_t a_{t+1} + q_t^m m_{t+1} + q_t^d d_{t+1} &\leq W_t \\ c_{t+1} &= a_{t+1} + m_{t+1} + d_{t+1} \end{aligned}$$

where W_t is wealth. The first-order conditions with respect to m_{t+1} and d_{t+1} are

$$\begin{aligned} \frac{U_{x,t+1} x_{d,t+1}}{U_{c,t+1}} &= \frac{q_t^d - q_t}{q_t} \equiv s_t^d \\ \frac{U_{x,t+1} x_{m,t+1}}{U_{c,t+1}} &= \frac{q_t^m - q_t}{q_t} \equiv i_t \end{aligned} \tag{55}$$

As in Section 3.3, if x is homothetic and concave we get

$$\begin{aligned} \frac{x_d\left(m_{t+1}, d_{t+1}\right)}{x_m\left(m_{t+1}, d_{t+1}\right)} &= \frac{s_t^d}{i_t} \equiv z_t \\ \Rightarrow \frac{m_{t+1}}{d_{t+1}} &= f\left(z_t\right) \end{aligned}$$

where f is increasing. Plugging back into (55):

$$\frac{U_x\left(c_{t+1}, d_{t+1} \cdot x\left(f\left(z_t\right), 1\right)\right)}{U_c\left(c_{t+1}, d_{t+1} \cdot x\left(f\left(z_t\right), 1\right)\right)} x_d\left(f\left(z_t\right), 1\right) = z_t i_t \tag{56}$$

defines a deposit demand function

$$\mathcal{D}_{t+1}\left(z_t, i_t, c^s_{t+1}\right)$$

Suppose now that a monopolist bank maximizes revenue taking c^s_{t+1} as given. There are several ways to justify why the bank takes c^s_{t+1} as given. First, this case is isomorphic to a model in which the representative savers has utility over an aggregate of deposits from a large number of horizontally differentiated banks b in monopolistic competition. Second, under a Greenwood et al. (1988) specification for U

$$U(c,x) = \tilde{U}(c + v(x)) \tag{57}$$

deposit demand is independent of consumption c^s_{t+1}; this justification is less appealing once we think of microfoundations for liquidity in the utility (e.g., Appendix B.1). The third route is to follow Drechsler et al. (2017) by taking the limit of vanishing liquidity services $U_x \longrightarrow 0$.

The bank maximizes revenue

$$\max_{s^d_t} s^d_t \mathcal{D}\left(\frac{s^d_t}{i_t}, i_t, c^s_{t+1}\right) \Leftrightarrow \max_{z_t} \; z_t \mathcal{D}\left(z_t, i_t, c^s_{t+1}\right)$$

and thus sets

$$\frac{\partial \log \mathcal{D}_{t+1}}{\partial \log z_t} = -1 \tag{58}$$

Denote

$$G(x,c) = \frac{U_x(c,x)}{U_c(c,x)}$$

Differentiating (56), we get the elasticity

$$\frac{\partial \log \mathcal{D}_{t+1}}{\partial \log z_t} = \frac{1 - \frac{\mathcal{D}_{t+1}(z_t,i_t)x(f(z_t),1)G'(\mathcal{D}_{t+1}(z_t,i_t)x(f(z_t),1))}{G(\mathcal{D}_{t+1}(z_t,i_t)x(f(z_t),1))} - \frac{\partial \log x_d(f(z_t),1)}{\partial \log z_t}}{\frac{\mathcal{D}_{t+1}(z_t,i_t)x(f(z_t),1)G'(\mathcal{D}_{t+1}(z_t,i_t)x(f(z_t),1))}{G(\mathcal{D}_{t+1}(z_t,i_t)x(f(z_t),1))} z_t f'(z_t)\, x_m(f(z_t),1)}$$

Hence if $\frac{xG'(x)}{G(x)}$ is constant, the elasticity $\frac{\partial \log \mathcal{D}_{t+1}}{\partial \log z_t}$ does not depend on i_t. Therefore the following result generalizes the CES case studied by Drechsler et al. (2017):

Proposition 11. *Suppose that there exist γ and a function φ such that*

$$\frac{U_x\left(c^s_{t+1}, x_{t+1}\right)}{U_c\left(c^s_{t+1}, x_{t+1}\right)} = \varphi\left(c^s_{t+1}\right) x_{t+1}^{-\gamma} \tag{59}$$

Then the monopolist bank sets a constant z^* (that does not depend on i_t), hence deposit pass-

through is "constant":[43]

$$\eta_t^d = \frac{d\log R_t^d}{d\log(1+i_t)} = 1 - \frac{d\log\left(1+s_t^d\right)}{d\log(1+i_t)} = 1 - z^* R_t^d \approx 1 - z^*$$

Importantly, (59) only involves the outer utility $U(c,x)$ and not the aggregator x that is just assumed to be homothetic and concave. Thus x can be CES as in Drechsler et al. (2017), hierarchical CES as in (10), or even a more general Kimball aggregator. Condition (59) is satisfied for instance if $U(c,x) = u(c)v(x)$ or $U(c,x) = u(c) + v(x)$ where u and v are CRRA sub-utilities. It also holds with the non-separable Greenwood et al. (1988) specification (57).

By contrast, s_t^d is increasing in i_t in the case of competitive but financially constrained banks (under the gross substitutes assumption). Under the assumption of gross substitutability between money and deposits that is common to both the models featuring market power (Drechsler et al., 2017) and those revolving around financial constraints (Di Tella and Kurlat, 2017), deposit pass-through is incomplete. The fact that whether supply adjusts or not is irrelevant means that incomplete pass-through is driven by deposit demand (and the portfolio choice between money, deposits and bonds). However, z_t has no reason to be constant in the case of constrained deposit supply. In fact, rewriting (56) as

$$\varphi(c_{t+1})\,\overline{D}^{-\gamma} x\left(f\left(\frac{s_t^d}{i_t}\right),1\right)^{-\gamma} \cdot x_d\left(f\left(\frac{s_t^d}{i_t}\right),1\right) = s_t^d \tag{60}$$

shows the following:

Corollary 2. *Holding supply fixed at $\overline{D}$, the ratio $z_t = \frac{s_t^d}{i_t}$ is a decreasing function of i_t.*

Market power and financial constraints also have different implications for *long-run* loan spreads. Suppose to simplify that the monopolist bank considered above cannot issue wholesale funding; it is straightforward to generalize the argument to costly wholesale funding as in Drechsler et al. (2017). Then in a steady state with lower nominal rate i, the monopolist bank will optimally supply more deposits. The equilibrium quantity of deposits $\mathcal{D}(z^*, i, c^s)$ is decreasing in i by (56). Therefore,

Corollary 3. *With a monopolist bank, both the steady state deposit spread τ^d and the steady state loan spread τ^l are lower when i is lower.*

This contrasts with the first fact in Section 2.1 and the stability of the total loan-deposit spread. In my long-run results in Section 4, the loan spread τ^l and the deposit spread τ^d move in opposite directions.

[43] In this discrete time formulation, pass-through still depends slightly on the level of rates through R_t^d but the effect is negligible and would disappear in continuous time.

D.2 Investment and firm borrowing

I now turn to an alternative setting in which bank loans also finance firm investment, so that flexible prices (or "natural") output is affected by financial frictions. I take a minimal model to illustrate how liquidity frictions can affect the economy's productive capacity: the result will generalize to more realistic models of investment, for instance following Elenev, Landvoigt and Nieuwerburgh (2018). Instead of having households borrowing from banks, consider in each period a unit continuum of new penniless firms producing the final good from capital and labor

$$y_t = A_t k_t^{\alpha} n_t^{\alpha}$$

where k_t is chosen at $t-1$. As in Gertler and Karadi (2011), all firms are bank-dependent and face the rental rate R_t^l; this can easily be generalized to allow for unconstrained firms able to issue bonds and thus facing a rental rate R_t. Given that total labor supply is equal to 1, equilibrium output is given by

$$Y_t = A_t^{\frac{1}{1-\alpha}} \left[\frac{\alpha}{R_t^l} \right]^{\frac{1}{1-\alpha}} .$$

The only difference with Section 4 is that the demand for loans is now

$$\mathcal{L}\left(R_t^l\right) = \left[\frac{\alpha A_t}{R_t^l} \right]^{\frac{1}{1-\alpha}} .$$

In this setting with endogenous output, a real rate shock also affects the marginal productivity of capital directly. To isolate the effect of low nominal rates, I focus on inflation. We then have the exact counterpart of Proposition 2:

Corollary 4. *In the constrained lending regime, a lower inflation target increases the steady state loan spread τ^l and thus reduces investment and output.*

D.3 Excess reserves

So far I have assumed a conventional monetary policy implementation through the outside money supply M_t, with no direct role for the banking system. Suppose now that excess reserves can be used to back deposit creation, and that reserves are more collateralizable than other assets (bonds and loans). The bank pledgeability constraint becomes

$$d_{t+1} \leq \theta^f f_{t+1} + \theta \left(a_{t+1} + \mathbf{q}_{t+1}^l \cdot \mathbf{l}^{t+1} \right)$$

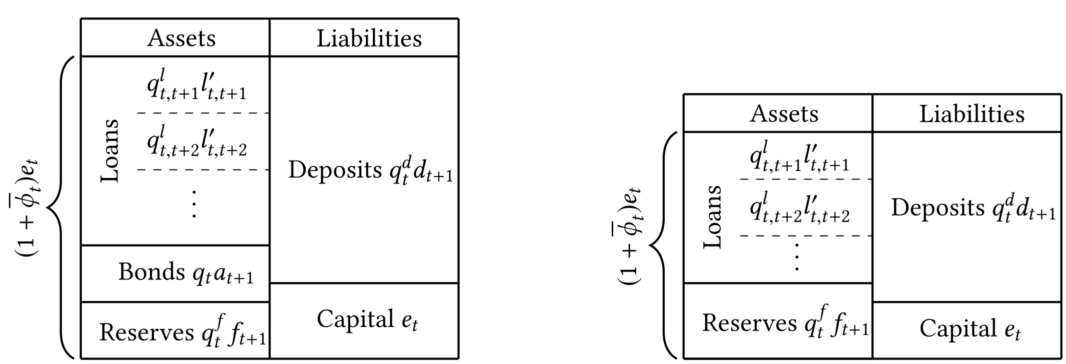
where f_{t+1} is reserve holdings at $t+1$ and $\theta^f \geq \theta$ is the pledgeability of reserves.[44]

Banks cannot short reserves, so their program becomes

$$
\begin{aligned}
V_t(e_t) \equiv \max_{a_{t+1}, d_{t+1}, \mathbf{l}^{t+1}} \quad & \frac{\beta u'\left(c^s_{t+1}\right)}{u'\left(c^s_t\right)} \left\{\rho e_{t+1} + (1-\rho) V_{t+1}(e_{t+1})\right\} \\
\text{s.t.} \quad & \mathbf{q}^l_t \cdot \mathbf{l}^{t+1} + q_t a_{t+1} + q^f_t f_{t+1} = e_t + q^d_t d_{t+1} \\
& e_{t+1} = \mathbf{q}^l_{t+1} \cdot \mathbf{l}^{t+1} + a_{t+1} + f_{t+1} - d_{t+1} \\
& \phi_t \leq \overline{\phi}_t \\
& a_{t+1} \geq 0 \\
& f_{t+1} \geq 0
\end{aligned}
$$

where the maximal leverage ratio is now $\overline{\phi}_t = \frac{\psi_t \theta^f \frac{q^d_t}{q^f_t} + (1-\psi_t)\theta \frac{q^d_t}{q^l_t}}{1 - \left[\psi_t \theta^f \frac{q^d_t}{q^f_t} + (1-\psi_t)\theta \frac{q^d_t}{q^l_t}\right]}$ and $\psi_t = \frac{q^f_t f_{t+1}}{q^f_t f_{t+1} + q_t a_{t+1} + \mathbf{q}^l_t \cdot \mathbf{l}^{t+1}}$ is the share of excess reserves in bank assets. There are still two regimes of bank lending at any point in time, and Figure 23 shows bank balance sheets in the two regimes. The central

Figure 23: Bank balance sheet in the two regimes



(a) Unconstrained lending

(b) Constrained lending

bank sets a price of reserves q^f_t and commits to supplying any quantity at that price (the net income is rebated lump-sum to savers, as for seigniorage). Banks' no-arbitrage condition between excess reserves and loans is

$$
\frac{q^d_t - q^l_{t,t+1}}{q^d_t - q^f_t} \geq \frac{1-\theta}{1-\theta^f}, \qquad \left[\frac{q^d_t - q^l_{t,t+1}}{q^d_t - q^f_t} - \frac{1-\theta}{1-\theta^f}\right] f_{t+1} = 0 \tag{61}
$$

[44] In the background, there could be a decentralized market for reserves in response to idiosyncratic withdrawal shocks, as in Bianchi and Bigio (2017).

Focusing on a case with positive excess reserves f_{t+1}, (61) binds and we can get rid of the reserve-deposit spread and reexpress the leverage ratio as a function of two endogenous objects, the reserve share ψ_t and the (endogenous) loan-deposit spread $\frac{q_t^d}{q_{t,t+1}^l}$:

$$\bar{\phi}_t = \frac{\frac{q_t^d}{q_{t,t+1}^l}\left[(1-\psi_t)\,\theta + \psi_t \frac{\theta f}{\frac{1-\theta f}{1-\theta} + \frac{q_t^d}{q_{t,t+1}^l}\left(\frac{\theta f - \theta}{1-\theta}\right)}\right]}{1 - \frac{q_t^d}{q_{t,t+1}^l}\left[(1-\psi_t)\,\theta + \psi_t \frac{\theta f}{\frac{1-\theta f}{1-\theta} + \frac{q_t^d}{q_{t,t+1}^l}\left(\frac{\theta f - \theta}{1-\theta}\right)}\right]} \tag{62}$$

The central bank can implement ψ (within an admissible range) by setting q^f so we can take a primal approach and directly use ψ as control variable. Differentiating (62) with respect to ψ_t holding $\frac{q_t^d}{q_{t,t+1}^l}$ fixed is enough to obtain long-run result, as the endogenous movement in $\frac{q_t^d}{q_{t,t+1}^l}$ cannot overturn the direct effect:

Proposition 12. *In a constrained lending steady state, an increase in ψ decreases τ^d and increases τ^l.*

This policy has a similar effect to a decrease in the inflation target. A large central bank balance sheet translates into larger commercial banks' balance sheets. As banks issue more deposits, the price of deposits is pushed down. However, this does not translate into more loans, because excess reserves (lending to the government) crowd out lending to the private sector. This crowding-out result arises because the "unconventional monetary policy" (the choice of ψ) is not targeted at buying loans directly. The central bank only sets an interest rate on excess reserves and lets private banks freely adjust their portfolios. In Gertler and Karadi (2011), unconventional monetary policy crowds-in lending because the central bank can purchase loans or equivalently subsidize lending to the private sector. See Di Maggio, Kermani and Palmer (2016) for an empirical comparison between QE1 and QE2.

D.4 Heterogeneous banks (coming soon)