2024

Essays on bank heterogeneity and monetary policy

BOSTON UNIVERSITY

GRADUATE SCHOOL OF ARTS AND SCIENCES

Dissertation

ESSAYS ON BANK HETEROGENEITY AND MONETARY POLICY

by

ANGEL DAVID FERNANDEZ ROJAS

B.A., Pontificia Universidad Católica del Perú, 2009
M.A., Pontificia Universidad Católica del Perú, 2013

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

2024

Approved by

First Reader

Robert G. King, Ph.D.
Professor of Economics

Second Reader

Adam Guren, Ph.D.
Associate Professor of Economics

Third Reader

Masao Fukui, Ph.D.
Assistant Professor of Economics

Acknowledgments

I would like to thank my main advisor Robert G. King for his excellent guidance and support over the past years. I am also grateful to Adam Guren, Masao Fukui, and Pascual Restrepo for their valuable and insightful comments and suggestions on my work.

Gracias a mis padres Angel y Maruja y a mi esposa Diana por su apoyo incondicional y amor infinito. Un agradecimiento especial a mi hermano Nestor Alonso por su bondad y alegría. Finalmente, quiero agradecer a mi hija Sofía por la felicidad, el amor y la alegría que trae a nuestras vidas.

Angel Fernández Rojas

Department of Economics

Boston University

ESSAYS ON BANK HETEROGENEITY AND MONETARY POLICY

ANGEL DAVID FERNANDEZ ROJAS

Boston University, Graduate School of Arts and Sciences, 2024

Major Professor: Robert G. King, Ph.D.
Professor of Economics

ABSTRACT

In this dissertation, I study how bank heterogeneity and the marginal propensity to lend affect the transmission of monetary policy.

In the first chapter, I develop a banking model with heterogeneous banks to study how heterogeneity in marginal propensities to lend and responses of deposits to monetary shocks affect the monetary transmission to bank lending. The marginal propensity to lend (MPL) measures how much lending increases after an idiosyncratic one unit increase in deposits. Banks face financial frictions to substitute deposits with wholesale funding, which exposes bank lending to idiosyncratic deposit shocks. When banks are heterogeneous in the degree of financial frictions they face, the aggregate response of bank lending to monetary shocks depends on a deposit heterogeneity channel, which comes from the covariance of MPLs and responses of deposits to monetary shocks. I use U.S. bank-level data to calibrate the model and I find that heterogeneity in the degree of financial frictions dampens monetary policy by at least 17%.

In the second chapter, I study how heterogeneity in the volatility of deposit withdrawal shocks affects the monetary transmission to bank lending. I develop a general equilibrium model where banks differ in their size and small banks are endowed with

a riskier distribution of deposit withdrawal shocks, consistent with the data. In the model, small banks experience a larger decline in deposits and lending after an increase in the policy rate. Moreover, bank size heterogeneity dampens monetary policy. I use U.S. bank-level data and I find that banks at the 90th percentile of the withdrawal risk distribution reduce lending by an extra 1% and deposits by an extra 0.7-0.9% relative to banks at the 10th percentile after a monetary shock that raises the Fed funds rate by 100 basis points. Moreover, aggregate lending falls by 0.9% due to withdrawal risk.

In the third chapter, I study the role of MPLs in the transmission of monetary policy in a general equilibrium model. I incorporate banks into a standard New Keynesian DSGE model. Banks face frictions to substitute deposits with wholesale funding. I use U.S. bank-level data to calibrate the model and I find that higher financial frictions that increase the aggregate MPL by 66% amplify the response of bank lending and investment to monetary shocks by 11% and 16%, respectively. Moreover, if the sensitivity of the marginal cost of funds also increases, the loan pass-through increases by 20%, which amplifies the response of bank lending and investment by 31% and 54%, respectively. Higher MPLs do not amplify the response of production in the short run but they do at longer horizons, due to the decline in investment.

Contents

List of Tables

List of Figures

List of Abbreviations

AD	....	Aggregate Deposits
DH	....	Deposit Heterogeneity
HFI	....	High Frequency Identified
IID	....	Independent and Identically Distributed
IV	....	Instrumental Variable
MP	....	Monetary Policy
MPL	....	Marginal Propensity to Lend
OLS	....	Ordinary Least Squares
SE	....	Standard Error
US	....	United States

Chapter 1

Monetary Policy, Bank Heterogeneity, and the Marginal Propensity to Lend

1.1 Introduction

The banking literature has provided evidence that supports two facts: (i) bank lending is exposed to idiosyncratic deposit shocks and (ii) the response of deposits to monetary policy is heterogeneous across banks. I define the marginal propensity to lend (MPL) as the increase in bank lending after an idiosyncratic deposit shock that increases deposits by one unit, which implies that MPL measures the exposure of bank lending to deposit shocks. In this chapter, I study how heterogeneity in MPLs and responses of deposits to monetary policy affect the monetary transmission to bank lending when banks have costly access to non-deposit funding. In the spirit of Auclert (2019), the response of aggregate bank lending to monetary policy crucially depends on the covariance between MPLs and the exposure of deposits to monetary policy.

First, I develop a banking model where heterogeneous banks supply loans and invest in liquid assets using deposits and wholesale funding. There are three key ingredients in the model. First, banks face financial frictions to substitute deposits with wholesale funding (Stein, 1998). This assumption implies that banks cannot perfectly offset deposit changes with wholesale funding (Kashyap and Stein, 1995), which exposes bank lending to deposit shocks, i.e. $MPL > 0$. Second, banks face bank-specific loan and deposit demands, which implies that banks have monopoly power in loan and deposit markets. This assumption implies that a deposit shock that increases deposits by one unit increases bank lending by less than one ($MPL < 1$) because more lending reduces the lending rate due to loan market power, which attenuates the increase in the supply of loans. Third, banks are subject to liquidity constraints.

Loan demand depends on the lending rate and the policy rate. A higher lending rate increases the cost for borrowers and reduces loan demand, while a higher policy rate increases the opportunity cost for borrowers (for instance, they can borrow from non-banks) and increases loan demand. Similarly, deposit demand depends on the deposit rate and the policy rate. A higher deposit rate increases the remuneration to depositors and increases deposit demand, while a higher policy rate increases the opportunity cost for depositors (for instance, they can save in bonds) and reduces deposit demand.[1]

Banks face different degrees of financial frictions to substitute deposits with wholesale funding. Banks that do not face financial frictions can perfectly offset a decline in deposits with wholesale funding, which implies that their lending is not exposed to bank-specific deposit changes and their MPL is zero. With frictions, banks can only partially offset changes in deposits with wholesale funding, which exposes their lending to bank-specific deposit changes and their MPL is strictly positive. Moreover, if the liquidity requirement is not too high, MPL always increases with the degree of financial frictions.[2]

At the bank level, the response of bank lending to a monetary shock depends on a

[1]In general, the direct effect of the policy rate on deposit (loan) demand can also encapsulate other channels, different from the deposit (lending) rate, that can affect deposit (loan) demand and are exposed to changes in the policy rate. For instance, if households demand deposits to consume, then a higher policy rate reduces consumption, which decreases deposit demand.

[2]If the liquidity requirement is too high, banks use wholesale funding mainly to meet the liquidity requirement, increasing the cost of wholesale funding, and reducing loans. Then, when frictions are high, higher frictions might decrease MPLs by reducing lending.

deposits channel and a loan demand channel. The deposits channel is measured by the impact of deposit changes on bank lending, while a loan demand channel is measured by the impact of the exposure of loan demand to the policy rate, net of the loan demand elasticity, on bank lending. An increase in the policy rate reduces deposits (substitution effect), which reduces bank lending by the MPL elasticity, while loan demand increases (substitution effect), which increases bank lending by $(1 - MPL)$ because more lending increases the demand for wholesale funding and the marginal cost of these funds. Then, we have that banks with higher MPLs are more exposed to bank-specific deposit changes and less exposed to loan demand changes.

At the aggregate level, the response of bank lending depends on two aggregate channels: the aggregate deposits channel and the aggregate loan demand channel, which measure the impact of aggregate deposits and aggregate loan demand on aggregate bank lending and two bank heterogeneity channels: a deposit heterogeneity (DH) channel, which measures the impact of the heterogeneity in the response of deposits to a monetary policy shock on aggregate bank lending, and a loan demand heterogeneity channel, which measure the impact of the heterogeneity in the response of loan demand to monetary policy on aggregate bank lending. Moreover, the covariance between MPLs and the response of deposits to monetary policy and between MPLs and the response of loan demand to monetary policy provide a set of sufficient statistics for the role of bank heterogeneity in the monetary transmission to bank lending. In the rest of the chapter, I study the deposit heterogeneity channel.

Financial frictions expose bank lending to idiosyncratic deposit changes and expose deposits to idiosyncratic lending changes. Banks that do not face financial frictions experience a decline in deposits and an increase in lending after a contractionary monetary shock. Then, banks that face financial frictions experience a lower decline in their deposits because their deposits are exposed to the increase in lending.

Intuitively, banks with costly access to wholesale funding reduce their dependence on wholesale funding and keep more deposits in their balance sheet by increasing deposit rates by more.

The key theoretical result of this chapter is that bank heterogeneity in the degree of financial frictions dampens monetary policy because there is a positive covariance between MPLs and exposures of deposits to monetary policy. This covariance arises endogenously in the model. An increase in the policy rate reduces aggregate deposits, which reduces aggregate bank lending in a model with a representative bank that faces financial frictions to substitute deposits with wholesale funding. However, when banks are heterogeneous those with higher MPLs care more about their deposit funding and keep more deposits in their balance sheets so that they experience a lower decline in their deposits. Then, those banks that experience a bigger decline in their deposits have lower MPLs, which dampens the monetary transmission to aggregate bank lending.

Second, I use bank-level data from U.S. Call Reports and estimate the model parameters targeting key moments from the data. To measure the deposit heterogeneity channel (DH channel), I provide an estimate of the average MPL elasticity and the (unweighted) covariance between MPLs and responses of deposits to monetary policy to target the mean and standard deviation of the distribution of the degree of financial frictions. First, I estimate the response of loans, deposits, and deposit rates to monetary shocks using HFI shocks from Miranda-Agrippino and Ricco (2021) and call these estimates β_j^l, β_j^d, and $\beta_j^{i^d}$. Second, I run an OLS regression of β_j^l on β_j^d to get a biased estimate of the average MPL elasticity. Third, I run an IV regression of β_j^l on β_j^d using $\beta_j^{i^d}$ as an instrument for β_j^d to get another biased estimate of the average MPL elasticity. Then, if moments higher than two and the covariance between MPLs and deposit demand parameters are negligible in the data, the bias in OLS regression is proportional to the bias in the IV regression, which implies that I get a system of two equations and two unknowns and I can solve for an unbiased estimate of the average MPL elasticity. Fifth, the OLS bias can be used to derive a lower bound for the covariance between MPLs and β_j^d.

The main finding of this chapter is that the DH channel increases aggregate bank lending by 0.11% after a monetary shock that increases the policy rate by 1%, while the aggregate deposits channel (AD) channel reduces aggregate bank lending by 0.62%. These estimates imply that bank heterogeneity in the degree of financial frictions dampens monetary policy by at least 17% when banks are heterogeneous in the degree of financial frictions they face to substitute deposits with wholesale funding. A contractionary monetary shock increases loan demand and decreases deposit demand, which implies that banks that do not face financial frictions experience an increase in lending and a decline in deposits. Financial frictions expose bank lending to idiosyncratic deposit changes and expose deposits to idiosyncratic lending changes. Then, a higher degree of financial frictions reduces deposits by less because deposits are exposed to higher loan demand, and increases bank lending by less because lending is more exposed to the fall in deposits.

Another important finding is that a higher mean in the distribution of the degree of financial frictions increases the size of the AD channel while keeping the size of the DH channel unchanged, which implies that the measure of this channel is robust to changes in the mean of the distribution. Intuitively, if all banks face higher frictions to lend, their lending is more exposed to deposit shocks, which increases the aggregate MPL elasticity in the economy without affecting the dispersion of frictions across banks. I also find that more heterogeneity reduces the size of the AD channel while increasing the size of the DH channel. This occurs because more heterogeneity increases the covariance between MPLs and changes in deposits and also increases loans for banks that face less frictions, which reduces aggregate MPL elasticity.

Related literature. This chapter contributes to the literature studying bank heterogeneity and monetary policy (Kashyap and Stein, 1995, 2000; Jamilov, 2021; Bellifemine et al., 2022; Coimbra and Rey, 2023; Bianchi and Bigio, 2022). Relative to these papers, I study how the interaction between MPLs and exposure of deposits affects the transmission of monetary policy to bank lending. Moreover, I quantify the size of this channel and evaluate the implications of changes in the distribution of bank heterogeneity.

This work builds on the literature on heterogeneous agents (Auclert, 2019; Kaplan et al., 2018; Kekre and Lenel, 2022). Instead of studying the marginal propensity to consume and the marginal propensity to invest, I study the marginal propensity to lend and how bank heterogeneity affects the monetary transmission to bank lending. In these papers, the impact of heterogeneity on the aggregate response of consumption depends on the covariance between marginal propensities to consume and exposure of balance sheets to monetary shocks. Similarly, the aggregate response of bank lending depends on the covariance between MPLs and exposure of deposits to monetary shocks.

This chapter also contributes to the literature that studies the monetary transmission to bank lending (Drechsler et al., 2017, 2021; Kashyap and Stein, 2000; Kishan and Opiela, 2000; Williams, 2020; Stein, 1998). Different from these papers, I study the role of deposits and bank heterogeneity in the monetary transmission to bank lending. In particular, the role of deposits is different at the bank level and at the aggregate level. This occurs because banks that have higher MPLs also have a lower exposure of deposits to monetary shocks, which dampens the role of deposits in the monetary transmission to bank lending.

Outline. The remainder of this chapter proceeds as follows. In Section 1.2, I present a model with heterogeneous banks. The model is estimated in Section 1.3 and the main results of this chapter are discussed. Section 1.4 concludes.

1.2 A model of banks, lending, and deposits

I present a model where banks supply loans and face financial frictions to substitute deposits with wholesale funding. This friction exposes the supply of loans to bank-specific deposit changes, which gives a role for deposits in the monetary transmission to bank lending. I assume that banks differ in the degree of financial frictions they face, which induces heterogeneity in the exposure of loans to bank-specific deposit changes. The response of aggregate bank lending to a monetary shock crucially depends on the covariance between the exposure of loans to deposits and the exposure of deposits to monetary shocks.

Overview of the model. The model is static. On the asset side, banks hold bonds and supply loans using two sources of funding: deposits and wholesale funding. Banks have market power in loan and deposit markets and face a liquidity constraint and financial frictions to use wholesale funding. There is a finite number of banks and they differ in the degree of financial frictions they face to substitute deposits with wholesale funding. Loan and deposit demands are bank-specific. The interest rate on wholesale funding is the same as the interest rate on bonds but there is a noninterest cost to use this source of funding. Monetary policy sets the interest rate on bonds and deposit and lending rates are endogenous.

Assets. There are J banks in the economy and they are indexed by j. Each bank j invests in liquid assets b_j and supply loans l_j, collect deposits d_j, and demand wholesale funding w_j from non-banks. They are subject to a liquidity constraint, which is constant across banks, i.e. $b_j \geq \bar{b}$. The Central Bank issues bonds and sets the interest rate on bonds, denoted by i. In wholesale markets, banks can borrow funds w_j at the policy rate i subject to a quadratic cost in the amount borrowed, which is given by $\frac{\phi_j}{2}w_j^2$, to capture the idea that it is increasingly costly to substitute deposits with non-deposit funding (Stein, 1998; Whited et al., 2022; Drechsler et al., 2017).

$$\begin{aligned} b_j + l_j &= d_j + w_j \\ b_j &\geq \bar{b} \\ w_j &> 0 \end{aligned} \tag{1.1}$$

Loan and deposit demand. Loan and deposit demands are bank-specific and are given by the following two equations.

$$\log l_j = -\varepsilon_j^l i_j^l + \gamma_j^l i + v_j^l \tag{1.2}$$

$$\log d_j = \varepsilon_j^d i_j^d - \gamma_j^d i + v_j^d \tag{1.3}$$

where i_j^l and i_j^d denote nominal lending and deposit rates, respectively, i is the policy rate, which is the only aggregate shock in the model, ε_j^l is the loan demand elasticity, ε_j^d is the deposit demand elasticity, γ_j^l, and γ_j^d measure the exposure of loans and deposits to changes in the policy rate, respectively, and v_j^l and v_j^d are idiosyncratic lending and deposit shocks, respectively. A higher lending rate reduces loan demand by ε_j^l, while a higher policy rate increases bank lending by γ_j^l due to a higher opportunity cost for borrowers (for instance, they can borrow funds from different financial intermediaries). Similarly, a higher deposit rate increases deposit demand by ε_j^d while a higher policy rate reduces deposit demand by γ_j^d due to a higher opportunity cost for depositors (for instance, they can save in bonds issued by the Central Bank).

Financial friction. Banks face a quadratic cost to use wholesale funding, which is given by $\frac{\phi_j}{2}w_j^2$, similar to Drechsler et al. (2017). The total cost of using wholesale funding is $(i + \frac{\phi_j}{2} w_j) w_j$. Then, banks pay a spread on wholesale funding that is increasing in the amount borrowed, which makes it increasingly costly to substitute deposits with wholesale funding. The marginal cost of wholesale funding is linear in the amount borrowed w_j and an additional unit of deposits reduces the marginal cost by ϕ_j. This parameter is key in the model because it measures the degree of financial frictions, which affects the exposure of loans to bank-specific deposit changes.

Bank problem. Banks choose the amount of bonds, loans, deposits, and wholesale funding that produce the highest net income, subject to loan and deposit demands, a liquidity constraint, and their balance sheet constraint.

$$\max_{l_j, d_j, w_j} i_j^l l_j - i_j^d d_j - i w_j - \frac{\phi_j}{2} w_j^2 \tag{1.4}$$
$$\text{s.t.} \quad (1.1) - (1.3)$$

Banks set lending rates and deposit rates as a mark-up $\frac{1}{\varepsilon_j^l}$ and a mark-down $\frac{1}{\varepsilon_j^d}$ on the policy rate and the marginal cost of wholesale funding $\phi_j w_j$.

$$i_j^l = i + \frac{1}{\varepsilon_j^l} + \phi_j w_j \tag{1.5}$$
$$i_j^d = i - \frac{1}{\varepsilon_j^d} + \phi_j w_j \tag{1.6}$$

Without financial frictions, $\phi_j = 0$, banks can use any amount of wholesale funding at the policy rate i. As a result, the marginal cost of wholesale funding is no longer bank-specific and is equal to an aggregate variable i. Then, changes in lending rates and deposit rates depend only on changes in the policy rate, which implies that bank lending is not exposed to bank-specific deposit changes, and deposits are not exposed to bank-specific loan changes.

With financial frictions, $\phi_j > 0$, the marginal cost of wholesale funding is bank-specific because banks can substitute deposits with wholesale funding at a rate that is increasing in the amount they borrow. Then, a bank-specific increase in deposits reduces the demand for wholesale funding, which reduces the marginal cost of lending for banks. A lower marginal cost[3] reduces lending rates and increases bank lending. Intuitively, more deposits reduce the spread banks pay on wholesale funding, which makes lending more profitable for banks.

Marginal Propensity to Lend. Measures the exposure of loans to a bank-specific deposit shock that increases deposits by one unit, i.e. $\frac{\partial l_j/\partial v_j^d}{\partial d_j/\partial v_j^d}$, and it is denoted by MPL_j.

$$MPL_j = \frac{\varepsilon_j^l \phi_j l_j}{1 + \varepsilon_j^l \phi_j l_j} \tag{1.7}$$

Without financial frictions, banks can fully offset changes in deposits with wholesale funding so that their supply of loans is not exposed to bank-specific deposit changes. As a result, we have $MPL_j = 0$. When there are frictions, they can only partially offset changes in deposits with wholesale funding, which exposes their supply of loans to bank-specific deposit changes and we have $MPL_j > 0$. As the degree of financial frictions (ϕ_j) is higher, their supply of loans is more exposed to bank-specific deposit changes, which implies higher MPLs.

MPL elasticity. Measures the exposure of loans to a bank-specific deposit shock that increases deposits by one percent, and it is denoted by λ_j^{MPL}. It can also be viewed as the elasticity of loans to idiosyncratic deposit shocks, i.e. $\lambda_j^{MPL} = \frac{\partial l_j/\partial v_j^d}{\partial d_j/\partial v_j^d}\frac{d_j}{l_j}$.

$$\lambda_j^{MPL} = \frac{\varepsilon_j^l \phi_j d_j}{1 + \varepsilon_j^l \phi_j l_j} \tag{1.8}$$

Notice that MPLs can take any value between zero and one, while MPL elasticities can take values above one. For instance, if the degree of financial frictions, denoted by

[3]Notice that the marginal cost of wholesale funding, the marginal benefit of deposits, and the marginal cost of lending are the same.

ϕ_j, is sufficiently high, MPL will be close to but below one while the MPL elasticity can be above one because those banks with higher ϕ_j have lower loan-deposit ratios in the model. Intuitively, more severe financial frictions increase the cost of lending, lowering loan supply and inducing banks to offer higher deposit rates to collect more deposits and reduce their dependence on wholesale funding.

Monetary shock. A change in the policy rate i changes bank lending due to a deposits channel and a loan demand channel at the bank level. The aggregate response of bank lending also depends on a bank heterogeneity channel.

$$\frac{\mathrm{d}\log l_j}{\mathrm{d}i} = \underbrace{\lambda_j^{MPL}\frac{\mathrm{d}\log d_j}{\mathrm{d}i}}_{\text{Deposits channel}} + \underbrace{(1 - MPL_j)(\gamma_j^l - \varepsilon_j^l)}_{\text{Loan demand channel}} \tag{1.9}$$

A monetary policy shock that increases the policy rate $\mathrm{d}i > 0$ and reduces deposits $\frac{\mathrm{d}\log d_j}{\mathrm{d}i} < 0$, reduces bank lending $\frac{\mathrm{d}\log l_j}{\mathrm{d}i} < 0$ due to the deposits channel only if $MPL_j > 0$. There is a deposits channel of monetary policy only if banks face financial frictions to substitute deposits with wholesale funding. In this case, banks cannot fully offset a decline in deposits with wholesale funding, which exposes their supply of loans to idiosyncratic deposit changes. The MPL elasticity measures the exposure of bank lending to idiosyncratic deposit changes. Without financial frictions, there is no role for deposits in the monetary transmission to bank lending.

Monetary policy also affects bank lending through a loan demand channel. Higher lending rates $\mathrm{d}i_j^l > 0$ reduce loan demand by ε_j^l and a higher policy rate increases the opportunity cost of loans, which increases bank lending by γ_j^l. The net change in loan demand increases bank lending by $(1 - MPL_j)$. Without financial frictions, changes in bank lending are only due to the loan demand channel and are given by $\gamma_j^l - \varepsilon_j^l$. With financial frictions, this channel is dampened by $(1 - MPL_j)$ because the increase (decline) in lending increases (decreases) the cost of wholesale funding.

Notice that banks with higher MPLs are more exposed to bank-specific deposit changes and less exposed to loan demand changes. As the degree of financial frictions increases, the role of deposits in the monetary transmission to bank lending is more important and the loan demand is less relevant. Intuitively, a higher degree of financial frictions reduces the ability of banks to increase their supply of loans using a source of funding different from deposits.

Aggregation. Define aggregate bank lending $l = \sum_j l_j$ and aggregate deposits $d = \sum_j d_j$. A monetary shock induces heterogeneous responses in deposits d_j and the role of the deposits channel and the loan demand channel in the response of aggregate bank lending depends on the distribution of MPL_j.

$$
\begin{aligned}
\frac{\mathrm{d}\log l}{\mathrm{d}i} = \; & \underbrace{\lambda^{MPL}\frac{\mathrm{d}\log d}{\mathrm{d}i}}_{\text{Aggregate deposits channel}} + \underbrace{(1-MPL)(\gamma^l - \varepsilon^l)}_{\text{Aggregate loan demand channel}} \\
& + \underbrace{\sum \frac{l_j}{l}\lambda_j^{MPL}\left(\frac{\mathrm{d}\log d_j}{\mathrm{d}i} - \frac{\mathrm{d}\log d}{\mathrm{d}i}\right)}_{\text{Deposit heterogeneity channel}} \\
& - \underbrace{\sum \frac{l_j}{l} MPL_j\left((\gamma_j^l - \varepsilon_j^l) - (\gamma^l - \varepsilon^l)\right)}_{\text{Loan demand heterogeneity channel}}
\end{aligned}
\tag{1.10}
$$

where $\frac{\mathrm{d}\log l}{\mathrm{d}i}, \frac{\mathrm{d}\log d}{\mathrm{d}i}$ denote the aggregate response of loans and deposits to a change in the policy rate i, respectively. Aggregate MPL, MPL elasticity, loan demand exposure to policy rate changes, and loan demand elasticity are loans-weighted averages of their individual statistics and are denoted by $MPL, \lambda^{MPL}, \gamma^l$, and ε^l, respectively. Notice that $\gamma_j^l - \varepsilon_j^l$ measures loan demand changes, at the bank level, in the absence of frictions in wholesale funding markets.

Relative to a representative bank economy, two new channels affect the transmission of monetary policy when banks are heterogeneous: a deposit heterogeneity channel, which comes from heterogeneity in MPLs and deposit changes, and a loan demand heterogeneity channel, which comes from heterogeneity in MPLs and frictionless loan demand changes. In the absence of financial frictions, bank heterogeneity is not relevant for aggregate bank lending because MPLs are equal to zero for all banks and monetary transmission works only through the aggregate loan demand channel.

An increase in the policy rate lowers deposits, which reduces aggregate bank lending by the MPL elasticity, λ^{MPL}, due to an aggregate deposits channel. If deposits decline by 1% for all banks, then bank lending decreases by the weighted average MPL elasticity λ^{MPL}. However, not all banks experience the same decline in deposits. If those banks that experience a greater decline in their deposits also have a lower MPL elasticity, the deposit heterogeneity channel dampens the response of bank lending to monetary policy, i.e. $\sum \frac{l_j}{l} \lambda_j^{MPL} \left(\frac{\mathrm{d} \log d_j}{\mathrm{d} i} - \frac{\mathrm{d} \log d}{\mathrm{d} i} \right) > 0$. In this case, an increase in the policy rate *increases* bank lending, relative to a representative bank economy, because those banks that experience a greater decline in their deposits tend to have lower exposure to deposit shocks, as measured by the MPL.

If the loan-deposit ratio is unrelated to the response of deposits $\frac{\mathrm{d} \log d_j}{\mathrm{d} i}$, a sufficient statistic of the deposit heterogeneity channel is the covariance between MPL elasticities and the exposure of deposits to the aggregate shock, $\mathrm{Cov}_J^d(\lambda_j^{MPL}, \frac{\mathrm{d} \log d_j}{\mathrm{d} i})$, where Cov_J^d is the deposits-weighted cross-sectional covariance taken over all J banks in the economy. This channel is relevant for the aggregate response of bank lending to monetary policy only if the sufficient statistic $\sum \frac{l_j}{l} \lambda_j^{MPL} \left(\frac{\mathrm{d} \log d_j}{\mathrm{d} i} - \frac{\mathrm{d} \log d}{\mathrm{d} i} \right)$ is different from zero. If this statistic is equal to zero, then heterogeneity in the response of deposits to monetary policy is irrelevant for the monetary transmission to bank lending.

An increase in the policy rate induces firms/households to increase their loan demand by the exposure of loan demand to policy rate changes, γ^l. However, the interest rate on loans i^l also increases, which reduces loan demand by the loan demand elasticity, ε^l. The costly access to wholesale funds dampens the impact of the increase in loan demand by $(1 - MPL)$. This is the aggregate loan demand channel, which summarizes the response of bank lending to a change in loan demand in a representative bank model. In a model with a representative bank and a frictionless wholesale funding market, there is no role for a deposits channel and the response of aggregate bank lending is given by the aggregate exposure of loan demand to policy rate changes, γ^l net of the aggregate (weighted-average) loan demand elasticity, ε^l.

The exposure of loan demand to policy rate changes and the loan demand elasticities can be heterogeneous across banks, which affects bank lending through a loan demand heterogeneity channel. A sufficient statistic for this channel is the cross-sectional covariance between MPLs and the exposure of loan demand to policy rate changes net of loan demand elasticities, $\text{Cov}^l_J(MPL_j, \gamma^l_j - \varepsilon^l_j)$, where Cov^l_J denotes the loans-weighted covariance taken over all J banks in the economy. If banks that experience a bigger increase in their loan demand after an increase in the policy rate have a lower MPL, then this channel dampens the response of aggregate bank lending to monetary policy. Equation (1.9) shows that $(1 - MPL_j)$ measures the exposure of bank lending to loan demand changes, then banks with a lower MPL have a higher exposure to loan demand changes, and if these banks also experience a bigger increase in their demand, aggregate bank lending decreases by less (increases by more) relative to a representative bank economy after a monetary tightening. This channel amplifies monetary policy if these banks have higher MPLs.

Heterogeneity and bank lending. For the rest of this chapter, I will study how heterogeneity in the degree of financial frictions ϕ_j affects the monetary transmission to *aggregate* bank lending. This parameter is key for MPLs and heterogeneity in ϕ_j induces an endogenous covariance between MPLs and responses of deposits to monetary shocks.

Assumption 1 (Bound for bank liquidity)**.** *For all banks, the aggregate liquidity constraint is sufficiently low such that* $\overline{b} < b^* = \min_{\phi_j}\{\frac{1}{\phi_j \varepsilon^l} + d(\phi_j)\}$.

The liquidity constraint in the model is the same for all banks. However, if the liquidity constraint is sufficiently large, then some banks might need to use wholesale funding mainly to meet the liquidity requirement, which makes lending too costly for them.

Assumption 2 (Elasticities)**.** *Exposure of loan demand to the policy rate is higher than loan demand elasticity, i.e.* $\gamma^l > \varepsilon^l$*, and exposure of deposit demand to the policy rate is higher than deposit demand elasticity, i.e.* $\gamma^d > \varepsilon^d$*.*

Under Assumption 2, in the absence of financial frictions, a 1% increase in the policy rate increases loan demand by $\gamma^l - \varepsilon^l$, while decreases deposit demand by $\gamma^d - \varepsilon^d$.

Proposition 1. *Under Assumption 1, the following holds to first order: MPL and MPL elasticities are increasing in the degree of financial frictions* ϕ_j*. Moreover, if Assumption 2 also holds, the response of deposits to an increase in the policy rate is also increasing in the degree of financial frictions* ϕ_j*.*

Proposition 1 states that a higher degree of financial frictions increases MPL, MPL elasticity, and the response of deposits to a monetary shock, which implies that this source of bank heterogeneity dampens the monetary transmission to bank lending through the deposit heterogeneity (DH) channel. Then, a positive covariance between MPLs and exposure of deposits to monetary policy shocks arises endogenously due to heterogeneity in the degree of financial frictions.

A monetary shock that increases the policy rate reduces deposits. Banks that face financial frictions find it increasingly costly to substitute deposits with wholesale funding and increase deposit rates to avoid a large outflow of deposits from their balance sheets. However, banks that do not face financial frictions do not care much about the decline in their deposits and do not increase their deposit rate by much. Then, banks that face a higher degree of financial frictions experience a *lower* decline in their deposits after a contractionary monetary shock.

Intuitively, banks that find it more costly to substitute a decline in deposits with wholesale funding try to keep deposits in their balance sheets by increasing the remuneration to depositors. Banks that do not face financial frictions do not care about a decline in their deposits because they can easily substitute deposits with wholesale funding. Then, banks that face a lower degree of financial frictions experience a larger decline in their deposits after a contractionary monetary shock. Moreover, these banks also have a lower MPL because their supply of loans is less exposed to idiosyncratic deposit shocks. Then, bank heterogeneity dampens monetary shocks because those banks that experience a larger decline in their deposits also have lower exposure of loans to deposit changes.

1.3 Measuring the DH channel

The model in the previous section shows that heterogeneity in the degree of financial frictions ϕ_j induces a positive covariance between MPLs and responses of deposits to monetary shocks, which implies that the DH channel is positive. To provide a measure of this channel, I assume that the natural logarithm of ϕ_j follows a normal distribution with mean μ and standard deviation σ and calibrate these two parameters using OLS and IV estimates of the average (unweighted) MPL elasticity $\mathrm{E}[\lambda_j^{MPL}]$. I define the time period to be a year and estimate the annual impact of the DH channel.

1.3.1 Data description

Bank-level data is from U.S. Call Reports, which provides quarterly data on bank balance sheets and income statements of U.S. commercial banks. I use information on the amount of loans, deposits, liquid assets, wholesale funding, lending rates, and deposit rates from 1994 to 2007. Lending rates are the ratio between interest income from loans and loans and deposit rates are the ratio of interest expenses on deposits divided by deposits.

Monetary shocks come from Miranda-Agrippino and Ricco (2021) and are robust to information frictions. They use monetary surprises in the fourth federal funds futures contracts within a 30-minute window around FOMC announcements and construct a monetary shock orthogonal to the central bank's information set, and own lags. Data on the Fed funds target rate, which is the policy rate in the model, comes from FRED.

1.3.2 Estimates of the average MPL elasticity

The model in the previous section shows that the response of loans to monetary shocks is affected by bank-specific changes in deposits for banks that face financial frictions to substitute deposits with wholesale funding. To provide a first view of the average impact of deposits on bank lending, I estimate the annual response of loans and deposits to a monetary shock for each bank and call these estimates β_j^l and β_j^d, respectively. The response of deposits over a year (4 quarters) to a monetary shock for each bank j is recovered from the following time series regression:

$$\frac{d_{j,t+3} - d_{j,t-1}}{d_{j,t-1}} = \alpha_j^d + \beta_j^d \mu_t^m + \nu_{jt}^d \tag{1.11}$$

and, similarly, the response of loans over a year (4 quarters) to a monetary shock is recovered from the following time series regression:

$$\frac{l_{j,t+3} - l_{j,t-1}}{l_{j,t-1}} = \alpha_j^l + \beta_j^l \mu_t^m + \nu_{jt}^l \tag{1.12}$$

where j indexes over banks, t indexes over quarters, and μ_t^m is a monetary policy shock from Miranda-Agrippino and Ricco (2021), which is normalized to have a 1% on impact effect on the Fed funds rate. Then, I group β_j^d into 50 bins, compute the mean of β_j^l and β_j^d within each bin, and plot these variables in Figure 1·1.

Figure 1·1: Response of loans and deposits to MP shock

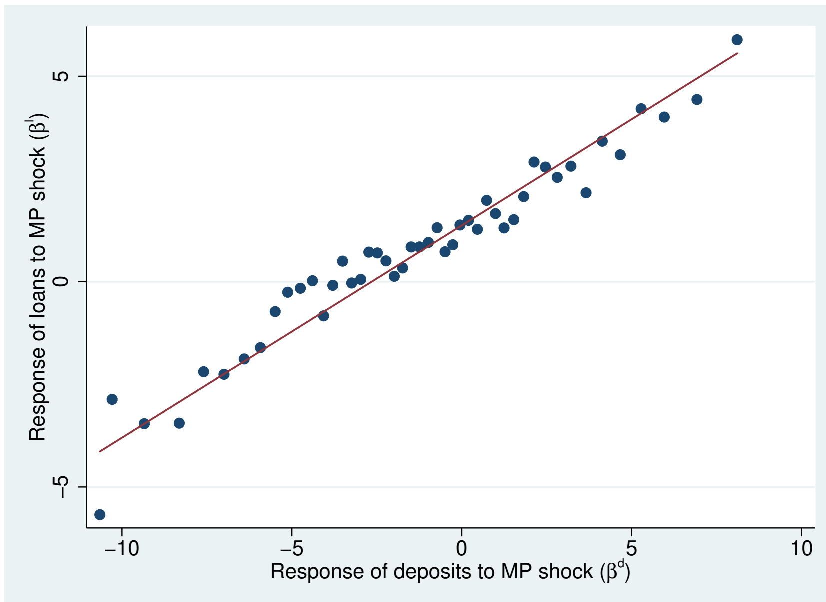


Figure 1·1 shows that banks that experience a higher increase in their deposits after a contractionary monetary shock also experience a higher increase in their lending. This result is consistent with the model from the previous section when banks face financial frictions to substitute deposits with wholesale funding. In the model, financial frictions expose bank lending to bank-specific deposit changes, which implies that MPLs matter for the monetary transmission to bank lending. Then, the slope of the line in Figure 1 provides information of the average MPL elasticity in the economy. However, the positive relation between the response of loans and deposits to monetary shocks can also arise from a positive covariance between changes in loan demand and changes in deposits.

OLS regression. Under heterogeneity in the vector $(\phi_j, \varepsilon_j^l, \varepsilon_j^d, \gamma_j^l, \gamma_j^d, \upsilon_j^l, \upsilon_j^d)$, the coefficient from a cross-sectional OLS regression of the response of loans to monetary shocks on the response of deposits to monetary shocks, denoted by λ^{OLS} is given by:

$$\begin{aligned}\lambda^{OLS} &= \mathrm{E}[\lambda_j^{MPL}] + \Big(1 - \mathrm{E}[MPL_j]\Big)\frac{\mathrm{Cov}(\gamma_j^l - \varepsilon_j^l, \beta_j^d)}{\mathrm{Var}(\beta_j^d)} \\ &+ \mathrm{E}[\beta_j^d]\frac{\mathrm{Cov}(\lambda_j^{MPL}, \beta_j^d)}{\mathrm{Var}(\beta_j^d)} - \mathrm{E}[(\gamma_j^l - \varepsilon_j^l)]\frac{\mathrm{Cov}(MPL_j, \beta_j^d)}{\mathrm{Var}(\beta_j^d)} \\ &+ \frac{\mathrm{E}[(\widehat{\lambda}_j^{MPL})(\widehat{\beta}_j^d)^2]}{\mathrm{Var}(\beta_j^d)} - \frac{\mathrm{E}[(\widehat{MPL}_j)(\widehat{\gamma}_j^l - \widehat{\varepsilon}_j^l)(\widehat{\beta}_j^d)]}{\mathrm{Var}(\beta_j^d)}\end{aligned} \tag{1.13}$$

where $\widehat{x}_j = x_j - \mathrm{E}[x_j]$ and $\mathrm{E}[x_j]$ is the cross-sectional average of x_j taken over all banks in the economy.

The first term in (1.13) is the average MPL elasticity, which is positive if at least some banks face financial frictions in the economy, i.e. $\mathrm{E}[\lambda_j^{MPL}] \geq 0$. The second term is positive in the model because banks that experience a higher increase in their loan demand increase their deposit rates by more, which increases their deposits by more, i.e. $\mathrm{Cov}(\gamma_j^l - \varepsilon_j^l, \beta_j^d) > 0$ and the average MPL is bounded by one, i.e. $(1 - \mathrm{E}[MPL_j]) \geq 0$. The third and fourth terms in (1.13) are expected to be negative if $\mathrm{E}[(\gamma_j^l - \varepsilon_j^l)] > 0$ because in the model there is a positive covariance between MPLs and the response of deposits to monetary policy and $\mathrm{E}[\beta_j^d]$ is expected to be negative. Finally, another source of bias in the estimation of $\mathrm{E}[\lambda_j^{MPL}]$ comes from two higher-order moments.

Remark 1. *If all banks can substitute deposits with wholesale funding costlessly,* $\lambda^{OLS} = 0$.

Without financial frictions, the pass-through from the policy rate to deposit and lending rates is one, which implies that the response of deposits to a 1% increase in the policy rate is the exposure of deposit demand to the policy rate γ_j^d net of the deposit demand elasticity ε_j^d and similarly for the response of loans, i.e. $\gamma_j^l - \varepsilon_j^l$. Given that the cross-sectional correlation of parameters is zero, responses of deposits and loans to monetary policy are unrelated, which implies a zero slope coefficient in the

OLS regression of β_j^l on β_j^d.

Table 1.1: OLS estimation

	Estimate
Average MPL elasticity	0.52 $[0.49, 0.55]$

Notes: This table shows an OLS estimate of the average MPL elasticity. In brackets: 95% bootstrap confidence intervals.

Consistent with Figure 1·1, Table 1.1 shows that banks that experience a 1% higher increase in deposits after a monetary shock also experience a 0.52% higher increase in their lending. This result is important because it shows that banks are financially constrained (on average) and that their lending is exposed to deposit shocks, consistent with other papers in the literature such as Khwaja and Mian (2008) and Drechsler et al. (2017). However, an OLS estimate is a biased estimate of the average MPL elasticity and the size of the bias is unknown.

If we assume that moments higher than two are negligible in the joint distribution of $(\lambda_j^{MPL}, MPL_j, \beta_j^d, \gamma_j^l, \varepsilon_j^l, \beta_j^{i^d})$, an estimate of a lower bound for the average MPL elasticity can be recovered. The idea is to do a cross-sectional IV regression of β_j^l on β_j^d using the deposit pass-through $\beta_j^{i^d}$ as an instrument of β_j^d. The coefficient from this IV regression is also biased but its bias is proportional to the OLS bias. Then, we have a system of two equations and two unknowns: a lower bound for the average MPL elasticity and an upper bound for the OLS bias.

To estimate how deposit rates change with monetary shocks for different banks, I estimate the change in deposit rates over a year (4 quarters) to a monetary shock for each bank j and call these estimates $\beta_j^{i^d}$.

$$i_{j,t+3}^d - i_{j,t-1}^d = \alpha_j^{i^d} + \beta_j^{i^d} \mu_t^m + \nu_{jt}^{i^d} \tag{1.14}$$

Then, to study the relation between changes in deposits and changes in deposit rates after a monetary shock, I group β_j^{id} into 50 bins, compute the mean of β_j^d and β_j^{id} within each bin, and plot these variables in Figure 1·2

Figure 1·2: Response of deposits and deposit rates to MP shock

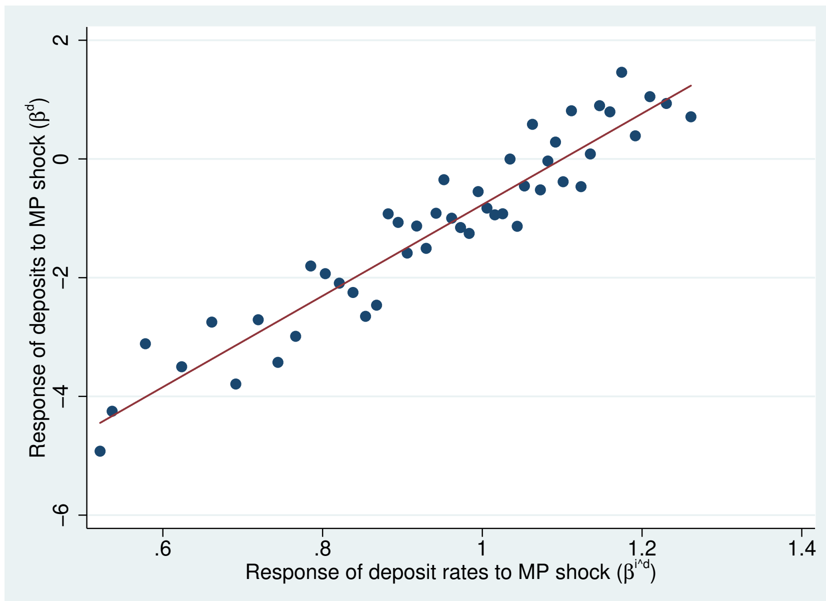


Figure 1·2 shows that there is a strong positive relation between changes in deposits and changes in deposit rates after a monetary shock. Banks that experience a higher increase in their deposit rates also experience a higher increase in their deposits after an increase in the policy rate. This result is consistent with the model and evidence from Drechsler et al. (2017). In the model, banks that face higher financial frictions increase their deposit rates by more to avoid a large outflow of deposits after a monetary contraction.

The IV coefficient is denoted by λ^{IV} and is given by:

$$\lambda^{IV} = \mathrm{E}[\lambda_j^{MPL}] + \Big(1 - \mathrm{E}[MPL_j]\Big)\frac{\mathrm{Cov}(\gamma_j^l - \varepsilon_j^l, \beta_j^{i^d})}{\mathrm{Cov}(\beta_j^d, \beta_j^{i^d})}$$
$$+ \mathrm{E}[\beta_j^d]\frac{\mathrm{Cov}(\lambda_j^{MPL}, \beta_j^{i^d})}{\mathrm{Cov}(\beta_j^d, \beta_j^{i^d})} - \mathrm{E}[(\gamma_j^l - \varepsilon_j^l)]\frac{\mathrm{Cov}(MPL_j, \beta_j^{i^d})}{\mathrm{Cov}(\beta_j^d, \beta_j^{i^d})} \tag{1.15}$$
$$+ \frac{\mathrm{E}[(\widehat{\lambda}_j^{MPL})(\widehat{\beta}_j^d)(\widehat{\beta}_j^{i^d})]}{\mathrm{Cov}(\beta_j^d, \beta_j^{i^d})} - \frac{\mathrm{E}[(\widehat{MPL}_j)(\widehat{\gamma}_j^l - \widehat{\varepsilon}_j^l)(\widehat{\beta}_j^{i^d})]}{\mathrm{Cov}(\beta_j^d, \beta_j^{i^d})}$$

Similar to the OLS estimate, the first term in (1.15) is the average MPL elasticity. The second term is positive because a higher increase in loan demand increases deposit rates by more. The third and fourth terms are expected to be negative if $\mathrm{E}[(\gamma_j^l - \varepsilon_j^l)] > 0$ because $\mathrm{E}[\beta_j^d]$ is expected to be negative and, in the model, banks that experience higher financial frictions have higher MPLs and increase their deposit rates by more after an increase in the policy rate to keep deposits in their balance sheets. Finally, another source of bias in the IV estimation depends on the joint distribution of $(\lambda_j^{MPL}, MPL_j, \beta_j^d, \gamma_j^l, \varepsilon_j^l, \beta_j^{i^d})$.

In equation (1.13), the first source of bias is the second term and comes from the covariance of loan demand parameters and responses of deposits to monetary policy, $\mathrm{Cov}(\gamma_j^l - \varepsilon_j^l, \beta_j^d)$, and loan demand parameters affect β_j^d only through changes in deposit pass-through $\beta_j^{i^d}$, then the first source of bias in the IV regression, which comes from the covariance of loan demand parameters and deposit pass-through, is proportional to the first source of bias in the OLS regression. The second source of bias is the sum of the third and fourth terms and comes from the covariance of MPLs and β_j^d, which is proportional to the covariance of MPLs and $\beta_j^{i^d}$ only if deposit demand parameters are unrelated to MPLs. The third source of bias comes from higher-order moments.

If higher-order moments are negligible and MPLs are unrelated to deposit demand parameters, i.e. $\mathrm{Cov}(\gamma_j^d, \lambda_j^{MPL}) = \mathrm{Cov}(\varepsilon_j^d, \lambda_j^{MPL}) = \mathrm{Cov}(\gamma_j^d, MPL_j) = \mathrm{Cov}(\varepsilon_j^d, MPL_j) =$

0, the IV bias in (1.15) is proportional to the OLS bias in (1.13). Then, we have a system of two equations and two unknowns: the average MPL elasticity and the OLS bias. However, in the model we have $\text{Cov}(\gamma_j^d, \lambda_j^{MPL}) < 0$, $\text{Cov}(\varepsilon_j^d, \lambda_j^{MPL}) > 0$, $\text{Cov}(\gamma_j^d, MPL_j) < 0$, $\text{Cov}(\varepsilon_j^d, MPL_j) > 0$. This occurs because higher deposit elasticity implies higher levels of deposits, which increases the level of loans and the marginal cost of wholesale funding, increasing MPLs, while higher exposure to deposits to the policy rate reduces deposits, which reduces lending and its dependence on wholesale funding, reducing MPLs. These results from the model imply that we have a lower bound for the average MPL elasticity. In the calibration, the average MPL elasticity helps to discipline the mean of the distribution of ϕ_j while the OLS bias helps to discipline its standard deviation. The next proposition provides a closed-form solution for these two elements.

Proposition 2. *Under heterogeneity in* $(\phi_j, \varepsilon_j^l, \varepsilon_j^d, \gamma_j^l, \gamma_j^d, \upsilon_j^l, \upsilon_j^d)$*, if moments higher than two are negligible in the joint distribution of* $(\lambda_j^{MPL}, MPL_j, \beta_j^d, \gamma_j^l, \varepsilon_j^l, \beta_j^{i^d})$*, we have the following result:*

$$\text{E}[\lambda_j^{MPL}] \geq \tilde{\lambda}^{MPL} = \lambda^{OLS} + \frac{\theta\varepsilon^d}{1-\theta\varepsilon^d}\Big(\lambda^{OLS} - \lambda^{IV}\Big) \tag{1.16}$$

where $\theta = \frac{\text{Cov}(\beta_j^{i^d}, \beta_j^d)}{\text{Var}(\beta_j^d)}$, $\varepsilon^d = \text{E}[\varepsilon_j^d]$*, and* $1 > \theta\varepsilon^d$*. The inequality is binding if additionally* $\text{Cov}(\gamma_j^d, \lambda_j^{MPL}) = \text{Cov}(\varepsilon_j^d, \lambda_j^{MPL}) = \text{Cov}(\gamma_j^d, MPL_j) = \text{Cov}(\varepsilon_j^d, MPL_j) = 0$.

Proposition 2 states that an estimate of a lower bound for the average MPL elasticity can be recovered by getting coefficients from three regressions: (1) OLS regression of β_j^l on β_j^d, (2) IV regression of β_j^l on β_j^d with $\beta_j^{i^d}$ as the instrument variable, and (3) OLS regression of $\beta_j^{i^d}$ on β_j^d. Moreover, if the correlation between deposit demand parameters and MPLs is sufficiently low, $\tilde{\lambda}^{MPL}$ is a good approximation of an unbiased estimate of the average MPL elasticity $\text{E}[\lambda_j^{MPL}]$.

The estimate $\tilde{\lambda}^{MPL}$ depends on the deposit demand elasticity, which is unknown but equation (1.16) can be used as an additional equation in the model to calibrate the mean of ϕ_j. Moreover, if we have an estimate of the average MPL elasticity, we can recover a lower bound for the covariance between MPL elasticities and responses of deposits to monetary shocks, which can be used to target the standard deviation of ϕ_j. In equation (1.13), if higher-order terms are negligible, the positive covariance between changes in loan demand and deposits implies the following:

$$\begin{aligned} \lambda^{OLS} - \mathrm{E}[\lambda_j^{MPL}] \geq & \left(\mathrm{E}[\beta_j^d] - (\gamma^l - \varepsilon^l)\mathrm{E}\left[\frac{l_j}{d_j}\right]\right)\frac{\mathrm{Cov}(\lambda_j^{MPL}, \beta_j^d)}{\mathrm{Var}(\beta_j^d)} \\ & - z\left(\gamma^l - \varepsilon^l\right)\mathrm{E}[\lambda_j^{MPL}] \end{aligned} \tag{1.17}$$

where $z = \frac{\mathrm{Cov}\left(\frac{l_j}{d_j}, \beta_j^d\right)}{\mathrm{Var}(\beta_j^d)}$ is the cross-sectional OLS regression of loan-deposit ratios, $\frac{l_j}{d_j}$, on responses of deposits to monetary shocks, β_j^d, and γ^l and ε^l are the cross-sectional averages of γ_j^l and ε_j^l, respectively. Then, we have a lower bound for the covariance $\mathrm{Cov}(\lambda_j^{MPL}, \beta_j^d)$.

$$\mathrm{Cov}(\lambda_j^{MPL}, \beta_j^d) \geq \mathcal{C} = \frac{\left(\lambda^{OLS} - \mathrm{E}[\lambda_j^{MPL}] + z\left(\gamma^l - \varepsilon^l\right)\mathrm{E}[\lambda_j^{MPL}]\right)\mathrm{Var}(\beta_j^d)}{\left(\mathrm{E}[\beta_j^d] - (\gamma^l - \varepsilon^l)\mathrm{E}\left[\frac{l_j}{d_j}\right]\right)} \tag{1.18}$$

In equation (1.18), the average MPL elasticity, the OLS estimate λ^{OLS}, the OLS estimate z, and the mean and variance of β_j^d can be used to estimate a lower bound for the covariance between MPLs and the response of deposits to monetary shocks, which is the measure of the DH channel when all banks have the same ratio of loans over deposits. This equation can be used to target the standard deviation of ϕ_j.

1.3.3 Estimation of the model

I assume that banks differ only in the degree of financial frictions, ϕ_j, and the natural logarithm of ϕ_j follows a normal distribution with mean μ and standard deviation σ. Then, there are eight parameters in the model: (1) loan demand elasticity, ε^l, (2) exposure of loan demand to the policy rate, γ^l, (3) deposit demand elasticity, ε^d, (4) exposure of deposit demand to the policy rate, (5) loan demand shock v^l, (6) deposit demand shock v^d, (7) mean μ, and (8) standard deviation σ. The model provides eight equations to target these parameters:

$$\mathrm{E}[i_j^l - i] = \frac{1}{\varepsilon^l} + \mathrm{E}[\phi_j w_j] \tag{1.19}$$

$$\mathrm{E}[i - i_j^d] = \frac{1}{\varepsilon^d} - \mathrm{E}[\phi_j w_j] \tag{1.20}$$

$$\mathrm{E}[\beta_j^l] = \mathrm{E}[\lambda_j^{MPL} \beta_j^d] + (1 - \mathrm{E}[MPL_j])(\gamma^l - \varepsilon^l) \tag{1.21}$$

$$\mathrm{E}[\beta_j^d] = \mathrm{E}\Big[\frac{\varepsilon^d \phi_j l_j (\gamma^l - \varepsilon^l) - (1 + \varepsilon^l \phi_j l_j)(\gamma^d - \varepsilon^d)}{1 + \varepsilon^l \phi_j l_j + \varepsilon^d \phi_j d_j}\Big] \tag{1.22}$$

$$\mathrm{E}[l_j] = \mathrm{E}[\exp\big(-\varepsilon^l i_j^l + \gamma^l i + v^l\big)] \tag{1.23}$$

$$\mathrm{E}[d_j] = \mathrm{E}[\exp\big(\varepsilon^d i_j^d - \gamma^d i + v^d\big)] \tag{1.24}$$

$$\tilde{\lambda}^{MPL} = \lambda^{OLS} + \frac{\theta \varepsilon^d}{1 - \theta \varepsilon^d}\Big(\lambda^{OLS} - \lambda^{IV}\Big) = \mathrm{E}\Big[\frac{\varepsilon_j^l \phi_j d_j}{1 + \varepsilon_j^l \phi_j l_j}\Big] \tag{1.25}$$

$$\mathrm{Cov}(\lambda_j^{MPL}, \beta_j^d) = \mathcal{C} = \frac{\Big(\lambda^{OLS} - \tilde{\lambda}^{MPL} + z\Big(\gamma^l - \varepsilon^l\Big)\tilde{\lambda}^{MPL}\Big)\mathrm{Var}(\beta_j^d)}{\Big(\mathrm{E}[\beta_j^d] - (\gamma^l - \varepsilon^l)\mathrm{E}\Big[\frac{l_j}{d_j}\Big]\Big)} \tag{1.26}$$

where $z = \frac{\mathrm{Cov}\left(\frac{l_j}{d_j}, \beta_j^d\right)}{\mathrm{Var}(\beta_j^d)}$, $\theta = \frac{\mathrm{Cov}(\beta_j^{i^d}, \beta_j^d)}{\mathrm{Var}(\beta_j^d)}$, $\lambda^{OLS} = \frac{\mathrm{Cov}(\beta_j^l, \beta_j^d)}{\mathrm{Var}(\beta_j^d)}$, and $\lambda^{IV} = \frac{\mathrm{Cov}(\beta_j^l, \beta_j^{i^d})}{\mathrm{Cov}(\beta_j^d, \beta_j^{i^d})}$.

In equation (1.19), the average spread between the lending rate and the policy rate from the data can be used to target loan demand elasticity ε^l. Similarly, equation (1.20) helps to target deposit demand elasticity ε^d using the average spread between the policy rate and the deposit rate from the data. The cross-sectional (unweighted) average of the bank-level estimates of the responses of loans to monetary shocks from equation (1.12) can be used to target γ^l in equation (1.21) and the average (unweighted) bank-level estimates of the response of deposits to monetary shocks in equation (1.11) can be used to target γ^d in equation (1.22). Equations (1.23) and (1.24) are used to target the average amount of loans and deposits, respectively. In equation (1.25), the estimate of the average (unweighted) MPL elasticity can be used to target mean μ, while equation (1.26) targets standard deviation σ.

To solve for the value of the parameters, I define a vector F, and each element in this vector is the difference between a data moment and a model moment, weighted by the inverse of the standard deviation of the data moment. Vector F takes the following form:

$$F = \begin{bmatrix} (\mathrm{E}[i^l_j] - mean(i^l_j))/std(i^l) \\ (\mathrm{E}[i^d_j] - mean(i^d_j))/std(i^d) \\ (\mathrm{E}[\beta^l_j] - mean(\beta^l_j)/std(\beta^l_j) \\ (\mathrm{E}[\beta^d_j] - mean(\beta^d_j)/std(\beta^d_j) \\ (\mathrm{E}[l_j] - mean(l_j)/std(l_j) \\ (\mathrm{E}[d_j] - mean(d_j))/std(d_j) \\ (\mathrm{E}[\lambda^{MPL}_j] - \tilde{\lambda}^{MPL})/std(\tilde{\lambda}^{MPL}) \\ (\mathrm{Cov}\big(\lambda^{MPL}_j, \beta^d_j) - \mathcal{C})/std(\mathcal{C}) \end{bmatrix} \tag{1.27}$$

The value of $F'F$ is minimized and the number of banks in the estimation procedure is $J = 20,000$. We use eight moments to target eight parameters.

Table 1.2: Data moments

Moment	Value	Standard deviation
Mean (i^l)	0.0845	0.0088
Mean (i^d)	0.0287	0.0059
Mean (β^d)	-1.1525	4.9004
Mean (β^l)	0.7731	5.7155
Mean (l_j)	2.5810	33.7136
Mean (d_j)	2.4663	27.8293
$\tilde{\lambda}^{MPL}$	0.7237	0.0279
$\mathcal{C}$	0.7897	0.3527

Table 1.2 shows data moments and their standard deviation. The model sets a policy rate equal to $i = 4\%$, roughly similar to the average Fed funds rate of 4.20%. The average lending rate is the Fed funds rate of 4% plus the average spread between the lending rate and the Fed funds rate, which implies an average lending rate of 8.45%. Similarly, the deposit rate is the Fed funds rate of 4% minus the average

deposit spread in the data, which implies an average deposit rate of 2.87%. From equation (1.11), the average response of deposits to monetary shocks is -1.15%, and the standard deviation of these estimates is 4.9. From equation (1.12), the average response of loans to monetary shocks is 0.77%, with a standard deviation of 5.71. After normalizing the average amount of wholesale funding equal to 1, the average amount of loans is 2.58, the average amount of deposits is 2.47, and the liquidity constraint is $\bar{b} = 0.89$. From equations (1.16) and (1.18), the average MPL elasticity $\tilde{\lambda}^{MPL}$ and the covariance $\mathcal{C}$ depend on model parameters: the elasticity of deposits, ε^d, and the exposure of loans to the policy rate net of the loan demand elasticity, $\gamma^l - \varepsilon^l$. To compute the standard error of these two estimates, I calibrate the model to a representative bank economy and get values $\varepsilon^d = 28.0$ and $\gamma^l - \varepsilon^l = 6.6$. These model parameters imply that the estimate of the average MPL elasticity is 0.72 and its standard error is 0.03, while the estimate of the covariance between MPLs and deposit betas is 0.79 with a standard error of 0.35.

Table 1.3: Estimated parameters

Parameter	Description	Value	Standard error
ε^l	Loan demand elasticity	49.6971	1.9511
γ^l	Exposure of loan demand to the policy rate i	56.3458	0.1198
v^l	Loan demand shock	2.8899	0.1831
ε^d	Deposit demand elasticity	27.9970	0.6276
γ^d	Exposure of deposit demand to the policy rate i	32.4598	1.8782
v^d	Deposit demand shock	0.9206	0.2382
μ	Mean of log of the degree of financial frictions ϕ_j	-3.5958	0.2663
σ	Std of log of the degree of financial frictions ϕ_j	1.8757	0.8221

Table 1.3 shows the estimated values and standard errors of the eight parameters in the model. Without frictions, banks increase loans by 6.6% (=56.3-49.7) after an increase of 1% in the policy rate, while deposits fall by 4.5%(=32.5-28.0). Banks that face financial frictions to substitute deposits with wholesale funding have a loan supply

that is affected by changes in deposits. This implies that bank lending increases by less (or decreases by more) for these banks. As the degree of financial frictions increases, the deposits channel becomes more important for these banks which induces a bigger decline in lending. These banks also experience a lower decline in deposits because higher loan demand increases the need for deposits given that they have costly access to wholesale markets.

Table 1.4: Model and data moments

Moment	Data	Model
Mean (i^l)	0.0845	0.0876
Mean (i^d)	0.0287	0.0318
Mean (β^d)	-1.1525	-1.5027
Mean (β^l)	0.7731	1.3942
Mean (l_j)	2.5810	2.5849
Mean (d_j)	2.4663	1.6902
$\tilde{\lambda}^{MPL}$	0.7237	0.7217
$\mathcal{C}$	0.7854	0.4424

Table 1.4 shows data and model moments. The average lending rate is 8.45% in the data and 8.76% in the model, while the average deposit rate is 2.87% in the data and 3.18% in the model. Deposits decrease by 1.50%, on average, in the model and by 1.15% in the data. Loans increase, on average, by 1.39% in the model and 0.77% in the data. The average amount of loans is 2.58 in the model and in the data and the average amount of deposits is 1.69 in the model and 2.47 in the data. The average MPL elasticity is 0.72 in the model and in the data and the covariance between MPLs and responses of deposits to monetary shocks is 0.79 in the data and 0.44 in the model.

1.3.4 Estimation results

The model assumes heterogeneity only in the degree of financial frictions ϕ_j and Table 1.5 shows the main results from the model calibration. In the model, a 1% increase in the policy rate reduces aggregate bank lending by 0.62% through the aggregate deposits channel (AD channel) while *increases* aggregate bank lending by 0.11% through the deposit heterogeneity channel (DH channel). This means that heterogeneity in the degree of financial frictions dampens the contractionary impact of monetary policy on bank lending by 17%.

Table 1.5: Estimates of aggregate MPL, and AD and DH channels

AD channel	-0.6248
DH channel	0.1062
Aggregate MPL elasticity	0.4875
Aggregate MPL	0.5741
Aggregate resp. of dep.	-1.2815
Aggregate resp. of loans	2.3131

Banks that face a higher degree of financial frictions reduce their lending and have higher MPL elasticities, which implies that smaller banks have higher MPLs. This correlation reduces the aggregate (weighted average) MPL elasticity in the economy, which is 0.49 in the model, lower than the average MPL elasticity of 0.72. Similarly, banks with high ϕ_j have more deposits and experience a lower decline in deposits after a monetary tightening, which implies that the aggregate decline in deposits after a 1% increase in the policy rate is lower than the average (unweighted) decline in deposits. Aggregate deposits fall by 1.28% while the (unweighted) average decline in deposits is 1.50%.

Table 1.5 shows the main result of this chapter: bank heterogeneity in the degree of financial frictions *increases* aggregate bank lending by 0.11% after a 1% in the policy rate. In the absence of frictions, bank lending increases while deposits fall after a contractionary monetary shock. With frictions, a higher loan demand induces banks to increase their deposits, which reduces the contractionary impact of monetary policy on deposits. Moreover, frictions expose bank lending to deposit shocks and this exposure is measured by MPLs. Then, we have that banks that are more exposed to deposit shocks experience a lower decline in deposits after a contractionary monetary shock, which dampens monetary policy by 17%.

1.3.5 Changes in the distribution of frictions

The distribution of financial frictions affects the size of the DH channel. In this section, I study how changes in the mean μ and standard deviation σ of the natural logarithm of ϕ_j affect key statistics such as the aggregate MPL elasticity, the aggregate response of deposits to monetary shocks, the AD channel and the DH channel.

Changes in the mean μ. The solid line in Figure 1·3 shows how aggregate MPL elasticities and aggregate responses of deposits to monetary shocks change when we allow the mean μ of the log of ϕ_j to vary. The range of μ goes from two standard errors below its estimated value to two standard errors above this value. The dashed line is the estimated value of μ. A higher mean implies that the cost of wholesale funding increases more rapidly with the amount borrowed. Then, banks are more exposed to deposit shocks, which increases aggregate MPL and banks keep more deposits on their balance sheet to reduce their dependence on wholesale funding, which reduces the aggregate decline of deposits to a contractionary monetary shock.

Figure 1·3: Aggregate MPL elasticity and response of deposits to MP - Changing μ

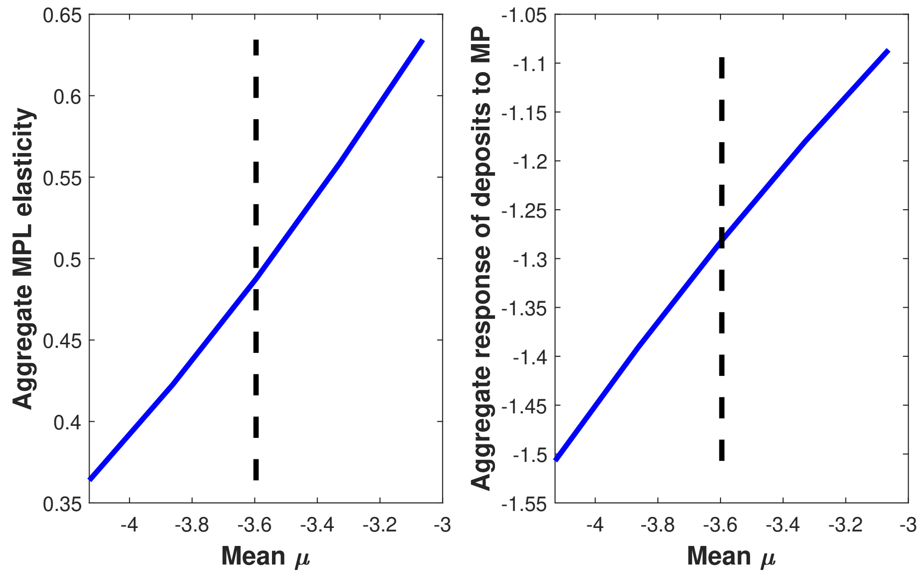


Figure 1·4 shows how AD and DH channels vary with μ. The AD channel induces a bigger decline in bank lending after a contractionary shock as μ increases. This occurs because a higher μ increases the exposure of bank lending to deposit shocks. The size of the DH channel is between 0.103 and 0.107, which implies that the size of this channel does not change much when μ changes. This result is important because it implies that the estimate of the DH channel is robust to changes in a key parameter μ.

Figure 1·4: AD and DH channels - Changing μ

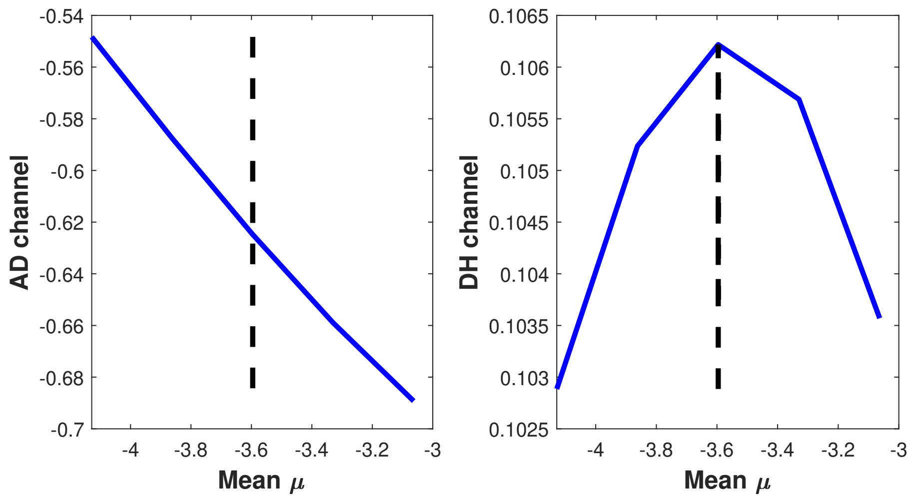
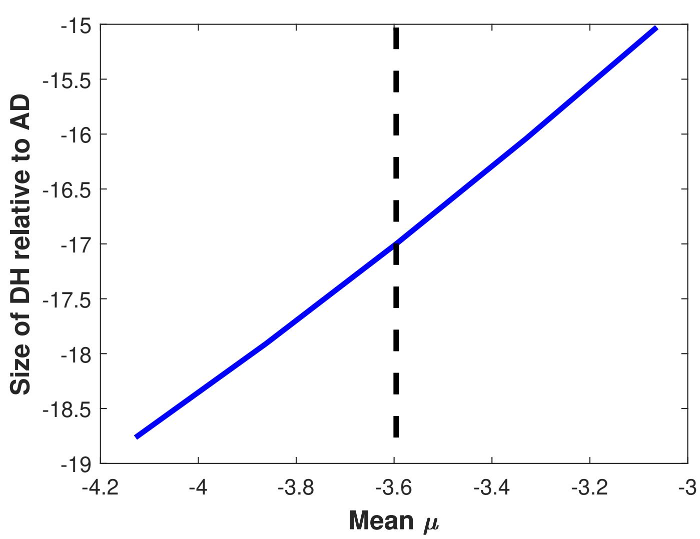


Figure 1·5 shows that the DH channel is relatively less important in economies with high μ. This occurs because the DH channel does not change much with μ while the AD channel increases (in absolute value) with μ. Then, bank heterogeneity in the degree of financial frictions dampens monetary policy by less in economies with a higher degree of financial frictions.

Figure 1·5: Relative size of DH channel - Changing μ



Changes in the standard deviation σ. The solid line in Figure 1·6 shows how aggregate MPL elasticities and aggregate responses of deposits to monetary shocks change when the standard deviation σ of the log of ϕ_j changes. The range of σ goes from two standard errors below its estimated value to two standard errors above this value. The dashed line is the estimated value of σ. A higher σ increases loans for banks with low ϕ_j and reduces loans for banks with high ϕ_j. Then, aggregate MPL elasticity falls because more loans go to banks with a supply of loans less exposed to idiosyncratic deposit shocks.

Figure 1·6 also shows that the aggregate response of deposits to monetary policy does not change much with changes in σ. As the degree of financial frictions increases, banks increase their deposits at a lower pace, which implies that a higher dispersion increases the average decline in deposits after a monetary contraction. Moreover, the aggregate decline in deposits is dampened because those banks that experience a bigger decline in their deposits also have lower deposits on their balance sheet.

Figure 1·6: Aggregate MPL elasticity and response of deposits to MP - Changing σ

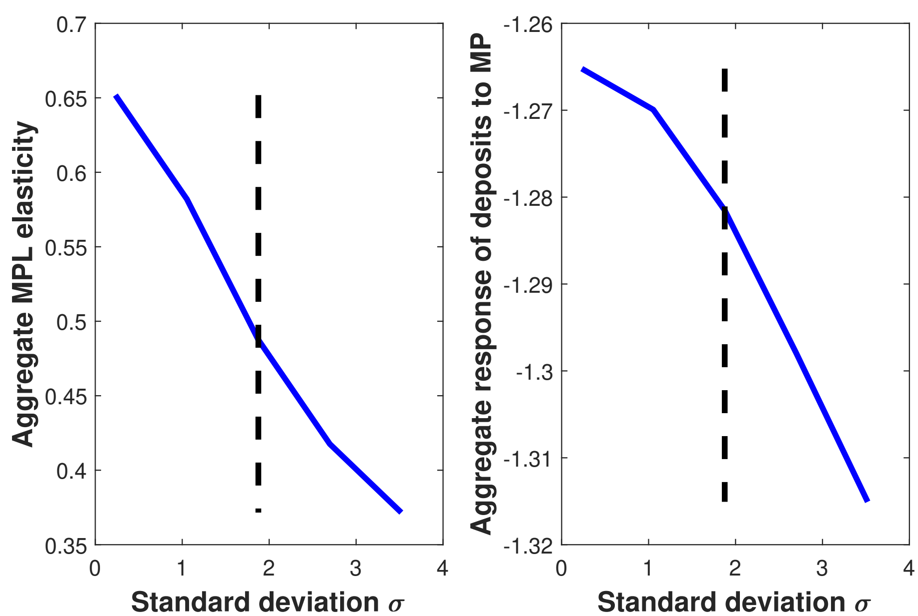


Figure 1·7 shows how AD and DH channels vary with σ. As σ increases, the

AD channel induces a lower decline in bank lending after a contractionary shock. This occurs because a higher σ reduces the exposure of aggregate bank lending to idiosyncratic deposit shocks. The size of the DH channel increases with σ. As σ increases, the dispersion of MPLs and responses of deposits to monetary shocks also increases, which increases the cross-sectional covariance between these two estimates. The size of this channel is close to zero for low values of σ, while it is 0.16 when $\sigma = 3.5$.

Figure 1·7: AD and DH channels - Changing σ

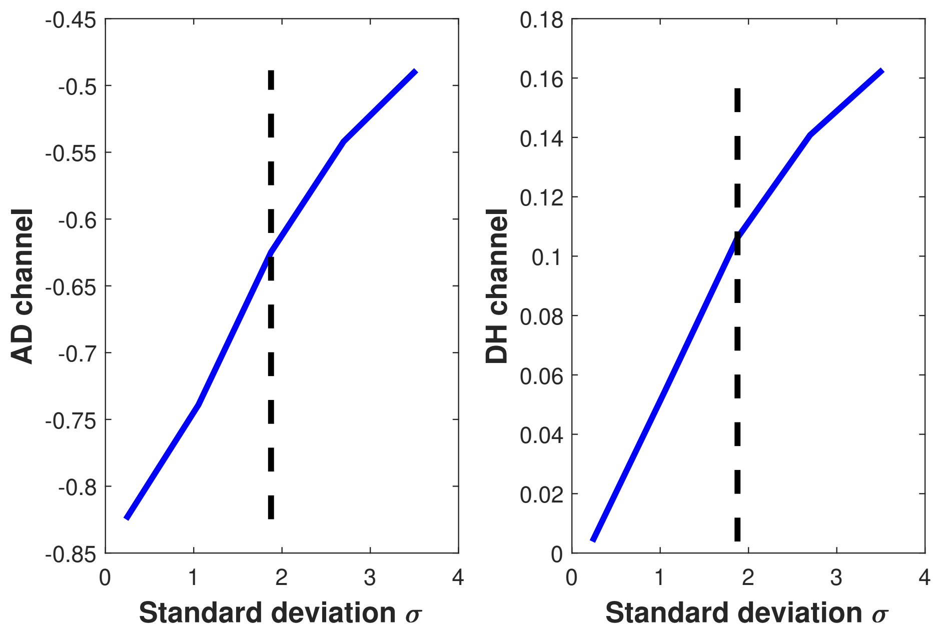
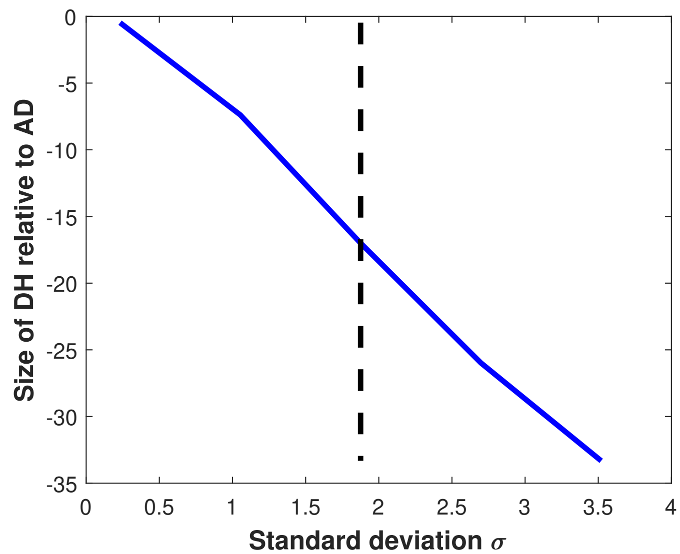


Figure 1·8 shows that the DH channel is relatively more important in economies with more heterogeneity, i.e. higher σ. This occurs because the DH channel increases with σ while the AD channel decreases (in absolute value) with σ. Then, bank heterogeneity in the degree of financial frictions dampens monetary policy by more in economies with more heterogeneity.

Figure 1·8: Relative size of DH channel - Changing σ



In summary, this section estimates that heterogeneity in the degree of financial frictions increases aggregate bank lending, through the DH channel, by at least 0.11% after a monetary shock that increases the policy rate by 1%, which dampens monetary policy transmission to bank lending by at least 17%. A higher degree of financial frictions, increases the importance of the AD channel, while keeping the DH channel unchanged. Moreover, a higher dispersion in the distribution of this friction across banks increases the DH channel, while reducing the AD channel.

1.4 Conclusion

This chapter studies the role of bank heterogeneity in the monetary transmission to bank lending when banks have costly access to non-deposit funding. The impact of bank heterogeneity on the bank lending channel is captured by the deposit heterogeneity channel, which comes from the covariance of MPLs and responses of deposits to monetary shocks. This covariance is endogenous in the model. The main finding of this chapter is that the deposit heterogeneity channel dampens monetary policy by at least 17% when banks are heterogeneous in the degree of financial frictions they face to substitute deposits with wholesale funding.

A contractionary monetary shock increases loan demand and decreases deposit demand, which implies that banks that do not face financial frictions experience an increase in lending and a decline in deposits. Financial frictions expose bank lending to idiosyncratic deposit changes and expose deposits to idiosyncratic lending changes. Then, as the degree of financial frictions increases deposits decrease by less because deposits are more exposed to the increase in loan demand, and bank lending increases by less because lending is more exposed to the fall in deposits.

Another important result is that a higher mean of the distribution of financial frictions increases the importance of the aggregate deposits channel, while keeping the deposit heterogeneity channel unchanged, which implies that the measure of this channel is robust to changes in the mean of the distribution. Moreover, more heterogeneity reduces the aggregate deposits channel while increasing the size of the deposit heterogeneity channel.

Chapter 2

Bank Heterogeneity, Reserve Management, and Monetary Policy

2.1 Introduction

The 2007-2008 financial crisis sheds light on the importance of understanding the macroeconomic impact of bank behavior. In particular, banks need reserves to create loans and issue deposits, and monetary policy plays a central role by affecting the trade-off between holding reserves and loans. Moreover, empirical evidence from the bank lending literature highlights the heterogeneous responses of banks to monetary policy. Therefore, this chapter proposes a simple macroeconomic model with banks facing idiosyncratic withdrawal risk.

I develop a general equilibrium model of a monetary economy with heterogeneous banks. The main contribution is the study of bank's reserves management under heterogeneous exposure to withdrawal shocks. In this model, the economy is populated by two types of banks: small and big banks, and small banks are endowed with a riskier withdrawal distribution. Banks provide loans, demand reserves, issue deposits, and demand discount window loans from the central bank subject to a capital requirement constraint. Depositors transfer deposits from one bank to another at any time to make transactions and banks must settle these transactions with reserves. In the case the outflow of deposits is larger than the amount of reserves banks hold in their balance sheet, they use discount window loans from the central bank.

In general equilibrium, workers consume and supply labor and deposits while entrepreneurs demand labor and loans to produce the final consumption good. Following an increase in the policy interest rate, small banks experience a larger decline in loans and deposits. Moreover, the increase in the spread between lending and deposit rates reduces aggregate output which further increases the gap in the response between large and small banks due to capital requirement constraints.

Another contribution of this chapter is that it explains the different responses between big and small banks after monetary policy actions. In the U.S. economy, there is a negative association between the standard deviation of the growth of deposits and the bank's size. Given that the volatility of deposit growth is a proxy for withdrawal uncertainty, my model predicts that big banks are less affected by monetary policy. This result is consistent with estimations provided by Kashyap and Stein (2000) and Kishan and Opiela (2000) in the bank lending literature.

This chapter distinguishes between partial and general equilibrium effects of monetary policy. In partial equilibrium, a contractionary monetary policy increases the marginal cost of lending, which reduces the supply of loans and increases the interest rate on loans and the demand for reserves. Moreover, there is an increase in the marginal cost of issuing deposits, which increases the spread between the interest rate on loans and deposits. Smaller banks, endowed with a higher withdrawal risk, respond by increasing their reserves and reducing their lending by more.

The marginal cost of issuing deposits is lower for big banks, which implies that they issue deposits up to its capital requirement constraint and therefore, its deposit demand is not directly affected by the aggregate deposit demand. Therefore, only small banks are affected by aggregate shocks to deposits. In general equilibrium, a contractionary monetary policy reduces the demand for labor, which reduces the aggregate loan demand and deposit demand. This decline is proportionally larger for small banks due to a higher exposure of their deposits to aggregate shocks. Moreover, the contraction in loan supply by small banks is partially offset by large banks, which reduces the response of the banking system and increases the gap in the response between both types of banks.

Using a quarterly bank-level dataset constructed by Drechsler et al. (2017) from U.S. Call Reports, I construct a measure of withdrawal uncertainty given by the standard deviation of the quarterly growth rate of deposits.[1] I find that an increase in 100 basis points in the Fed funds rate, reduces lending from banks in the 75th percentile by an extra 0.4% with respect to banks in the 25th percentile and by an extra 0.8% if we compare banks in the 90th-10th percentile. Moreover, this result is robust to the inclusion of variables that measure bank size, liquidity and concentration. I also provide evidence that bank leverage is higher for big banks, which implies that they are closer to the capital requirement constraint, a result of the model proposed.

This chapter contributes to the growing literature on banks and monetary policy in macroeconomic models. Several papers model the behavior of financial intermediaries and their interactions with monetary and macroprudential policies. Brunnermeier and Sannikov (2016) develop a theory of inside and outside money and the role of monetary policy and macroprudential policies to enhance price and financial stability. Bianchi and Bigio (2022) develop a dynamic general equilibrium model of banks' liquidity management and the credit channel and introduce an OTC interbank market to study the role of monetary policy in altering the liquidity premium. In their model, banks hold optimally a positive amount of reserves. Moreover, they find that the main drivers of the financial crisis were an early disruption in the interbank market, followed by a substantial and persistent decline in credit demand.

This chapter also contributes to the bank lending literature. Some relevant papers

[1] This growth rate has been winsorized at 1% and 99% level at every quarter to avoid the presence of outliers.

are (Kashyap and Stein, 2000) and Kishan and Opiela (2000). The first one finds that a monetary policy has a stronger effect on lending by small banks with lower liquidity ratios while the second one finds that lending by small and undercapitalized banks is more responsive to monetary policy. A related work by Drechsler et al. (2017) provides evidence for a deposits channel of monetary policy where banks widen their spreads on deposits after an increase in the policy rate, which reduces deposits and lending. Moreover, banks that raise deposits in more concentrated markets contract their lending by more.

Outline. The remainder of this chapter proceeds as follows. Section 2.2 presents some banking sector facts for the U.S. economy. Consistent with these facts, Section 2.3 describes the model and presents partial and general equilibrium results. Section 2.4 provides empirical evidence consistent with the main results from the model. Section 2.5 concludes.

2.2 Banking Sector Facts

In the U.S. economy, small banks face higher withdrawal uncertainty and have a higher capital-asset ratio. Using quarterly data from U.S. commercial banks from 1994Q1 to 2006Q4, a measure of bank size is constructed as the average of the relative size of assets. Let S_j denotes the size of bank j, A_{jt} denotes total assets from bank j in quarter t, Then bank size is defined as follows:

$$S_j = \frac{\sum_{t=T_{0j}}^{T_j} \frac{m_{jt}}{h_t} 100}{T_j - T_{0j}}$$

where m_{jt} indicates the relative position of total assets A_{jt} in increasing order such that $m_{jt} = 1$ for the smallest bank in terms of total assets in quarter t and $m_{jt} = h_t$ for the largest bank in quarter t, h_t is the total number of banks in quarter t and T_{0j} and T_j are the initial and final periods for bank j, respectively. Therefore, if a given bank is the largest bank in every quarter in which it is operating, the size variable will assigned a value of 100 to this bank and similarly if a given bank is the smallest in terms of total assets in every quarter in which this bank is present, the size variable will assign a value close to zero given the large number of banks in U.S. For example if $h_t = 5000$ in every quarter, then the size variable S_j will assign a value of $\frac{100}{5000} = 0.02$.

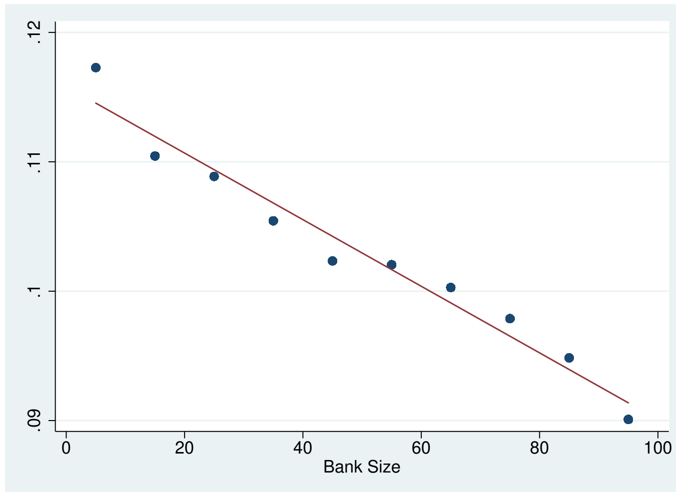
The capital-asset ratio is constructed as the average of the ratio of total bank equity (book equity) and total assets. The withdrawal uncertainty variable in computed as the standard deviation of the quarterly growth rate of total deposits after being winsorized at 1st and 99th percentiles in every year to ensure that uncertainty is not driven by large and sudden increments or decrements in total deposits. Also, I am dropping observations where the growth rate of equity is above 95th and below the 5th percentile since the volatility of equity induces volatility of deposits. Let D_{jt} denote total deposits from bank j in quarter t and $\frac{D_{jt}-D_{jt-1}}{D_{jt-1}}$ is the quarterly growth rate of deposits at quarter t and denote its winsorized value at period t as g_{jt} with mean $\bar{g}_j$. Then withdrawal uncertainty σ_j is defined as follows:

$$\sigma_j = \left(\sum_{t=T_{0j}}^{T_j} \frac{(g_{jt} - \bar{g}_j)^2}{T_j - T_{0j}} \right)^{\frac{1}{2}}$$

Figure 2·1 plots the relationship between the average bank capital-asset ratio and bank size and it is constructed as follows: First, group the bank size variable into ten equal-sized bins, then compute the mean of the bank capital-asset ratio and bank size within each bin. Finally, create a scatterplot of these means. This figure shows that this ratio decreases as bank size increases, which means that small banks have proportionally higher capital-asset ratios than big banks. Furthermore, it shows that big banks are closer to a given capital requirement constraint.

Figure 2·2 plots the relationship between withdrawal uncertainty σ_j and bank size

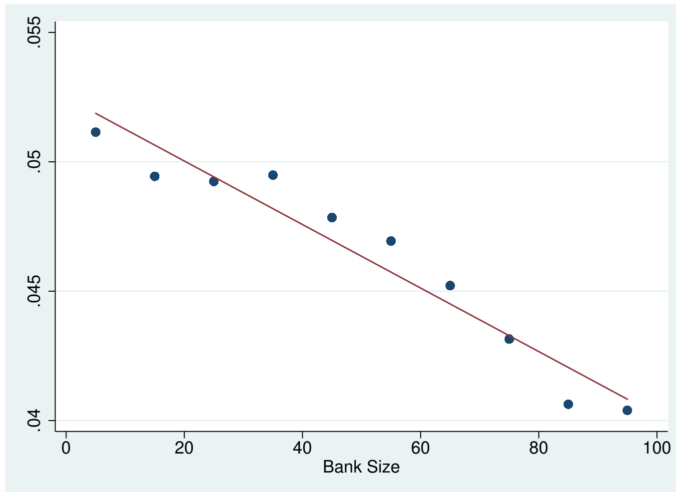
Figure 2·1: Bank Capital-Asset Ratio



Notes: This figure plots the relationship between average bank capital-asset ratio and bank size. It is constructed as follows: First, group the bank size variable into ten equal-sized bins, then compute the mean of the bank capital-asset ratio and bank size within each bin. Finally, create a scatterplot of these means. It can be noticed that small banks have a capital-asset ratio around 11.5% while for big banks this ratio is around 9%.

and it is constructed using ten equal-sized bins of bank size as in figure 1. This figure shows that small banks have proportionally larger withdrawal risk. A withdrawal uncertainty of 0.05 indicates that most of the variation in the deviation from its mean of the quarterly growth rate deposits is between 10% and -10%. This means for example that if a given bank with zero mean in the quarterly growth rate of deposits receives a withdrawal shock of one standard deviation, its deposits will decrease by 5% in that quarter.

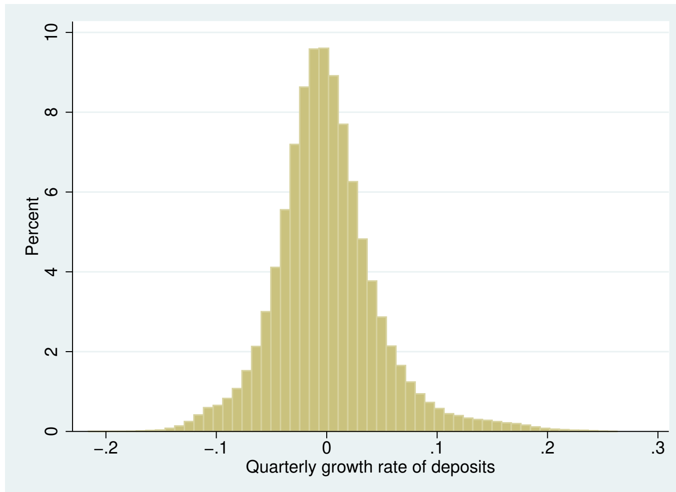
Figure 2·2: Withdrawal risk



Notes: This figure plots the relationship between withdrawal risk σ_j and bank size. It is constructed as follows: First, group the bank size variable into ten equal-sized bins, then compute the mean of withdrawal uncertainty and bank size within each bin. Finally, create a scatterplot of these means. Small banks have a withdrawal risk of around 5% while for big banks this variable is around 4%.

Figure 2·3 plots the histogram of the demeaned quarterly growth rate of deposits. The figure shows that the maximum withdrawal shock is far away from the -1 bound, in this case this is around -0.2, which means that the most negative withdrawal shock is a decrease in the quarterly growth rate deposits of 20%, which occurs with probability zero.

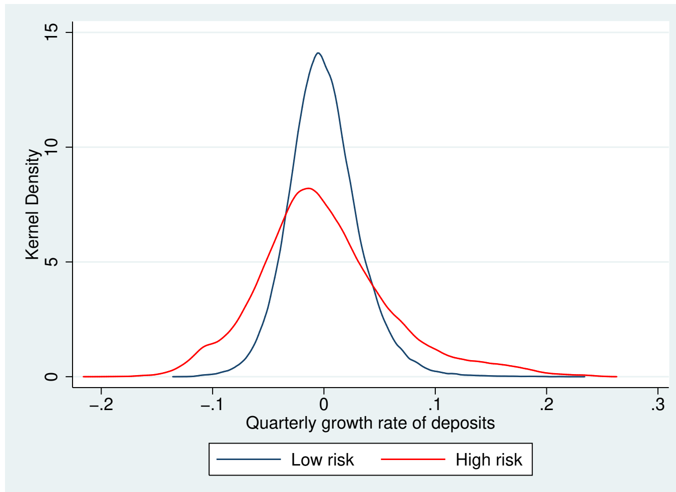
Figure 2·3: Histogram of quarterly growth rate of deposits



Notes: This figure plots the histogram of the demeaned quarterly growth rate of deposits after being winsorized at 1st and 99th percentiles every year.

Figure 2·4 shows that banks above the 50th percentile of σ_j face a fat-tailed distribution. We can also observe that the distribution is asymmetric for banks with higher risk since an outflow of deposits is relatively more likely than an inflow of deposits. Moreover, the largest possible inflow shock is similar for both types of banks while the largest possible outflow shock is very different, around -22% for high-risk banks and around -14% for low-risk banks.

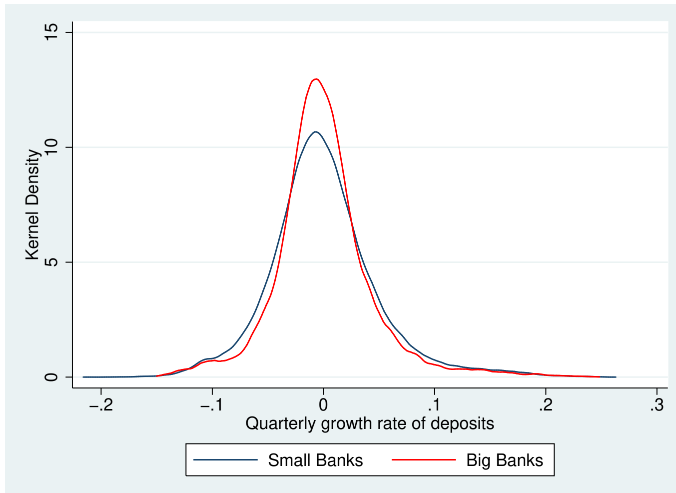
Figure 2·4: Density of quarterly growth rate of deposits by risk



Notes: This figure plots the density of the demeaned quarterly growth rate of deposits after being winsorized at 1st and 99th percentiles every year by type of uncertainty. Banks with low risk are those below the 50th percentile of σ_j and banks with high risk are the ones above the 50th percentile.

Figure 2·5 shows the density of the demeaned quarterly growth rate of deposits by bank size. Big banks are the ones above the 95th percentile using the size variable S_j and small banks are the ones below the 75th percentile. Although Figure 2·2 shows a negative relationship between bank size and withdrawal risk, it is not the case that only small banks face riskier distributions. Figure 2·5 shows that small banks are subject to a similar ratio of inflow shocks but they are more exposed to withdrawal shocks, which means that small banks are more likely to receive a shock that reduces their deposits in a given quarter. The probability of receiving a shock that increases its deposits in a given quarter is similar for both types of banks.

Figure 2·5: Density of quarterly growth rate of deposits by bank size



Notes: This figure plots the density of the demeaned quarterly growth rate of deposits after being winsorized at 1st and 99th percentiles every year by bank size. Big banks are the ones above the 95th percentile using the size variable S_j and small banks are the ones below the 75th percentile.

Table 2.1 shows the percentiles from the density of the demeaned quarterly growth rate of deposits. Small banks are more likely to receive a withdrawal shock (a shock that induces an outflow of deposits). For instance, with a probability of 10% small banks are hit by a shock that reduces their total deposits by 5.5% while deposits from big banks decrease by 4.6%. Therefore, small banks will need a proportionally higher reserve-deposit ratio in equilibrium. In general, banks with a higher withdrawal risk are hit by larger withdrawal shocks and will need to hold a proportionally higher reserve-deposit ratio.

Table 2.1: Quarterly growth rate of deposits by bank size and risk

Percentile	Small Banks	Big Banks	Low-risk	High-risk
1%	-0.112	-0.111	-0.072	-0.123
5%	-0.074	-0.065	-0.049	-0.091
10%	-0.055	-0.046	-0.038	-0.069
15%	-0.043	-0.036	-0.031	-0.056
25%	-0.029	-0.023	-0.021	-0.038
50%	-0.003	-0.003	-0.002	-0.006
Std. Dev.	0.049	0.044	0.032	0.060

Notes: This table shows the demeaned quarterly growth rate of deposits after being winsorized at 1st and 99th percentiles every year. Big banks are above 95th percentile and small banks are below 75th using Sj.

Numbers in Table 2.1 can also be interpreted as equilibrium reserve-deposit ratios. For example, if banks find it optimal to have an exposure of 15% to withdrawal risk, this implies a reserve-deposit ratio of 3.1% for low-risk banks and 5.6% for high-risk banks. If there is a shock such that banks find it optimal to reduce its exposure to withdrawal risk to 5%, then the new reserve-deposit ratios would be 4.9% for low-risk banks and 9.1% for high-risk banks. This implies that the sensitivity of high-risk banks to a shock is proportionally higher given that the increment in the reserve-deposit ratio is 1.8% for low-risk banks and 3.5% for high-risk banks. If banks find it optimal to move from an exposure of 5% to only 1%, the increment in the reserve-deposit ratio is 3.2% for high-risk banks and 2.3% for low-risk banks.

If we compare small and big banks, an exposure of 15% to withdrawal risk implies a reserve-deposit ratio of 4.3% for small banks and 3.6% for big banks. If there is a shock such that banks find it optimal to reduce its exposure to withdrawal risk to 5%, then the new reserve-deposit ratios would be 7.4% for small banks and 6.5% for big banks. This implies that the sensitivity of small banks to shocks is proportionally higher given that the increment in the reserve-deposit ratio is 3.1% for small banks and 2.9% for big banks. The opposite holds if banks reduce its exposure from 5%

to 1%, where small banks increase the reserve-deposit ratio by 3.8% while big banks increase this ratio by 4.6%. Hence, big banks have lower volatility of deposits on average but during periods where it is optimal to have a low exposure to withdrawal risk, big banks' reserve-deposit ratio will be more sensitive to shocks.

2.3 Model

In this section, I develop a simple general equilibrium model to understand the effect of heterogeneous withdrawal risk on the monetary policy transmission mechanism. First, I model the behavior of banks and derive some partial equilibrium results. Next, workers and entrepreneurs are introduced to derive general equilibrium results.

2.3.1 Banks

In this economy, there is a continuum of risk-neutral banks indexed by j. They receive an endowment with a fixed nominal value N_j. Each bank issues deposits D_j, creates loans L_j and demands reserves M_j from the Central Bank. Deposits, loans and reserves earn nominal interest rates i^d, i^l and i^m, respectively. Deposits are subject to an idiosyncratic withdrawal shock ω_j with support $[\underline{\omega}\,,\overline{\omega}]$ where $-\underline{\omega}$ is the largest deposit inflow shock and $\overline{\omega}$ is the largest withdrawal shock. There is a capital requirement constraint $D_j \leq \phi N_j$ where N_j determines the scale of bank j and ϕ is an upper limit to bank leverage $\frac{D_j}{N_j}$. Reserves are issued by the Central Bank and they serve as the numeraire in this economy.

This economy exists for two periods $t = 1, 2$. There are two stages in period $t = 1$. In the first stage, banks sell its asset endowment N_j to the Central Bank in exchange for reserves and decide optimal amounts of deposits, loans, and reserves. In the second stage, an idiosyncratic withdrawal shock occurs and banks can only use reserves or discount window loans to meet this shock. Reserves increase if there is an inflow of deposits and decrease if there is an outflow of deposits. Banks demand discount window loans from the Central Bank at a high interest rate $i^w > i^l$ if the outflow of deposits is higher than the amount of reserves. Intuitively, withdrawal risk and a high interest rate $i^w > i^l$ induce an ex-ante strictly positive demand for reserves and the capital requirement induces a limit in the issuance of deposits by banks. In period $t = 2$, banks repurchase their asset from the Central Bank and consume their net worth.

A withdrawal of deposits from one bank implies a transfer of reserves from this bank to the receptor bank. Assuming that one unit of deposits is settled with one unit of reserves, profits from Bank j, Z_j take the following expression:

$$\begin{aligned} Z_j = (1+i^l)L_j + (1+i^m)M_j - (1+i^d)D_j + (i^d - i^m)\omega_j D_j \\ -(i^w - i^m)\Big[\omega_j D_j - M_j\Big]\mathbb{1}\{\omega_j > e_j\} + T_j \end{aligned} \tag{2.1}$$

where $e_j = \frac{M_j}{D_j}$ is the reserve-deposit ratio for bank j and T_j are transfers/taxes from the Central Bank. Notice that if there is an outflow of deposits, banks save the interest paid on deposits and lose the interest paid on reserves they transfer to another bank. Moreover, if ω_j is large enough banks need to demand discount window loans and pay an extra cost given by the spread $i^w - i^m$.

After bank j decides the optimal amount of deposits, loans and reserves, a withdrawal/inflow of deposits occurs such that for bank j the new level of deposits is $D_j(1-\omega_j)$ and reserves after shock are $M_j - \omega_j D_j$. A positive value $\omega_j > 0$ indicates a withdrawal of deposits and a negative value $\omega_j < 0$ indicates an inflow of deposits. Hence, if reserves are larger than the withdrawal of deposits, i.e. $M_j \geq \omega_j D_j$, bank j receives interest on reserves $(1+i^m)(M_j - \omega_j D_j)$ and pays interest on deposits $(1+i^d)D_j(1-\omega_j)$, which implies a net payment of $(1+i^m)M_j-(1+i^d)D_j+(i^d-i^m)\omega_j D_j$.

If the withdrawal of deposits is strictly larger than the amount of reserves, i.e. $M_j < \omega_j D_j$, bank j receives additional funds $W_j = \omega_j D_j - M_j$ from the Central

Bank at an interest rate i^w. Bank j pays interest on deposits $(1+i^d)D_j(1-\omega_j)$ and interest on discount window loans $(1+i^w)W_j$. In this scenario, the net payment is $(1+i^m)M_j-(1+i^d)D_j+(i^d-i^m)\omega_j D_j-(i^w-i^m)W_j$.

Bank j maximizes expected profits Z_j subject to a balance sheet and a capital requirement constraint.

$$\max_{L_j,D_j,M_j}\Bigg\{(1+i^l)L_j+(1+i^m)M_j-(1+i^d)D_j+T_j - \overbrace{(i^w-i^m)\int_{\frac{M_j}{D_j}}^{\overline{\omega}}(\omega D_j-M_j)f_j(\omega)d\omega}^{\textbf{expected cost of reserve deficiency}}\Bigg\} \tag{2.2}$$

$$s.t.\quad M_j+L_j=D_j+N_j$$

$$D_j\leq \phi N_j$$

The expected cost of reserve deficiency is inversely related to the reserve-deposit ratio held by banks. As this ratio increases, the expected cost decreases since higher reserves reduce the need for expensive borrowing from the Central Bank.

Banks consume their net worth in $t=2$, which implies that their consumption depends on realized profits. This implies that a large withdrawal of deposits can hit the non-negative lower bound of banker's consumption in the second period. This setting assumes that the withdrawal risk faced by banks is sufficiently low to guarantee an always non-negative consumption from bankers.[2]

[2]If this condition doesn't hold, the effects of monetary policy on lending are asymmetric. The reserve-deposit ratio can be insensitive to a reduction in the discount window loan rate or any other expansionary policy. Angrist et al. (2018) provide evidence of asymmetry in the response of economic activity in the U.S. to monetary policy interventions, in particular, they find that output and inflation are less sensitive to expansionary policies.

2.3.2 Partial Equilibrium

In this economy, banks are competitive and take interest rates on loans and deposits as given. I consider an equilibrium with heterogeneous banks in terms of bank equity and withdrawal risk. There are two types of banks, those with a low equity (small banks) and those with a high equity (big banks).

Assumption 1: The cumulative distribution of the withdrawal shock for big banks F_b second-order stochastically dominates the distribution for small banks F_s. Both distributions satisfy the following properties:

$$\int_{\underline{\omega}}^{\overline{\omega}} [F_s(\omega) - F_b(\omega)] d\omega = 0 \tag{2.3}$$

$$\int_{\underline{\omega}}^{y} [F_s(\omega) - F_b(\omega)] d\omega > 0 \quad \forall \underline{\omega} < y < \overline{\omega} \tag{2.4}$$

where small banks are indexed by s and big banks are indexed by b. Depositors use deposits to make payments such that deposits are transferred from one bank to another and banks must settle these transactions with reserves (Bianchi and Bigio, 2022). We can think of deposit withdrawal shocks as shocks to bank branches. Then, small banks have fewer bank branches than big banks, which induces a heterogeneous exposure to withdrawal shocks at the bank level.

Bank Behavior

Bank j solves optimization problem (2.2) taking as given interest rates on loans, deposits, reserves, and discount window loans. Then, first order conditions for reserves M_j and deposits D_j are the following:

$$(i^l - i^m) = (i^w - i^m) \int_{e_j}^{\overline{\omega}} f_j(\omega) d\omega \tag{2.5}$$

$$(i^l - i^d) = (i^w - i^m) \int_{e_j}^{\overline{\omega}} \omega f_j(\omega) d\omega + \lambda_j \tag{2.6}$$

where λ_j is the Lagrange multiplier associated to the capital requirement constraint and f_j is the density of the withdrawal shock ω_j.

Equation (2.5) is the first-order condition for reserves. An additional unit of reserves reduces the expected cost of borrowing from discount window loans, where $i^w - i^m$ is the net borrowing cost of one unit of discount window loans and $\int_{e_j}^{\overline{\omega}} f_j(\omega) d\omega$ is the probability that bank j borrows from the Central Bank given a reserve-deposit ratio e_j. An increase in reserves reduces the amount bank j borrows from the Central Bank in case a sufficiently large withdrawal of deposits occurs. This marginal benefit must be equal to the marginal cost of holding an additional unit of reserves, which is given by $i^l - i^m$ since bank j needs to reduce loans in one unit to increase reserves in one unit.[3]

Equation (2.6) is the first-order condition for deposits. An additional unit of deposits receives a net payment of $i^l - i^d$ since one unit of deposits increases lending in one unit. The marginal cost of issuing deposits is given by the spread $i^w - i^m$ and the expected demand for discount window loans. Given a withdrawal shock, a larger amount of deposits induces a larger outflow of deposits, which increases the demand for discount window loans and, therefore, the cost of issuing deposits.

Bank behavior from small and big banks implied by equations (2.5) and (2.6) is summarized in the following result:

Result 1: *Under Assumption 1, the following must hold:*

(i) The reserve-deposit ratio is higher for small banks, i.e. $e_s > e_b$

(ii) The capital requirement constraint is binding for big banks, i.e. $D_b = \phi N_b$ *and* $D_s \leq \phi N_s$.

[3] In equilibrium $i^l > i^m$. Otherwise, banks would prefer to hold reserves instead of loans.

(iii) Deposits for small banks is $D_s = \frac{1}{\psi}\big[D - (1-\psi)\phi N_b\big]$ *where* D *is the aggregate deposit supply and* ψ *is the share of small banks in the economy.*
(iv) The supply of loans is $L_s = N_s + (1-e_s)D_s$ *and* $L_b = (1+\phi(1-e_b))N_b$*.*

The marginal cost of holding reserves is the spread between the interest rates on loans and reserves, which is the same for all banks. Therefore, the expected marginal cost of holding reserves is the same across banks. As a result, small banks need a higher reserve-deposit ratio given that they face a higher withdrawal uncertainty. Result $1(i)$ indicates that the expected marginal cost of holding reserves is the same across banks only if small banks hold a higher reserve-deposit ratio.

The marginal benefit of deposits is the spread $i^l - i^d$ which is is the same for all banks, then the expected marginal cost of issuing deposits should be also the same for all banks. But a higher withdrawal uncertainty implies a higher expected demand for discount window loans. This implies that for any level of deposits, big banks always have a lower marginal cost. Hence, issuing deposits is always profitable for big banks (as long as small banks exist) and they would issue deposits up to their upper limit for leverage. This is result $1(ii)$ and it implies that big banks are insensitive to changes in the aggregate deposit supply.

Result $1(iii)$ follows by equilibrium in the deposit market. The deposit demand by small banks should be the difference between the aggregate deposit supply and the amount of deposits demanded by big banks. Moreover, from result $1(iv)$, lending by big banks is completely determined by their reserve-deposit ratio and net worth. Hence, changes in the aggregate deposit supply only affects lending by small banks.

Banking Sector Equilibrium

A partial equilibrium analysis of loan and deposit markets requires a loan demand and a deposit supply. Hence, the following two assumptions are introduced:

Assumption 2: The central bank sets a policy rate i such that the interest rate on reserves is $i^m = i + s^m$ and the interest rate on discount window loans is $i^w = \theta i + s^w$ where $\theta > 1$, and $s^w > 0$ and $s^m < 0$ are set by the Central Bank[4].

Assumption 3: Aggregate loan demand is $L = \Phi_l(\frac{1+i^l}{1+\pi})^{-\nu_l}$ and aggregate deposit supply is $D = \Phi_d(\frac{1+i^d}{1+\pi})^{\nu_d}$ where $\Phi_d, \Phi_l, \nu_l, \nu_d$ are strictly positive coefficients and π denotes inflation.

The assumption that small banks face a riskier distribution of withdrawal shocks implies a higher *level* of the reserve-deposit ratio for small banks. Nevertheless, this assumption does not imply that small banks have higher *sensitivity* to monetary policy. Using equation (2.5), this effect can be illustrated in the following equation:

$$\mathrm{d}e_s = \left[\frac{f_b(e_b)}{f_s(e_b)}\frac{f_s(e_b)}{f_s(e_s)}\right]\mathrm{d}e_b \tag{2.7}$$

which is the relationship between the change in the reserve-deposit ratio of small and big banks. The expression in brackets can be either greater or lower than one but under regular conditions, it should be greater than one. For example, the ratio $\frac{f_s(e_b)}{f_s(e_s)}$ is greater than one if we assume that $f_s(.)$ is decreasing[5] given that $e_s > e_b$. Nevertheless, the ratio $\frac{f_b(e_b)}{f_s(e_b)}$ is greater than one only if e_b is low enough.

Only in the case that e_b is high enough,[6] big banks are more responsive than small banks. For example, if the spread between the loan rate and the interest rate on reserves $i^l - i^m$ is small enough, banks might find it optimal to hold very high reserve-deposit ratios such that big banks are more responsive to shocks than small

[4]In general equilibrium, if $\theta \leq 1$ a higher policy rate increase lending and reduces the spread between the loan and deposit rates.

[5]Over strictly positive values of ω

[6]It is not sufficient that the ratio $\frac{f_b(e_b)}{f_s(e_b)}$ takes values below 1 to generate a regime with proportionally more responsive big banks.

banks.

The model generates endogenously two regimes, a regime with relatively low reserve-deposit ratios (regime I) and a regime with high reserve-deposit ratios (regime II). Small banks are more responsive to monetary policy in the first regime and big banks are more sensitive in the second one.

Result 2: *Under Assumptions 1-3, regime I and if θ is sufficiently high, after an increase in the policy rate i:*

(i) The reserve-deposit ratio increases for small and big banks and the response is larger for small banks, i.e. $\frac{\partial e_s}{\partial i} > \frac{\partial e_b}{\partial i}$

(ii) The interest rate on loans i^l and the interest rate on deposits i^d increase and total loans decrease.

(iii) The spreads $i^l - i^m$, $i^l - i^d$ and $i^d - i^m$ increase.

(iv) The total amount of deposits D increases.

(v) The decline in lending is larger in the case of small banks as long as deposits D do not increase enough, i.e. $\frac{\partial L_s/L_s}{\partial i} < \frac{\partial L_b/L_b}{\partial i}$

(vi) The response of i^l and L increases as withdrawal uncertainty increases, i.e. $\frac{\partial L/L}{\partial \sigma_j \partial i} < 0$ *and* $\frac{\partial i^l}{\partial \sigma_j \partial i} > 0$ *where σ_j is a measure of withdrawal uncertainty for bank j.*

Notice that an increase in the policy rate increases the interest rate on reserves and the interest rate on discount window loans. This assumption is consistent with a standard increase in the policy rate by the central bank in an economy with external financing frictions. A standard increase of 1% in the target rate (Fed funds rate), increases the interest rate on reserves by 1% and the interest rate on discount window loans by 1%. However, the interest rate on non-deposit funding for banks increases by more than 1% when external financing is costly (Kashyap and Stein, 1995). Consistent with external financing frictions, I assume that the interest rate on discount window loans increases by more than 1% after a 1% increase in the target rate.

An increase in the interest rate on reserves induces banks to increase their demand for reserves and reduce lending, which implies a higher interest rate on loans and a proportionally larger increase in the reserve-deposit ratio for small banks[7]. A higher reserve-deposit ratio reduces the expected cost of issuing an additional unit of deposits which tends to reduce the spread between loans and deposits but the increase in the spread between the discount window rate and the interest rate on reserves[8] has an opposite and stronger effect. Also, a higher interest rate on reserves increases the marginal benefit of deposits for banks, which increases deposit rates and the aggregate deposit supply.

A higher withdrawal uncertainty increases the expected cost of borrowing and banks find it optimal to hold a higher reserve-deposit ratio. If only small banks face a higher uncertainty, they reduce lending even more after a contractionary policy shock and there is a higher increase in the interest rate on loans. The higher spread between loans and reserves increases lending by big banks. Hence, the gap in the response between small and large banks increases. If only big banks face a higher uncertainty, then the higher spread between loans and reserves increases lending by small banks and the gap between small and big banks decreases. In both cases, there is a larger contraction in aggregate lending and a higher increase in the interest rate on loans.

2.3.3 General Equilibrium

In this economy, output in period $t = 1$ is given as endowment while output in period $t = 2$ in produced by entrepreneurs. This economy is populated by workers and entrepreneurs. Workers supply labor and deposits in period $t = 1$ and consume in

[7]This holds under regular conditions and regime I

[8]It follows from the assumption $\theta > 1$. Without this assumption, an increase in the interest rate on reserves is expansionary given that $i^l - i^d$ decreases after an increase in the reserve-deposit ratio.

periods $t = 1, 2$. Entrepreneurs need loans to finance working capital and produce, they demand loans and labor in period $t = 1$ and produce and consume in $t = 2$.

Timing

At the beginning of $t = 1$, banks receive an endowment with value N_j. This endowment is sold to the Central Bank in exchange for reserves. Then, banks choose optimal amount of reserves, deposits and loans in $t = 1$. Entrepreneurs demand labor to produce and workers receive an endowment of goods $\overline{y}_1$, supply labor, receive a wage, supply deposits and consume. Then, an idiosyncratic withdrawal shock is realized and some banks demand discount window loans from the Central Bank. At the beginning of $t = 2$, workers receive an interest rate on deposits, entrepreneurs pay an interest on loans and banks receive an interest rate on reserves and pay an interest rate on discount window loans. Also, banks repurchase its endowment. Entrepreneurs produce goods and consume, workers use deposits/savings to consume, the Central Bank make profits from monetary operations in $t = 1$, which are transferred to banks. Banks consume its endowment and buy goods using its profits from financial intermediation in $t = 1$ and transfers received from the Central Bank (see figure 6).

Entrepreneurs

There is a continuum of identical entrepreneurs that produce a single good y_2 using labor n_1 subject to a technology $y_2 = A_2 n_1^{\alpha}$ where A_2 is a productivity shock and $\alpha < 1$. They maximize profits given by $c_{e2} = A_2 n_1^{\alpha} - (\frac{1+i^l}{1+\pi}) w_1 n_1$ where w_1 is the real wage and c_{e2} is the consumption by entrepreneurs in period $t = 2$. They do not consume in period $t = 1$. Then, from first order condition $(\frac{1+i^l}{1+\pi}) w_1 = \alpha \frac{y_2}{n_1}$.

In period $t = 1$, firms demand loans from banks to finance working capital. Labor generates output in period $t = 2$ and entrepreneurs sell goods and consume in this

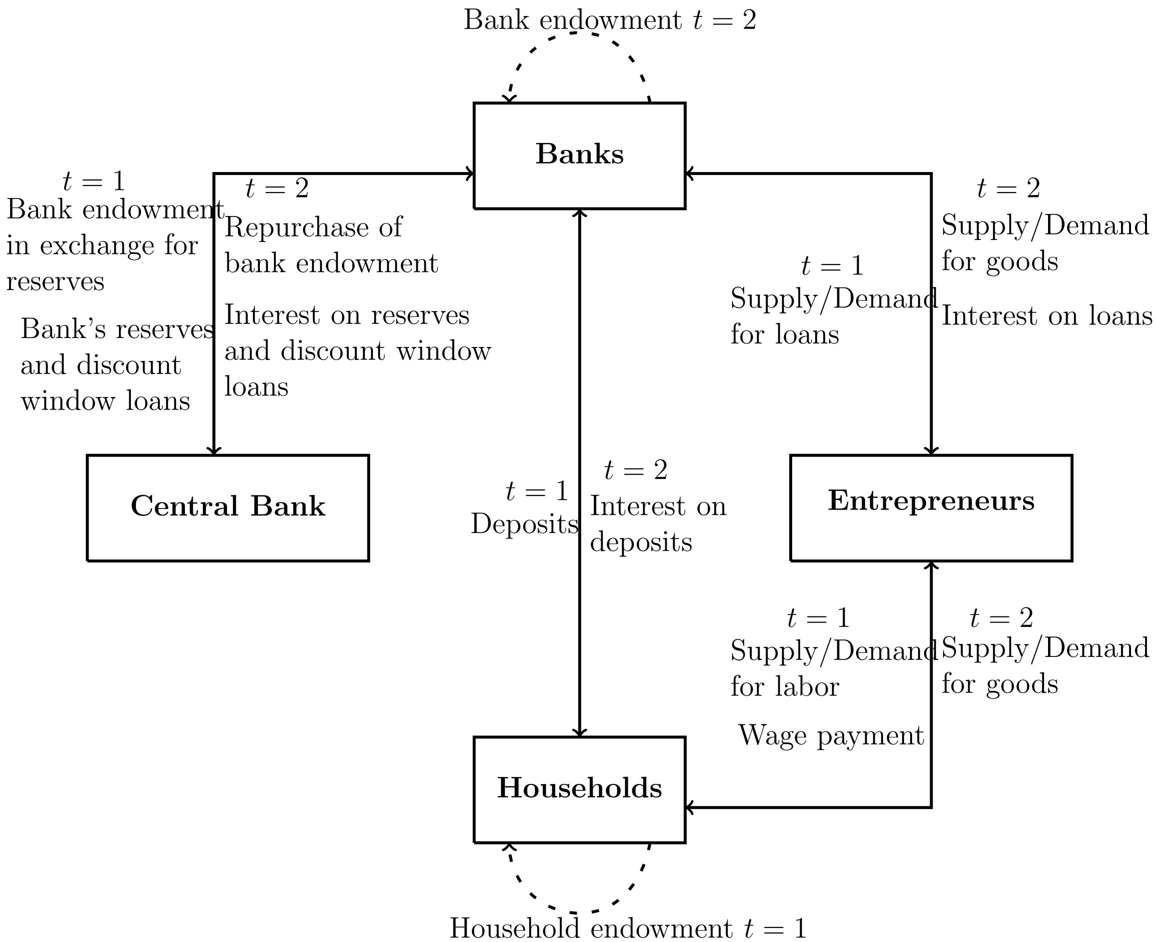


Figure 2·6: Diagram of the model

period. The optimal demand for nominal loans is given by:

$$L_e = p_1(\alpha A_2)^{\frac{1}{1-\alpha}} \left(\frac{1+i^l}{1+\pi}\right)^{\frac{-1}{1-\alpha}} w_1^{\frac{-\alpha}{1-\alpha}} \tag{2.8}$$

Workers

Workers consume in periods $t = 1, 2$. They supply labor to firms and receive a wage in period $t = 1$. They also supply a strictly positive amount of deposits at the end of this period to consume in the following given that they do not receive a wage in period $t = 2$. Deposits are equal to savings in this model and workers use savings from $t = 1$ to consume in period $t = 2$. They also receive an endowment of goods in $t = 1$, denoted by $\bar{y}_1$ and solve the following problem:

$$\max_{c_{w1}, D_w} \left\{ u(c_{w1}) - v(n_{w1}) + \beta u(c_{w2}) \right\} \tag{2.9}$$

$$s.t. \quad c_{w1} + \frac{D_w}{p_1} = w_1 n_{w1} + \bar{y}_1$$

$$c_{w2} = \frac{1+i^d}{1+\pi} \frac{D_w}{p_1}$$

where c_{wt} is real consumption by workers in period t, w_t is real wage in period t, D_w is the nominal deposit supply, and n_{wt} is labor supply by workers in period t, and π denotes the inflation rate from $t = 1$ to $t = 2$. I assume a utility function $u(c) = \frac{c^{1-\sigma^c}}{1-\sigma^c}$ for consumption and disutility function $v(n) = \chi \frac{n^{1+\gamma}}{1+\gamma}$ for labor. The optimal demand for deposits is:

$$\frac{D_w}{p_1} = \frac{w_1 n_{w1} + \bar{y}_1}{1 + (\frac{1+i^d}{1+\pi})^{-\frac{1}{\sigma^c}+1} (\beta)^{-\frac{1}{\sigma^c}}} \tag{2.10}$$

I assume that $\sigma^c < 1$ to have a positive relationship between the interest rate on deposits and deposit supply. The optimal labor supply in period $t = 1$:

$$w_1 = \chi (c_{w1})^{\sigma^c} (n_{w1})^{\gamma} \tag{2.11}$$

I assume that $\gamma > 0$ so that there is a positive relationship between labor supply and real wage.

Central Bank

In this economy, banks need reserves to create loans. I assume that bank j is endowed with an asset with nominal value N_j, which is sold to the Central Bank in exchange for reserves M_j. Hence, the Central Bank issues reserves M to buy bank's endowment $N = \int N_j \, dj$ and hold discount window loans W in period $t = 1$. In period $t = 2$, Central Bank generates profits or looses from monetary operations at $t = 1$, which are redistributed as transfers/taxes T_2 to bankers such that $T_2 = i^w W - i^m \tilde{M}$[9], where $\tilde{M}$ is the amount of reserves after withdrawal shocks are realized.

Equilibrium

The market clearing conditions for every market are the following:

In the goods market,

$$\bar{y}_1 = c_{w1} \tag{2.12}$$

$$y_2 + \frac{N}{p_2} = c_{w2} + c_{e2} + c_{f2} \tag{2.13}$$

where $\bar{y}_1$ is an endowment of goods in period $t = 1$, c_{wt} denotes the consumption of workers in period t, c_{e2} denotes the consumption of entrepreneurs in $t = 2$, and c_{f2} denotes the consumption of bankers in $t = 2$. I assume that bankers and entrepreneurs do not consume in $t = 1$. Notice that consumption by bankers, c_{f2}, equals aggregate profits from banks $Z_{banks} = \int Z_j \, dj$ [10]. The following holds:

$$c_{f2} = \frac{Z_{banks}}{p_2} = \frac{(i^l - i^d)}{1 + \pi} \frac{D}{p_1} + \frac{N}{p_2} \tag{2.14}$$

In the labor market,

$$n_1 = n_{w1} \tag{2.15}$$

[9]Since the Central Bank does not consume in any period, profits or looses must be redistributed to other agents. In this case, the redistribution to bankers is the same as a redistribution to entrepreneurs. If these are redistributed to workers, equilibrium allocations are different. The redistribution to bankers can alter the equilibrium allocation of quantities and prices only if $\bar{\omega}$ and i^w are sufficiently high.

[10]I assume that uncertainty of withdrawal shocks is such that profits from bankers are positive even when they receive the largest withdrawal shock.

where

$$n_{w1} = \chi^{-\frac{1}{\gamma}} w_1^{\frac{1}{\gamma}} (c_{w1})^{-\frac{\sigma^c}{\gamma}} \quad \text{and} \quad n_1 = (\alpha A_2)^{\frac{1}{1-\alpha}} \left[\Big(\frac{1+i^l}{1+\pi} \Big) w_1 \right]^{\frac{-1}{1-\alpha}} \tag{2.16}$$

In the deposit market, the aggregate deposit supply by workers equals the aggregate demand by banks.

$$\psi D_s + (1-\psi) D_b = D_w = \frac{p_1 w_1 n_{w1} + p_1 \bar{y}_1}{1 + (\frac{1+i^d}{1+\pi})^{-\frac{1}{\sigma^c}+1} (\beta)^{-\frac{1}{\sigma^c}}} \tag{2.17}$$

Similarly in the loan market,

$$\psi L_s + (1-\psi) L_b = L_e = p_1 (\alpha A_2)^{\frac{1}{1-\alpha}} \Big(\frac{1+i^l}{1+\pi} \Big)^{\frac{-1}{1-\alpha}} w_1^{\frac{-\alpha}{1-\alpha}} \tag{2.18}$$

where,

$$D_b = \phi N_b \quad \text{and} \quad D_s = \frac{1}{\psi} \big[D - (1-\psi) \phi N_b \big] \tag{2.19}$$

$$L_b = (1 + \phi(1-e_b)) N_b \quad \text{and} \quad L_s = N_s + (1-e_s) D_s \tag{2.20}$$

Result 3: *In general equilibrium, under assumptions 1-2, regime I and if θ is sufficiently high, an increase in the policy rate i:*

(i) Decreases output y_2, loans L and deposits D due to higher loan-deposit spread.

(ii) The decline in lending is unambiguously larger in the case of small banks. i.e. $\frac{\partial L_s/L_s}{\partial i} < \frac{\partial L_b/L_b}{\partial i}$

(iii) The effect on i^l, L and output y_2 increases as withdrawal uncertainty increases, i.e. $\frac{\partial L/L}{\partial \sigma_j \partial i} < 0$, $\frac{\partial y_2/y_2}{\partial \sigma_j \partial i} < 0$ *and* $\frac{\partial i^l}{\partial \sigma_j \partial i} > 0$ *where σ_j is a measure of withdrawal risk for bank j. Moreover, its effect is asymmetric.*

(iv) Increases concentration in the market for loans.

An increase in the policy rate increases the spread between loans and deposits and reduces loan and labor demand, which reduces wages and deposit supply. Output and consumption in period $t = 1$ are fixed but the reduction in the demand for labor and wages leads to a contraction in deposits. However, given that the nominal interest rate on deposits increases, inflation needs to increase to reduce the real rate on deposits and satisfy the equilibrium condition[11] $L = D$. Therefore, the decline in lending is partly offset by an increase in inflation. Big banks are not sensitive to aggregate deposit supply, hence the decline in total deposits implies that there is an even larger decline in lending by small banks. Moreover, a weaker deposit supply dampens the decline in lending from large banks, increasing market concentration.

In general equilibrium, an increase in withdrawal risk has asymmetric effects on lending. An increase in the withdrawal risk from small banks increases the gap in the response between small and big banks after a contractionary monetary policy but an increase in the withdrawal risk from big banks does not necessarily reduce this gap. If only withdrawal risk from small banks increases, aggregate lending and deposits decrease, which increases the interest rate on loans. A higher i^l increases lending by big banks and the decline in aggregate deposits reduces lending by small banks even more. If only withdrawal risk from big banks increases, they decrease lending, which reduces output and deposits and increases i^l. A higher i^l induces small banks to lend more but a decline in aggregate deposits reduces its loan supply. Therefore, in this case, the gap in the response between small and large banks might not decrease. Moreover, this gap might increase if, for example, market shares from big banks are sufficiently high.

The responses of loan supply by small and big banks after an increase in the policy

[11]It follows from equation (2.12)

rate are the following:

$$\frac{\mathrm{d}L_s}{\mathrm{d}i}\frac{1}{L_s} = \frac{(1-e_s)D_s}{L_s}\Bigg[\underbrace{\left(\frac{\mathrm{d}D}{\mathrm{d}i}\frac{1}{D}\right)}_{\substack{\textbf{Aggregate}\\ \textbf{deposits}}}\underbrace{\left(\frac{1}{\psi}\frac{D}{D_s}\right)}_{\substack{\textbf{Inverse of}\\ \textbf{market share}}}\Bigg] - \frac{D_s}{L_s}\left(\frac{\mathrm{d}e_s}{\mathrm{d}i}\right) \tag{2.21}$$

$$\frac{\mathrm{d}L_b}{\mathrm{d}i}\frac{1}{L_b} = -\left(\frac{\mathrm{d}e_b}{\mathrm{d}i}\right)\frac{\phi N_b}{L_b} \tag{2.22}$$

An increase in the policy rate increases reserve-deposit ratios, which reduces lending. Under regular conditions, the reserve-deposit ratio for small banks is more sensitive, i.e. $\mathrm{d}e_s > \mathrm{d}e_b$. The implication is that small banks reduce lending proportionally more than big banks. Under Assumption 2, this policy is contractionary and reduces aggregate deposit supply, which further declines lending from small banks. Moreover, banks with lower deposit market share experience a larger contraction in lending.

Result 4: *An increase in the deposit market share of big banks:*

(i) reduces the responsiveness of the banking system to shocks.

(ii) increases the responsiveness of small banks to shocks.

General equilibrium effects amplify the effects of monetary policy on loans, output, and consumption through a weaker deposit supply and this induces an even larger response from small banks. Since big banks are not affected by fluctuations in aggregate deposits they partially offset the decline in aggregate loans by lending more, which reduces the responsiveness of the banking system. Moreover, a higher market share of big banks implies that there is a larger contraction in deposits from small banks after a contractionary monetary policy.

The aggregate effect of an increase in the policy rate on lending is:

$$\frac{\mathrm{d}L}{\mathrm{d}i}\frac{1}{L} = -\frac{1}{e_s}\left[\frac{\psi D_s}{D}\frac{\mathrm{d}e_s}{\mathrm{d}i} + \underbrace{\frac{(1-\psi)D_b}{D}}_{\substack{\textbf{market share}\\ \textbf{of big banks}}}\frac{\mathrm{d}e_b}{\mathrm{d}i}\right] \tag{2.23}$$

Given that, under regular conditions, the reserve-deposit ratio for small banks is more sensitive, i.e. $\mathrm{d}e_s > \mathrm{d}e_b$, a higher market share of big banks reduces the responsiveness of aggregate loans to monetary policy. By equation (2.21), a lower market share of small banks increases its sensitivity to monetary policy. Equation (2.23) also shows that a higher reserve-deposit ratio for small banks (keeping fixed withdrawal risk) reduces the aggregate response of lending.

A closed-form solution of the model

Assumption 4: The variable $1-\omega_j$ for bank j follows a log-normal distribution, i.e. $\log(1-\omega_j) \sim \mathcal{N}\big(\mu_j,\, \sigma_j^2\big)$ where $\mu_j = -\frac{1}{2}\sigma_j^2$.

Under Assumption 4, the main equations of the model are the following:

$$i^l - i^m = (i^w - i^m)\Phi\left(\frac{\log(1-e_s) + \frac{1}{2}\sigma_s^2}{\sigma_s}\right) \tag{2.24}$$

$$i^d - i^m = (i^w - i^m)\Phi\left(\frac{\log(1-e_s) - \frac{1}{2}\sigma_s^2}{\sigma_s}\right) - \lambda_s \tag{2.25}$$

$$\log(1-e_s) = \frac{\sigma_s}{\sigma_b}\log(1-e_b) + \frac{1}{2}\sigma_s(\sigma_b - \sigma_s) \tag{2.26}$$

$$\frac{D}{P_1} = \Theta^d (R^d)^{\varepsilon^d} \tag{2.27}$$

$$\frac{L}{P_1} = \Theta^l (R^l)^{-\varepsilon^l} \tag{2.28}$$

$$L = D \tag{2.29}$$

$$\psi N_s + (1-\psi)N_b = \psi e_s D_s + (1-\psi)e_b\phi N_b \tag{2.30}$$

$$D_s = \frac{1}{\psi}\big[D - (1-\psi)\phi N_b\big] \tag{2.31}$$

$$1+\pi = \frac{P_2}{P_1} = \frac{1}{P_1} \tag{2.32}$$

$$R^l = \frac{1+i^l}{1+\pi} \tag{2.33}$$

$$R^d = \frac{1+i^d}{1+\pi} \tag{2.34}$$

$$y_2 = \Theta^l (R^l)^{-\varepsilon_l+1} \tag{2.35}$$

There are 12 equations and 12 unknown variables. The endogenous variables[12] are: $e_s, e_b, D_s, D, L, P_1, \pi, R^l, R^d, i^l, i^d, y_2$.

The first three equations (2.24)-(2.26) define the reserve-deposit ratio for small and big banks and the nominal interest rate on deposits in the case $\lambda_s = 0$. This variable is the lagrange multiplier associated with the capital requirement constraint for small banks $D_s \leq \phi N_s$. If this constraint is non-binding, $\lambda = 0$. $\Phi(.)$ denotes the cumulative distribution function of a standard normal variable.

Equation (2.24) equates the marginal benefit of increasing reserves for small banks, which is the spread $i^w - i^m$ times the probability of a large withdrawal shock and its marginal cost given by the spread $i^l - i^m$ lost from reducing loans in one unit. A similar condition holds for big banks, which implies a positive relation between reserve-deposit ratios in both banks, expressed in equation (2.26).

Equation (2.25) equates the marginal cost of issuing deposits, $i^d - i^m$, and its marginal benefit, which is the probability of a large withdrawal shock multiplied by the spread $i^w - i^m$ times the expected demand for discount window loans. If the capital requirement constraint is not binding, $\lambda_s > 0$.

Equation (2.35) determines real output in the second period. In the model, output

[12] The list of variables is on the Appendix.

in the first period is an endowment and it is given. It appears in the parameters Θ^l and Θ^d. These parameters depend on a number of exogenous variables. Moreover, the parameters of the model implies that $\varepsilon_l > 1$, which means that an increase in the real rate of loans reduces aggregate output in the second period.

Equations (2.27), (2.28), and (2.29) are the deposit supply, the loan demand, and the equilibrium condition in the loan market, respectively. Equation (2.30) shows the equilibrium condition for reserves before a withdrawal shock is realized in the first period, where the nominal supply of reserves is fixed. After a withdrawal shock is realized, this supply is endogenous as it depends on the demand for discount window loans. Equation (2.31) defines the amount of deposits issued by small banks as the difference between the total deposit supply and the optimal demand of deposits from big banks.

Finally, equations (2.32)-(2.34) define inflation and the real rates of loans and deposits. Notice that the price level in the second period is fixed at one.

The main results of the previous sections hold in this specification. Under the assumption $\sigma_s > \sigma_b$, from equation (2.26), the reserve-deposit ratio is higher for small banks, the ones with a riskier distribution of withdrawal shocks. Also, if the reserve-deposit ratio for big banks increases by 1%, this ratio increases by more than 1% for small banks.

$$\mathrm{d}e_s = \left[\frac{\sigma_s}{\sigma_b} \left(\frac{1 - e_s}{1 - e_b} \right) \right] \mathrm{d}e_b \tag{2.36}$$

Equation (2.36) shows that the reserve-deposit ratio for small-risky banks is relatively *more sensitive* if the ratios are sufficiently small. Finally, in this model, bank size is measured in terms of equity, i.e. $N_b > N_s$

Bank heterogeneity and macroeconomic aggregates

Equation (2.23) provides an important link between the aggregate response of loans and the individual behavior of banks after an increase in the policy rate. This equation can be used to show that bank heterogeneity is relevant for the aggregate behavior of loans and, in general, for macroeconomic aggregates such as output, consumption, and inflation.

Result 5: *Under Assumption 4, bank heterogeneity reduces the aggregate effect of monetary policy on aggregate lending and output.*

The implication of bank heterogeneity for the aggregate behavior of loans after an increase in the interest rate on reserves is given by the following equation:

$$\frac{\mathrm{d}L}{\mathrm{d}i}\frac{1}{L} = \overbrace{\frac{\overline{e}}{e_s}}^{\substack{\textbf{macro}\\\textbf{exposure}\\<1}} \left[\overbrace{-\frac{1}{\overline{e}}\frac{\mathrm{d}\overline{e}}{\mathrm{d}i}}^{\substack{\textbf{Aggregate}\\\textbf{channel}}} + \overbrace{\overbrace{\frac{\overline{\sigma}-\tilde{\sigma}}{\overline{\sigma}}\frac{1}{\overline{e}}\frac{\mathrm{d}\overline{e}}{\mathrm{d}i}}^{\substack{\textbf{nonlinear risk}\\>0}} + \overbrace{\frac{\mathrm{cov}(\sigma_j-\tilde{\sigma},e_j)}{\overline{\sigma}(1-\overline{e})}\frac{1}{\overline{e}}\frac{\mathrm{d}\overline{e}}{\mathrm{d}i}}^{\substack{\textbf{heterogeneous risk}\\>0}}}^{\textbf{withdrawal risk heterogeneity channel}>0} \right] \tag{2.37}$$

$$\text{where:} \quad \overline{e} = \alpha e_s + (1-\alpha)e_b \quad , \quad \tilde{\sigma} = \alpha\,\sigma_s + (1-\alpha)\sigma_b < \overline{\sigma}$$

$$\text{and} \quad \overline{\sigma} \quad \text{solves:} \qquad \log(1-e_s) = \frac{\sigma_s}{\overline{\sigma}}\log(1-\overline{e}) + \frac{1}{2}\sigma_s(\overline{\sigma}-\sigma_s)$$

where α is the deposit market share of small banks. An increase in the policy rate reduces aggregate lending through four channels. The aggregate channel is the decline in lending predicted by a model with a representative bank. A higher policy rate increases reserves and reduces lending from the "average bank". The macroeconomic exposure channel reduces aggregate lending because in a model with heterogeneous banks only some banks (those with a bank leverage strictly below its upper limit) respond to changes in the aggregate deposit supply. A nonlinear risk channel also reduces the aggregate effect of monetary policy since the withdrawal risk estimate of the “average bank” is strictly higher than the weighted average of the individual withdrawal risk measures. Finally, banks with higher withdrawal risk hold higher reserve-deposit ratios, which attenuates the aggregate decline in lending.

2.4 Empirical evidence

In this section, I provide evidence of the importance of heterogeneity in withdrawal risk in the monetary transmission mechanism for the U.S. economy.

2.4.1 Data

I use bank-level quarterly data from U.S. Call Reports from 1994Q1 to 2006Q4 to avoid the financial crisis period[13]. During this period, the total number of commercial banks is 8,109[14]. The measure of withdrawal risk used in this section is the standard deviation of the quarterly growth rate of deposits after being winsorized at the 1st and 99th percentile every year. This is a bank-level variable and it is constant over time. I denote this measure as σ_j where j index banks. Bianchi and Bigio (2022) calibrate the standard deviation of withdrawal shocks to match the distribution of excess reserves and the resulting value was 0.12, a bit higher than the volatility of the quarterly growth rate of deposits in this dataset, which is 0.046.

2.4.2 Panel Estimation

I estimate the following panel regression:

$$\begin{aligned}\Delta y_{jt} = \alpha_t + \lambda_j + \beta_1 \Delta i_t^{mp} \times \sigma_j + \beta_2 \Delta i_t^{mp} \times \text{HHI}_{jt-1} + \beta_3 \text{HHI}_{jt-1} \\ + \beta_4 \Delta i_t^{mp} \times Ch_{jt-1} + \beta_5 Ch_{jt-1} + \beta_6 \Delta i_t^{mp} \times S_j\end{aligned} \tag{2.38}$$

[13]The model developed in the previous section assumes no aggregate uncertainty
[14]After dropping banks with less than 20 observations.

where j is an index for banks, t indexes over quarters, α_t is a time fixed effect, λ_j is a bank fixed effect, Δi_t^{mp} is a monetary policy shock from Nakamura and Steinsson (2018)[15]. The variable σ_j is a measure of withdrawal risk, and Δy_{jt} is the log change of loans or deposits between t and $t-1$. HHI_{jt-1} is a measure of concentration in the deposits market from Drechsler et al. (2017). Ch_{jt-1} is the cash-deposit ratio lagged one quarter and S_j is a measure of bank size that goes from 0 to 1.

Drechsler et al. (2017) construct a bank-level measure of deposit market concentration as the average of the concentration of bank branches, weighted by its share of the bank's deposits. This is a bank-level variable with annual frequency[16]. Bank liquidity is the lagged value of the cash-deposit ratio, denoted by Ch_{jt-1}. Cash is defined as "Cash and balances due from depository institutions". Bank size variable S_j from Section 2.2 is also included.[17]

Table 2.2 shows the estimates from regression (2.38) under four different specifications. The first one only uses withdrawal risk as a source of heterogeneity in the response of loans to monetary policy. The second one adds a deposit concentration measure denoted by HHI. The third one adds the cash-deposit ratio and the final specification also uses a bank size variable. The estimated coefficient for the withdrawal risk measure is highly significant with a value around 21, which means that an increase of 100 basis points in the Fed funds rate contracts lending from a bank with a withdrawal risk measure higher by 0.05 by an extra $21 \times 0.05 = 1.05\%$.

[15]The main results from this section hold if I use the first difference of the Fed funds target rate instead of this variable.

[16]In any given year the first two quarters use the bank-level HHI from the previous year and the last two quarters use the ones from the current year.

[17]I divided by 100, so that it takes values between 0 and 1 such as the cash-deposit ratio and HHI measure. The maximum value of σ_j is 0.148

Table 2.2: Loan growth and monetary policy in the U.S.

	Loan growth			
	(1)	(2)	(3)	(4)
$\Delta i_t^{mp} \times \sigma_j$	-21.655***	-21.510***	-20.875***	-22.563***
	[2.292]	[2.405]	[2.422]	[2.450]
HHI	N	Y	Y	Y
Cash-Deposit	N	N	Y	Y
Bank size	N	N	N	Y
Bank f.e.	Y	Y	Y	Y
Quarter f.e.	Y	Y	Y	Y
Observations	312,281	280,320	280,320	280,320
R-squared	0.1781	0.1733	0.1745	0.1746

Notes: This table shows the estimates from equation (2.38) for the log change of loans using quarterly data and a winsorized measure of withdrawal risk σ_j. First, I compute the quarterly growth rate of total deposits. Second, winsorize this variable at 1st and 99th percentiles every year. Third, compute the standard deviation of the winsorized quarterly growth rate of deposits for every bank and denote this value as σ_j. Fourth, use the winsorized log change of loans at 1st and 99th percentiles as the dependent variable. Fixed effects are denoted by f.e. and Yes is denoted by Y. Standard errors are denoted in brackets and clustered by bank. The use of control variables HHI, Cash-Deposit ratio and Bank size depends on the specification. Significance at 1% is denoted by ***.

Table 2.3 shows withdrawal risk measures by percentile. For example, a bank at the 10th percentile has a measure of withdrawal risk equal to 0.026, lower than a bank at the 90th percentile. The coefficients from Table 2 imply that a monetary policy shock that increases the Fed funds rate in 100 basis points reduce lending contemporaneously from banks at the 90th percentile by an extra 1% with respect to banks at the 10th percentile. This extra decline in lending is computed as the difference between withdrawal risk measures times the estimated coefficient, i.e. $21\times(0.074-0.026)=1.008$. Similarly, if we compare banks at the 25th and 75th percentile, a monetary shock that increases the Feds Funds rate in 1% decreases lending by an extra 0.5% for banks at the 75th percentile.

Table 2.3: Withdrawal risk by percentile

	Percentile			
	10%	25%	75%	90%
Withdrawal risk	0.026	0.033	0.056	0.074

Notes: This table shows estimates for σ_j by percentile. Procedure: Winsorize quarterly growth rate of deposits at 1st and 99th percentiles every year and calculate its standard deviation.

Table 2.4 shows that a monetary policy shock that increases the Fed funds rate by 100 basis points reduces deposits contemporaneously from banks at the 90th percentile by an extra 0.7-0.9% with respect to banks at the 10th percentile. Consistent with the model presented in the previous section, the decline in deposits is proportionally larger for banks with higher withdrawal risk due to higher *sensitivity* of the reserve-deposit ratio and *non-binding* leverage constraint (higher macroeconomic exposure).

Table 2.4: Deposit growth and monetary policy in the U.S.

	Deposit growth			
	(1)	(2)	(3)	(4)
$\Delta i_t^{mp} \times \sigma_j$	-14.025***	-17.389***	-19.538***	-17.038***
	[2.611]	[2.768]	[2.786]	[2.858]
HHI	N	Y	Y	Y
Cash-Deposit	N	N	Y	Y
Bank size	N	N	N	Y
Bank f.e.	Y	Y	Y	Y
Quarter f.e.	Y	Y	Y	Y
Observations	312,281	280,320	280,320	280,320
R-squared	0.1359	0.1308	0.1373	0.1375

Notes: This table shows the estimates from equation (2.38) for the log change of deposits using quarterly data and a winsorized measure of withdrawal risk σ_j. First, I compute the quarterly growth rate of total deposits. Second, winsorize this variable at 1st and 99th percentiles every year. Third, compute the standard deviation of the winsorized quarterly growth rate of deposits for every bank and denote this value as σ_j. Fourth, use the winsorized log change of loans at 1st and 99th percentiles as the dependent variable. Fixed effects are denoted by f.e. and Yes is denoted by Y. Standard errors are denoted in brackets and clustered by bank. The use of control variables HHI, Cash-Deposit ratio, and Bank size depends on the specification. Significance at 1% is denoted by ***.

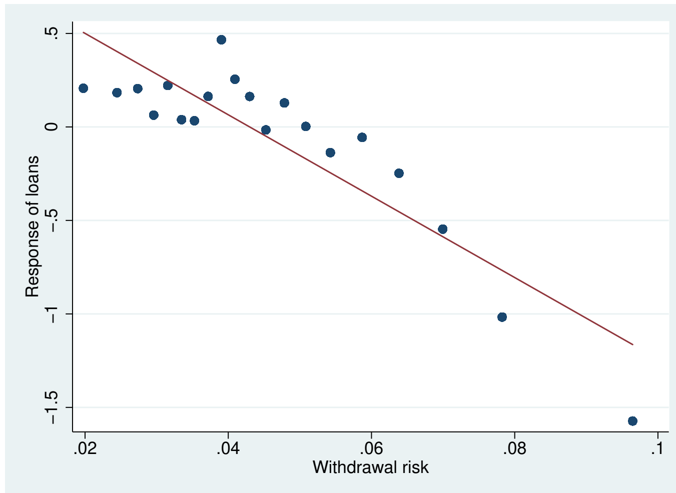
2.4.3 OLS estimation

I estimate the response of lending after a monetary policy shock by OLS for each bank.

$$\Delta L_{jt} = \lambda_j + \beta_j \Delta i_t^{mp} \tag{2.39}$$

where ΔL_{jt} is the log change of loans for bank j in period t and Δi_t^{mp} is the monetary policy shock from Nakamura and Steinsson (2018).

Figure 2·7: Response of loans and Withdrawal risk



Notes: This figure plots the relationship between the response of loans given by β_j from equation (2.39) and the withdrawal risk σ_j. It is constructed as follows: First, group the withdrawal risk variable into twenty equal-sized bins, then compute the mean of β_j and σ_j within each bin. Finally, create a scatterplot of these means. Banks with a higher withdrawal risk decrease lending by a larger proportion after a monetary tightening.

Figure 2.7 shows a binned scatterplot of β_j from regression (2.39) and a withdrawal risk measure σ_j. This figure shows that a higher withdrawal risk is associated with a larger contraction in lending after a contractionary monetary policy shock. This figure is consistent with the results from the model: The decline in lending after a contractionary monetary policy shock is proportionally larger for banks with a higher withdrawal risk. Moreover, if withdrawal risk is low enough, the capital requirement constraint becomes binding which implies that there is a group of banks for which the response in lending only depends on its reserve-deposit ratio. The figure shows that if withdrawal risk is low enough, the response of loans between banks is similarly low. This response is more negative as withdrawal risk increases, which is consistent with the model since the capital requirement constraint becomes non-binding if withdrawal risk is sufficiently high. Also, it is important to notice that this is only the contemporaneous effect of monetary policy on lending[18]. The effect of withdrawal risk is likely to be more important if we study the effect of monetary policy at longer horizons.

2.4.4 Sources of Bank heterogeneity

Table 2.5 shows the full set of estimates from equation (2.38). In this specification, there are four sources of heterogeneous bank response after a monetary shock: Withdrawal risk, Bank size, Cash-Deposit ratio, and Bank concentration.

Table 2.5: Sources of heterogeneous behavior in the U.S.

	Loan growth	Deposit growth
$\Delta i_t^{mp} \times \sigma_j$	-22.563***	-17.038***
	[2.450]	[2.858]
$\Delta i_t^{mp} \times \text{HHI}_{jt-1}$	0.307	-0.411*
	[0.230]	[0.240]
$\Delta i_t^{mp} \times Ch_{jt-1}$	-2.571**	7.342***
	[1.029]	[1.007]
$\Delta i_t^{mp} \times S_j$	-0.599***	0.887***
	[0.115]	[0.118]
Bank f.e.	Y	Y
Quarter f.e.	Y	Y
Observations	280,320	280,320
R-squared	0.1746	0.1375

Notes: This table shows the estimates from equation (2.38) for the log change of loans and deposits using quarterly data and a winsorized measure of withdrawal risk σ_j. Significance at 1%, 5%, and 10% is denoted by *, **, and *** respectively.

Table 2.6 shows the percentiles of the variables relevant for the heterogeneous behavior of loans. Using Tables 2.5 and 2.6, we can conclude that the heterogeneity in the withdrawal risk is the most relevant variable. For loans, if we compare banks at

[18]Some preliminary regressions, not shown in this document, show that withdrawal risk can explain the dynamic behavior of lending

the 90th and 10th percentile, a monetary shock that increases the Fed funds rate by 100 basis points reduces lending by an extra 1% for banks with a higher withdrawal risk, by an extra 0.2% for banks with a higher cash-deposit ratio, and by an extra 0.5% for big banks.

Table 2.6: Percentiles for control variables

	Percentile			
	10%	25%	75%	90%
Withdrawal risk	0.026	0.033	0.056	0.074
Cash-Deposit ratio	0.026	0.035	0.067	0.098
HHI	0.12	0.16	0.30	0.42
Bank size	0.10	0.25	0.75	0.90

Notes: This table shows the values of the control variables by percentile.

Deposits decline by an extra 0.8% for banks with higher withdrawal risk, by an extra 0.5% for banks with a lower cash-deposit ratio, by an additional 0.7% for small banks, and by an additional 0.1% for banks that raise deposits in more concentrated markets.

Results from above show that withdrawal risk is relevant for the behavior of loans and deposits. Moreover, this variable is quantitatively important to explain the different responses of lending and deposits across banks.

2.4.5 Aggregate effect of withdrawal risk

As a first approximation to study the aggregate effect of withdrawal risk, I calculate the sensitivity of aggregate lending to withdrawal risk as the weighted average of individual bank coefficients. First, I estimate by OLS for each bank the effect of monetary policy on lending over one year:

$$\Delta \log(L_{jt}) = \lambda_j + \beta_{j0}\Delta i_t^{mp} + \beta_{j1}\Delta i_{t-1}^{mp} + \beta_{j2}\Delta i_{t-2}^{mp} + \beta_{j3}\Delta i_{t-3}^{mp} \tag{2.40}$$

where i_t^{mp} is the measure of monetary policy shocks from Nakamura and Steinsson (2018) and L_{jt} is lending from bank j in quarter t. Using the coefficients from regression (2.40), I define loan beta $\beta_j = \hat{\beta}_{j0} + \hat{\beta}_{j1} + \hat{\beta}_{j2} + \hat{\beta}_{j3}$ as the individual effect of monetary policy on lending over one year. Then, I estimate by OLS:

$$\beta_j = \text{constant} + \theta\sigma_j \tag{2.41}$$

where σ_j is a measure of withdrawal risk for bank j. Table 2.7 shows the estimates from regression (2.41) and, consistent with previous estimation results, a higher withdrawal risk reduces lending proportionally more after a monetary tightening over one year. In particular, banks at the 90th percentile reduce lending by an extra 2.4%, over one year, relative to banks at the 10th percentile after a monetary shock that increases the Fed funds rate by 1%.

Table 2.7: Response of loans to monetary policy and withdrawal risk

	Loan betas
Withdrawal risk	-50.053***
	[6.903]
Observations	8,047
R-squared	0.0065

Notes: Cross-sectional OLS regression of loan betas β_j on withdrawal risk σ_j.

Finally, using the predicted value for β_j from regression (2.41), I calculate a measure of the aggregate effect of withdrawal risk as the weighted average of individual bank responses to monetary policy shocks explained by withdrawal risk.

$$\hat{\beta} = \sum_j \left[\frac{L_j}{\sum_j L_j} \right] \hat{\beta}_j \tag{2.42}$$

The estimated coefficient is equal to -0.89, i.e. $\hat{\beta} = -0.89$, which implies that aggregate lending declines by 0.9% after a monetary policy shock that increases the Fed funds rate in 1% over a one-year period due to the *withdrawal risk channel*.

2.4.6 A robust measure of withdrawal risk

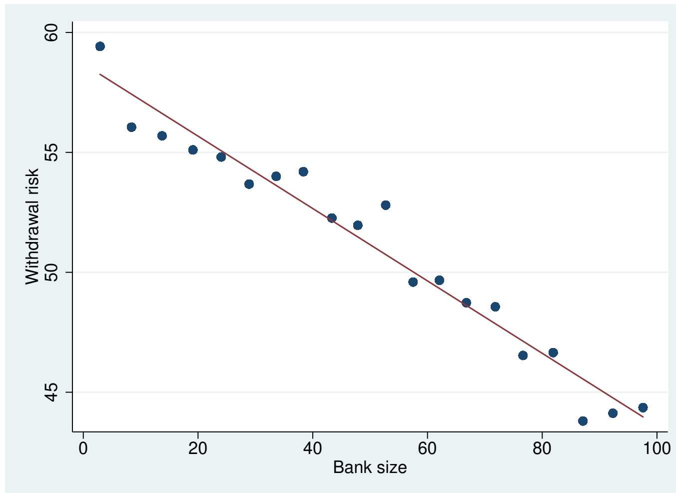
By definition from the model, a withdrawal shock is a shock to deposits that is iid across time and banks. This implies that equity and loans are independent of these shocks (contemporaneously). Hence, any variation in equity or loans that increase or decrease deposits should not be included in the construction of a measure of withdrawal risk. In the previous sections, I have used a measure of withdrawal risk that controls for volatility of bank equity since for some banks, especially big banks, volatility of equity induces volatility of deposits.

In this section, I use a robust measure of withdrawal risk. The idea is that the number of times (relative to the number of observations) the quarterly growth rate of deposits of bank j is either very high or very low is a better indicator of exogenous volatility of deposits than the actual value of this growth rate. I define a robust measure of withdrawal risk for bank j:

$$\sigma_j = \left[\frac{\sum_t \mathbb{1}\{\Delta \log \text{dep}_{jt} \notin [\text{25th-75th percentile of } \Delta \log \text{dep}]\}}{\text{number of observations for bank } j} \right] \times 100 \qquad (2.43)$$

Similar to the measure of withdrawal risk used in the previous sections, I find that big banks have lower withdrawal risk (see Figure 2·8).

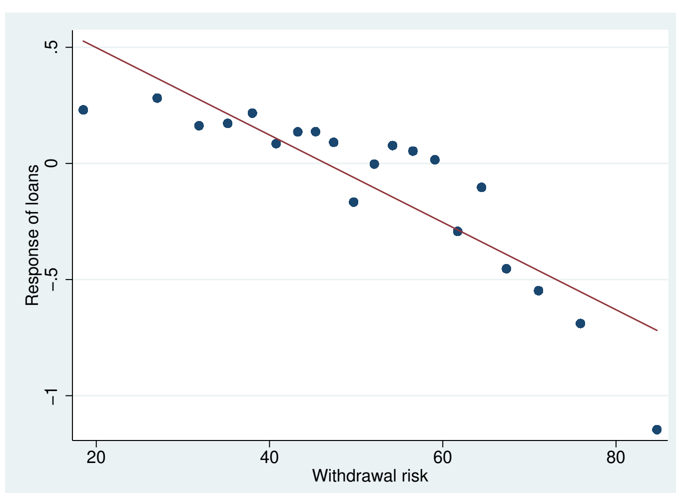
Figure 2·8: Robust withdrawal risk and bank size



Notes: This figure plots the relationship between the new withdrawal risk measure and bank size.

Figure 2·9 shows the relationship between the response of loans after a monetary policy shock and this new measure of withdrawal risk and the main conclusion from the previous sections is the same: Banks with higher withdrawal risk are more sensitive to monetary policy shocks.

Figure 2·9: Response of loans and robust withdrawal risk



Notes: This figure plots the relationship between the new withdrawal risk measure and the response of loans after a monetary policy shock.

Table 2.8 shows the percentiles for the new withdrawal risk measure and the implied β_j coefficient associated with that measure. If we compare banks at the 90th and 10th percentile, we find that banks at the 90th percentile reduce lending by an extra 0.82% after a monetary policy shock that increases the Fed funds rate in 1%. Banks at the 75th percentile reduce lending by an extra 0.45% relative to banks at the 25th percentile. These estimates are similar to those found in the previous sections.

Table 2.8: Percentiles for robust withdrawal risk and loan betas

Percentile	Robust withdrawal risk	Loan betas
10%	29.4	0.32
25%	39.3	0.14
75%	63.0	-0.31
90%	73.2	-0.50

Notes: This table shows estimates of the robust withdrawal risk measure by percentile and the loan betas β_j associated with that measure.

2.5 Conclusion

The heterogeneous exposure to withdrawal shocks in the banking system leads to heterogeneous responses after a change in the monetary policy rate. General equilibrium effects exacerbate the heterogeneity in the responses of loans. In particular, a monetary tightening reduces aggregate deposits, which further reduces funding sources for small banks endowed with a riskier distribution of withdrawal shocks.

Bank heterogeneity is relevant to explain the aggregate behavior of loans after a monetary policy shock. In particular, a model with a representative bank will overestimate the aggregate effect of monetary policy on lending. As bank heterogeneity increases, the effect of monetary policy on lending declines.

In the U.S. economy, banks at the 90th percentile of the withdrawal risk distribution reduce lending by an extra 1% and deposits by an extra 0.7-0.9% relative to banks at the 10th percentile after a monetary policy shock that increases the Fed funds rate in 100 basis points. The importance of withdrawal risk for the heterogeneous responses of lending and deposits is robust to the inclusion of control variables such as bank size, liquidity, and concentration. Moreover, this variable is quantitatively more important to explain the heterogeneous responses of lending and deposits across banks.

The aggregate effect of withdrawal risk on lending is also important. Aggregate lending declines by 0.9% after a monetary policy shock that increases the Fed funds rate in 1% over a one-year period due to the withdrawal risk channel.

Chapter 3

Banks and the Marginal Propensity to Lend in General Equilibrium

3.1 Introduction

Financial frictions expose bank lending to idiosyncratic deposit shocks. This is important because it provides a role for deposits in the transmission of monetary policy to macroeconomic aggregates. The marginal propensity to lend measures the exposure of bank lending to deposit shocks. In this chapter, I study the role of the marginal propensity to lend in the transmission of monetary policy in a general equilibrium model with a representative bank.

In the model, banks provide loans using deposits and wholesale funding and face frictions to substitute deposits with wholesale funding. Then, banks cannot fully compensate for a decline in deposits with an increase in wholesale funding, which exposes their bank lending to deposit changes. Banks collect deposits from patient households and demand wholesale funding from impatient households. Patient households cannot provide wholesale funding to banks and have a discount factor higher than impatient households. Intermediate goods producers borrow funds from banks to invest in capital and produce output using capital and labor from patient households.

In partial equilibrium, a higher degree of financial frictions increases MPLs, reduces the response of deposits to monetary shocks, and increases the pass-through from the policy rate to the lending rate. In general equilibrium, at steady state, the interest rate on deposits and the policy rate are given by the discount factor of patient and impatient households. Then, the marginal cost of wholesale funding does not change with the degree of financial frictions, which dampens the role of frictions in the transmission of monetary policy. This occurs because the loan rate pass-through depends on the marginal cost of wholesale funding. If the marginal cost does not change, the loan pass-through is (almost) unchanged. If we keep the demand for wholesale funding constant, then higher fricitons increase the MPLs and loan pass-through, which amplifies the transmission of monetary policy to macroeconomic aggregates.

In general equilibrium, higher frictions increase the aggregate MPL and also reduce the response of deposits to monetary shocks. Higher MPLs amplify the transmission of monetary policy. However, the lower exposure of deposits to monetary shocks dampens the transmission of monetary policy. This occurs because higher frictions increase the exposure of banks to deposit changes. Banks try to reduce the exposure of their deposits to aggregate shocks by reducing their demand for wholesale funding, which dampens the response of deposits and lending to monetary shocks. Then, economies with high MPLs and a similar response of deposits to monetary shocks experience a larger decline in macroeconomic aggregates after a contractionary monetary shock. This chapter evaluates the impact of higher MPLs alone by increasing higher frictions and keeping the demand for wholesale funding unchanged at steady state. The results are compared with an economy with higher frictions and a baseline economy, calibrated to the U.S. economy.

Related literature. This chapter contributes to the literature studying financial intermediaries in macroeconomic models (Gertler and Karadi, 2011; Gertler and Kiyotaki, 2010; Bianchi and Bigio, 2022; Begenau, 2020; Jamilov and Monacelli, 2023; Jamilov, 2021; Bellifemine et al., 2022; Brunnermeier and Sannikov, 2016; Balloch and Koby, 2019; Polo, 2021; Whited et al., 2022; Ulate, 2021; Wang, 2018). The contribution of this chapter is to provide a DSGE model with banks to study the role of the marginal propensity to lend and quantify the impact of higher aggregate MPLs in the context of the U.S. economy.

This chapter also contributes to the literature that studies the bank lending channel (Drechsler et al., 2017; Kashyap and Stein, 1995, 2000; Kishan and Opiela, 2000; Williams, 2020; Stein, 1998). The contribution of this chapter is to study the general equilibrium implications of higher MPLs.

Outline. The remainder of this chapter proceeds as follows. In Section 3.2, I present a general equilibrium model with a representative bank. The model is calibrated in Section 3.3 and the main results of this chapter are discussed. Section 3.4 concludes.

3.2 The Model

This section develops a general equilibrium model with a representative bank that faces frictions to substitute deposits with wholesale funding to study the role of the marginal propensity to consume in the transmission of monetary policy. The core framework is a standard New Keynesian DSGE model developed by Christiano et al. (2005) and Smets and Wouters (2007). To this, I add banks that collect deposits from patient households, demand wholesale funding from impatient households subject to a quadratic cost, and provide loans to firms that produce intermediate goods. There are six types of agents: impatient households, patient households, banks, intermediate goods producers, capital producers, and retailers. In addition, there is a Central Bank that conducts monetary policy. Financial frictions expose bank lending to idiosyncratic deposit shocks, which amplify the impact of monetary policy on macroeconomic aggregates.

3.2.1 Banks

There is a continuum of banks, indexed by j. Banks invest in liquid assets B subject to a liquidity constraint $B_j \geq \bar{B}$, supply loans L_j, collect deposits D_j and use wholesale funding F_j subject to a quadratic cost $\frac{\phi}{2P}F_j^2$, where P is the price level so that the spread on wholesale funding is linearly increasing in the real amount borrowed. They earn interest rates i, i^l on liquid assets and loans, respectively, and pay an interest rate i^d on deposits. The bank balance sheet of bank j is given by

$$B_j + L_j = D_j + F_j \tag{3.1}$$
$$B_j \geq \bar{B}$$

Similar to Ulate (2021), depositors and borrowers have CES preferences across banks. Each bank faces an upward-sloping deposit demand and a downward-sloping loan demand.

$$\log L_j = -\varepsilon^l \Big(\log\big(1 + i_j^l\big) - \log\big(1 + i^l\big) \Big) + \log L \tag{3.2}$$
$$\log D_j = \varepsilon^d \Big(\log\big(1 + i_j^d\big) - \log\big(1 + i^d\big) \Big) + \log D \tag{3.3}$$

Banks choose the amount of liquid assets, loans, deposits, and wholesale funding to maximize their net income.

$$\max_{B_j, L_j, D_j, F_j} iB_j + i_j^l L_j - i_j^d D_j - iF_j - \frac{\phi_j^f}{2P} F_j^2 \tag{3.4}$$
$$\text{s.t.} \quad (3.1) - (3.3)$$

Banks set lending and deposit rates as a mark-up and a mark-down, respectively, on the policy rate and its marginal cost of wholesale funding.

$$1 + i_j^l = \left(\frac{\varepsilon^l}{\varepsilon^l - 1}\right)\left(1 + i + \phi_j^f \frac{F_j}{P}\right) \tag{3.5}$$

$$1 + i_j^d = \left(\frac{\varepsilon^d}{1 + \varepsilon^d}\right)\left(1 + i + \phi_j^f \frac{F_j}{P}\right) \tag{3.6}$$

Notice that if banks can substitute deposits with wholesale funding without cost, $\phi_j^f = 0$, lending decisions are independent of deposit decisions, which implies that idiosyncratic deposit shocks do not affect bank lending. Then, in the absence of frictions, there is no role for deposits in the monetary transmission to bank lending.

3.2.2 Patient Households

There is a continuum of patient households of measure one. Each household supplies labor n_t, saves in deposits D_t, and consumes c_t subject to habit formation with parameter h, similar to Christiano et al. (2005) and Smets and Wouters (2007). This type of household cannot save in government bonds nor by lending funds to banks in the form of wholesale funding. Preferences are given by

$$\mathrm{E}_0 \sum_{t=0}^{\infty} \beta^t \left(\frac{(c_t - hc_{t-1})^{1-\sigma} - 1}{1 - \sigma} - \chi \frac{n_t^{1+\frac{1}{\eta}}}{1 + \frac{1}{\eta}} \right) \tag{3.7}$$

with $0 < \beta < 1$, $0 < h < 1$, and $\sigma, \chi, \eta > 0$. They earn an interest rate i^d on deposits, W_t is the nominal wage, Π_t are net payouts to the household from ownership of financial and non-financial firms, and T_t are nominal lump-sum taxes. Then, the budget constraint of this type of household is given by

$$P_t c_t + D_t = W_t n_t + \Pi_t - T_t + (1 + i_{t-1}^d) D_{t-1} \tag{3.8}$$

The marginal utility of consumption is ϱ_t, and the optimality conditions for consumption and labor are given by

$$\varrho_t = (c_t - hc_{t-1})^{-\sigma} - \beta h \mathrm{E}_t (c_t - hc_{t-1})^{-\sigma} \tag{3.9}$$

$$1 = \mathrm{E}_t \Big(\beta \frac{\varrho_{t+1}}{\varrho_t} (1 + i_t^d) \frac{P_t}{P_{t+1}} \Big) \tag{3.10}$$

$$\chi n_t^{\frac{1}{\eta}} = \varrho_t \frac{W_t}{P_t} \tag{3.11}$$

3.2.3 Impatient Households

There is a continuum of impatient households of measure one. They have a lower discount factor, i.e. $\beta_u < \beta$, and save by lending funds to banks in the form of wholesale funding $F_{u,t}$ and by lending funds to the government in the form of bonds $B_{u,t}$. In both cases, they earn an interest rate i_t, which is the nominal rate on bonds issued by the government and it is set by the Central Bank. This type of household supplies labor $n_{u,t}$, and consumes $c_{u,t}$. Preferences are given by

$$\mathrm{E}_0 \sum_{t=0}^{\infty} \beta_u^t \left(\frac{(c_{u,t})^{1-\sigma_u} - 1}{1 - \sigma_u} - \chi \frac{n_{u,t}^{1+\frac{1}{\eta}}}{1 + \frac{1}{\eta}} \right) \tag{3.12}$$

with $0 < \beta_u < 1$, and $\sigma_u, \chi, \eta > 0$. They earn an interest rate i on bonds and wholesale funding, and W_t is the nominal wage. Then, the budget constraint of this type of household is given by

$$P_t c_{u,t} + F_{u,t} + B_{u,t} = W_t n_{u,t} + (1 + i_{t-1})(B_{u,t-1} + F_{u,t-1}) \tag{3.13}$$

The optimality conditions for consumption and labor are the following:

$$1 = \mathrm{E}_t \Big(\beta_u \Big(\frac{c_{u,t}}{c_{u,t+1}} \Big)^{\sigma_u} (1 + i_t) \frac{P_t}{P_{t+1}} \Big) \tag{3.14}$$

$$\chi n_{u,t}^{\frac{1}{\eta}} = \frac{W_t}{P_t} \Big(\frac{1}{c_{u,t}} \Big)^{\sigma_u} \tag{3.15}$$

3.2.4 Intermediate goods firms

On the production side of the economy, competitive nonfinancial firms produce intermediate goods, which are sold to retail firms. At the end of period $t-1$, an intermediate goods producer borrows funds from banks to invest in capital k_t and use labor $n_t + n_{u,t}$ and capital to produce output y_t^m. The production function is given by

$$y_t^m = a_t k_t^{\alpha} (n_t + n_{u,t})^{1-\alpha} \tag{3.16}$$

where a_t denotes total factor productivity. Let P_t^m be the price of intermediate goods output. Then, the firm chooses labor to maximize nominal profits

$$\Pi_t^m = P_t^m y_t^m - W_t(n_t + n_{u,t}) - Z_t k_t \tag{3.17}$$

where firms pay a dividend of Z_t to each unit of capital borrowed from banks. The optimality condition for labor is

$$(1-\alpha)\frac{P_t^m}{P_t}\frac{y_t^m}{n_t + n_{u,t}} = \frac{W_t}{P_t} \tag{3.18}$$

The dividend Z_t is such that nominal profits are equal to zero.

$$Z_t = P_t^m \alpha \frac{y_t^m}{k_t} \tag{3.19}$$

Then, similar to Gertler and Karadi (2011) and Ulate (2021), the nominal return on loans for banks is

$$1 + i_t^l = \frac{P_t}{P_{t-1}} \frac{\frac{Q_t}{P_t}(1-\delta) + \frac{P_t^m}{P_t}\alpha\frac{y_t^m}{k_t}}{\frac{Q_{t-1}}{P_{t-1}}} \tag{3.20}$$

3.2.5 Capital producing firms

Capital producers are competitive and buy capital from intermediate goods producers and build new capital. There are flow adjustment costs $f(.)$ to produce new capital. The value of a unit of new capital is Q_t. Let I_t be investment, then the evolution of capital is

$$k_{t+1} = (1-\delta)k_{t-1} + I_t \tag{3.21}$$

Discounted real profits from capital producers are given by

$$\mathrm{E}_t \sum_{\tau=t}^{\infty} \beta^{\tau-t} \Lambda_{t,\tau} \left[\left(\frac{Q_\tau}{P_\tau} - 1 \right) I_\tau - f\left(\frac{I_\tau}{I_{\tau-1}} \right) I_\tau \right] \tag{3.22}$$

where $\Lambda_{t,\tau}$ is the patient household's stochastic discount factor between periods t and τ. Following Christiano et al. (2005) and Gertler and Karadi (2011), the adjustment cost function satisfies $f(1) = f'(1) = 0$ and $f''(1) > 0$. Then, the optimal condition for investment is given by

$$\frac{Q_t}{P_t} = 1 + f\Big(\frac{I_t}{I_{t-1}}\Big) + f'\Big(\frac{I_t}{I_{t-1}}\Big)\Big(\frac{I_t}{I_{t-1}}\Big) - \mathrm{E}_t \beta \frac{\varrho_{t+1}}{\varrho_t} f'\Big(\frac{I_{t+1}}{I_t}\Big)\Big(\frac{I_{t+1}}{I_t}\Big)^2 \tag{3.23}$$

3.2.6 Retail firms

Following Gertler and Karadi (2011) and Ulate (2021), there is a continuum of retail firms of measure one. They use one unit of intermediate goods and transform them into one unit of a differentiated variety of retail good without an additional cost. Final output Y_t is a CES aggregate of retail goods.

$$y_t = \left(\int_0^1 y_t(s)^{\frac{\theta-1}{\theta}} ds \right)^{\frac{\theta}{\theta-1}} \tag{3.24}$$

where $y_t(s)$is output by retailer s. The demand for a variety and the price index are given by

$$y_t(s) = \left(\frac{P_t(s)}{P_t}\right)^{-\theta} y_t \tag{3.25}$$

$$P_t = \left(\int_0^1 P_t(s)^{1-\theta} ds\right)^{\frac{1}{1-\theta}} \tag{3.26}$$

Each firm can adjust its price with probability $1-\gamma$. Then, the pricing problem of each firm s is to choose the optimal reset price $P_t^*(s)$ to solve:

$$\max \mathrm{E}_t \sum_{\tau=0}^{\infty} \gamma^\tau \beta^\tau \Lambda_{t,t+\tau} \frac{P_t}{P_{t+\tau}} \left[P_t^*(s) - P_{t+\tau}^m\right] y_{t+\tau}(s) \tag{3.27}$$

The optimal reset price is given by the following system of equations

$$\theta \Gamma_t^1 = (\theta - 1)\Gamma_t^2 \tag{3.28}$$

$$\Gamma_t^1 = \varrho_t \frac{P_t^m}{P_t} y_t + \gamma\beta \mathrm{E}_t \left(\frac{P_t}{P_{t+1}}\right)^{-\theta} \Gamma_{t+1}^1 \tag{3.29}$$

$$\Gamma_t^2 = \varrho_t \frac{P_t^*}{P_t} y_t + \gamma\beta \mathrm{E}_t \frac{P_t^*}{P_{t+1}^*} \left(\frac{P_t}{P_{t+1}}\right)^{-\theta} \Gamma_{t+1}^2 \tag{3.30}$$

The evolution of the aggregate price in the economy is given by

$$1 = (1-\gamma)\left(\frac{P_t^*}{P_t}\right)^{1-\theta} + \gamma\left(\frac{P_{t-1}}{P_t}\right)^{1-\theta} \tag{3.31}$$

The relation between the final output and intermediate output is

$$y_t^m = y_t v_t^p \tag{3.32}$$

The dispersion of prices is given by

$$v_t^p = \gamma\left(\frac{P_{t-1}}{P_t}\right)^{-\theta} v_{t-1}^p + (1-\gamma)\left(\frac{P_t^*}{P_t}\right)^{-\theta} \tag{3.33}$$

3.2.7 Monetary Policy

The Central Bank sets the interest rate on government bonds i_t and follows a Taylor Rule with interest rate smoothing.

$$i_t = (1 - \rho_i)(\bar{i} + \psi_\pi \pi_t) + \rho_i i_{t-1} + \varepsilon_t^{mp} \tag{3.34}$$

where $\bar{i}$ is the steady state nominal interest rate, $\rho_i \in [0, 1]$, and ε_t^{mp} is a monetary policy shock.

3.2.8 Resource Constraint

Output is divided between consumption, investment, adjustment costs to capital, and the quadratic cost to use wholesale funding. Then, the aggregate economy resource constraint is given by

$$y_t = c_t + c_{u,t} + I_t + f\left(\frac{I_t}{I_{t-1}}\right)I_t + \frac{\phi}{2}\left(\frac{F_{t-1}}{P_{t-1}}\right)^2 \frac{1}{1 + \pi_t} \tag{3.35}$$

In equilibrium, total loans are equal to the total value of capital in the economy.

$$L_t = Q_t k_{t+1} \tag{3.36}$$

Also, the wholesale market clears, and $B_{u,t} = 0$.

$$F_t = F_{u,t} \tag{3.37}$$

3.3 Model Results

In this section, I calibrate the model and study the role of MPLs in the transmission of monetary policy to macroeconomic aggregates.

3.3.1 Calibration

I use data from US banks to calibrate the banking side of the model. The degree of financial frictions ϕ^f is calibrated using the estimate of the average MPL elasticity from Chapter 1, which implies a pass-through from changes in wholesale funding to the lending rate of 0.02. Another important parameter is the inverse of the IES for impatient households, which is set to target the aggregate response of wholesale funding to changes in the policy rate, consistent with Drechsler et al. (2017).

Table 3.1: Calibrated parameters

Parameter	Value	Description	Target or source
ϕ^f	0.0045/4	Sensitivity of MC to w	Response of Δi^l to $\Delta F = 0.02$
ε^l	42*4	Loan demand elasticity	Loan spread=4.4%
ε^d	32*4	Deposit demand elasticity	Deposit spread=1.1%
$\overline{B}$	6.32	Liquidity requirement	Ratio Liquidity/Deposits=0.45
β	0.9925	Discount factor - PH	Deposit spread = 1.1%
β	0.99	Discount factor - IH	Policy rate of 4% annual
σ_u	1/20	Inverse of the I.E.S. - IH	Response of F to i = 5% annual
σ	1	Inverse of the I.E.S. - PH	Ulate (2021)
hh	0.815	Habit parameter	Gertler and Karadi (2011)
χ	3.409	Relative utility weight of labor	Gertler and Karadi (2011)
η	1	Frisch elasticity of labor supply	Ulate (2021)
α	0.333	Capital share	Ulate (2021)
δ	0.025	Depreciation rate	Ulate (2021)
η_I	1.728	Elasticity of q to investment	Ulate (2021)
γ	0.75	Prob. of not changing prices	Ulate (2021)
θ	6	Elasticity of substitution	Ulate (2021)
ψ_π	3.5	Taylor rule - Inflation coefficient	Ulate (2021)
ρ_i	0.8	Taylor rule - Smoothing parameter	Ulate (2021)

3.3.2 The role of MPLs

In this section, I compare three economies to study the role of MPLs in the transmission of monetary policy to real variables. The first economy is the baseline economy, with parameters calibrated according to Table 3.1. The second economy faces a high degree of financial frictions, i.e. high ϕ^f, but the rest of the parameters are identical to the baseline economy. The third economy faces a high degree of financial frictions, identical to the second economy but the amount borrowed from wholesale markets is identical to the baseline economy at steady state. In this case, I include a linear component in the cost of wholesale funding, κ, so that the amount borrowed at steady state is identical to the baseline economy. The rest of the parameters are identical to the baseline economy.

A key difference between the second and the third economy is the sensitivity of the marginal cost of wholesale funding to changes in the policy rate. At steady state, the marginal cost of wholesale funding is the same in the three economies but the sensitivity of the marginal cost to the policy rate is higher only in the third economy, which implies that the pass-through from the policy rate to the lending rate is (almost) identical in the baseline economy and the second economy, and significantly higher in the third economy. Lending and deposit rates are given by

$$1 + i^l = \left(\frac{\varepsilon^l}{\varepsilon^l - 1}\right)\left(1 + i + \phi^f \frac{F}{P} + \kappa\right) \tag{3.38}$$

$$1 + i^d = \left(\frac{\varepsilon^d}{1 + \varepsilon^d}\right)\left(1 + i + \phi^f \frac{F}{P} + \kappa\right) \tag{3.39}$$

In the baseline economy, we have $\phi^f = 0.0011, \kappa = 0$, in the second economy, we have $\phi^f = 0.0112, \kappa = 0$, and in the third economy we have $\phi^f = 0.0112, \kappa = -0.0489$, where κ is chosen such that wholesale funding in this economy is identical to the amount borrowed in the baseline economy at the steady state.

The baseline economy has an MPL of 0.56 and an MPL elasticity of 0.68, the second economy has an MPL of 0.92 and an MPL elasticity of 1.76, and the third economy has an MPL of 0.92 and an MPL elasticity of 1.14. Figure 3·1 shows the response of production when the policy rate increases by 1% on impact for the three economies. Financial frictions amplify the transmission of monetary policy at longer horizons. However, on impact, the second economy experiences a lower decline in production. This occurs because more frictions reduce the amount borrowed from wholesale funds such that the marginal cost of funding, $\phi_f f$, is constant. This implies that the total cost of wholesale funding falls at steady state. Then, an increase in the policy rate lowers inflation, which increases the real cost of wholesale funding. However, this increase is lower in the second economy, which relatively increases resources for consumption and dampens the decline of production.

Figure 3·1: Elasticity of Production

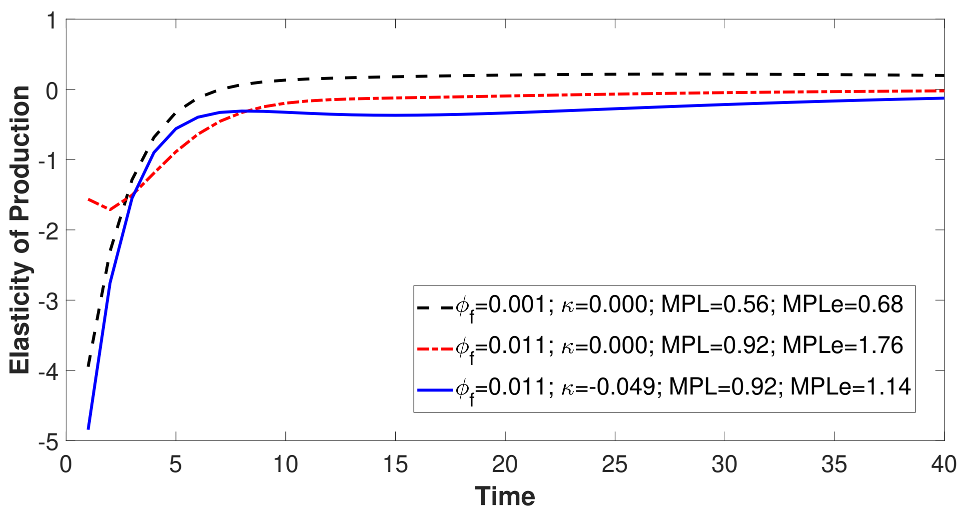


Figure 3·1 shows that, on impact, the third economy experiences a 23% larger decline in production, relative to the baseline economy while the second economy experiences a 60% lower decline in output. Then, economies with high MPLs amplify the response of production to changes in the policy rate if their loan pass-through also increases. Higher frictions do not amplify the response of production to monetary shocks in the short run.

Figure 3·2 shows the response of production after a monetary shock that increases the policy rate by 25 basis points. In the short run, higher MPLs do not amplify the transmission of monetary policy to production. They only amplify the response of production at longer horizons. Frictions that do not increase the loan pass-through dampen monetary policy in the short run and experience a larger decline in output at longer horizons.

Figure 3·2: Production

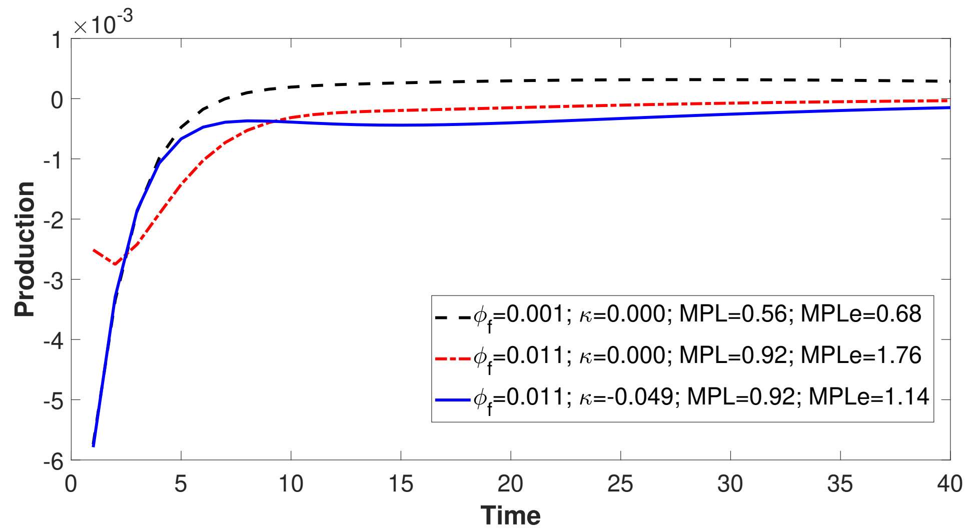


Figure 3·3 shows that financial frictions amplify the monetary transmission to bank lending. On impact, the second economy experiences a decline in lending 11% larger than the baseline economy, while the third economy experiences a 31% larger decline in lending. At longer horizons, the decline in bank lending is much larger in the economy with higher loan pass-through.

Figure 3·3: Bank lending

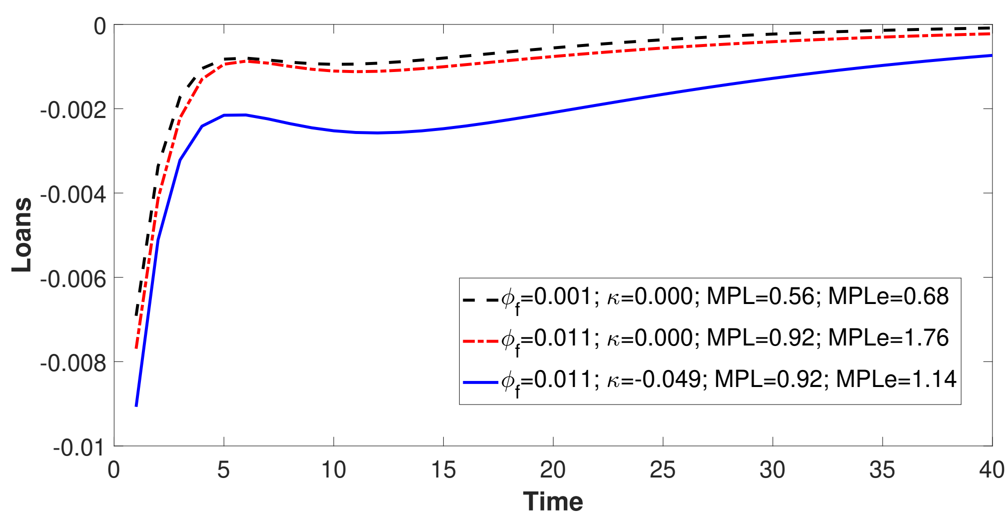
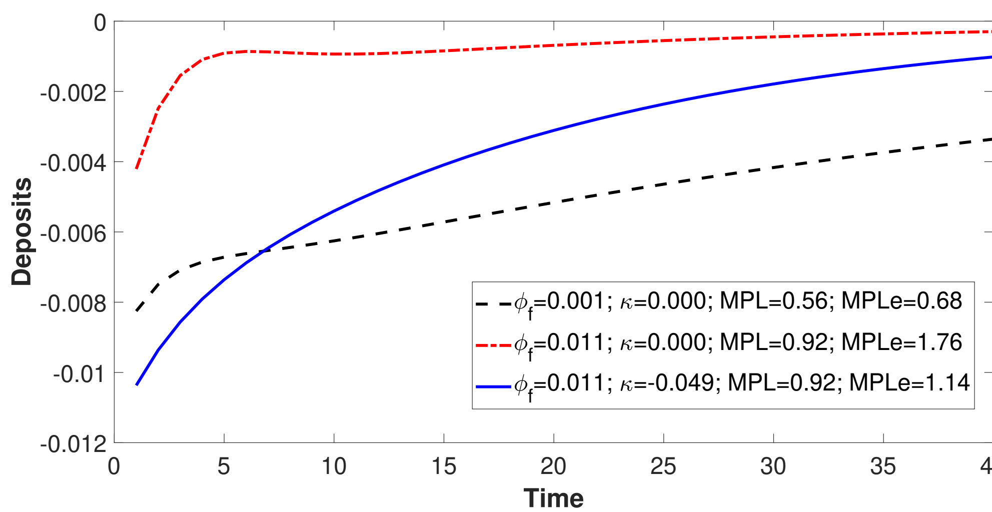


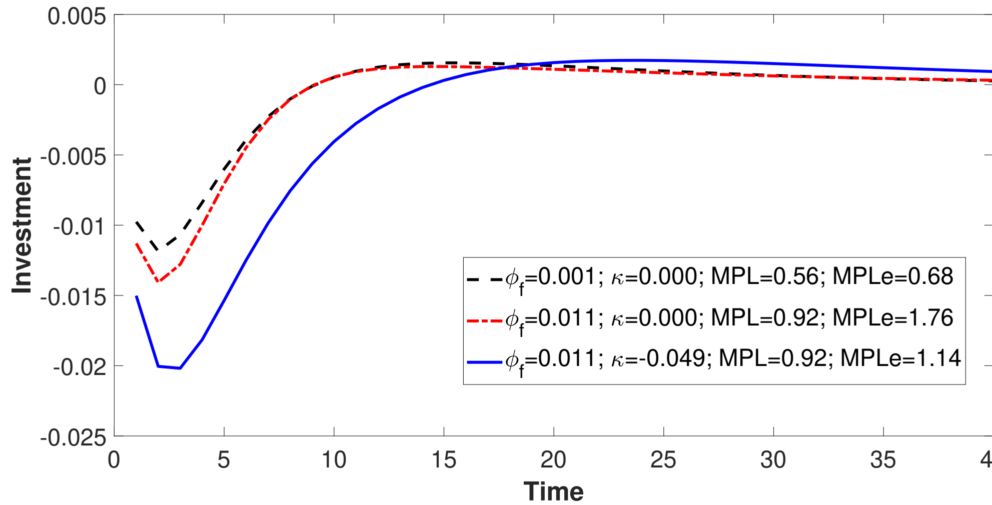
Figure 3·3 shows that the decline in bank lending is significantly larger in the third economy, relative to the second economy. This occurs because a higher degree of frictions increases the exposure of bank lending to deposit shocks. Then, banks try to reduce the exposure of their deposits to monetary shocks, which implies that deposits fall by less in economies with a high degree of financial frictions (see Figure 3·4). The lower decline in deposits dampens the transmission of monetary policy to bank lending.

Figure 3·4: Deposits



Related to the response of bank lending, Figure 3·5 shows that financial frictions also amplify the transmission of monetary policy to investment. On impact, the decline in investment is 16% larger in the second economy and 54% larger in the third economy. Then, economies with high MPLs amplify the transmission of monetary policy to bank lending, capital, and investment. If the sensitivity of the marginal cost is higher, the amplification is even larger.

Figure 3·5: Investment



Consistent with Drechsler et al. (2017), Figure 3·6 shows that wholesale funding increases with an increase in the policy rate. However, a higher degree of financial frictions dampens the response of wholesale funding to monetary shocks, especially in the third economy because of the higher sensitivity of their marginal cost, which reduces the demand for wholesale funding.

Figure 3·6: Wholesale funding

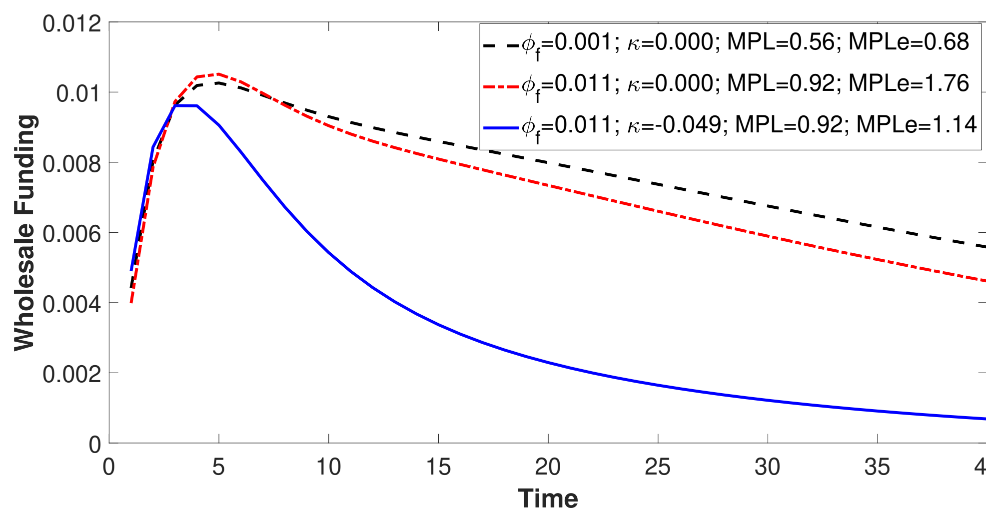


Figure 3·7 shows the response of the lending rate after a monetary shock divided by the initial response of the policy rate, which I define as the loan pass-through. Then, the pass-through is (almost) identical in the baseline and second economies, and it is significantly higher in the third economy. This occurs because higher frictions in the second economy reduce the demand for wholesale funding such that the marginal cost of wholesale funding and its sensitivity to shocks is unchanged, relative to the baseline economy, at steady state. The sensitivity of the marginal cost of wholesale funding to the policy rate is higher in the third economy, which increases the loan pass-through.

Figure 3·7: Loan pass-through

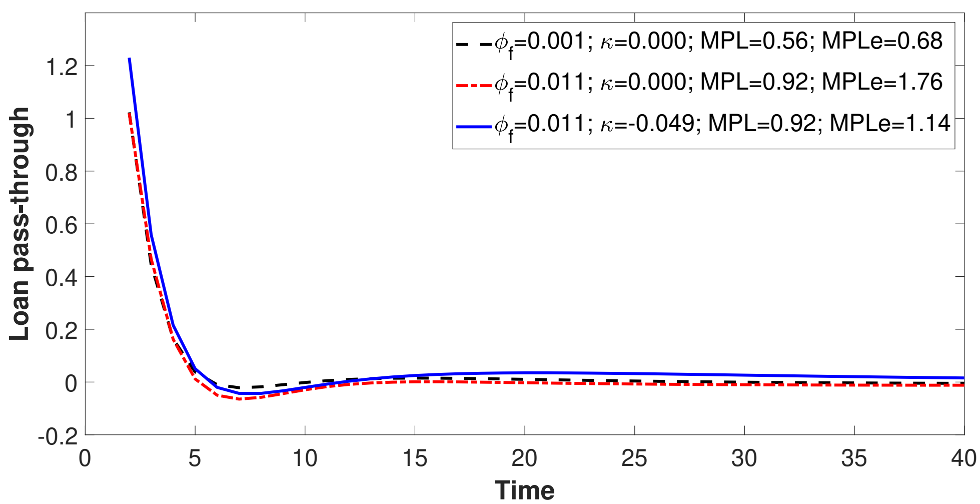
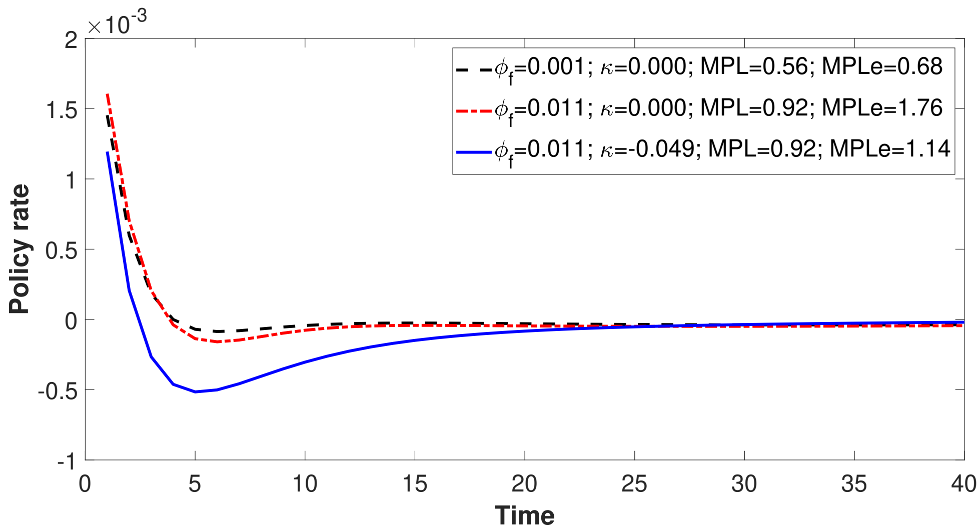


Figure 3·8 shows the response of the policy rate after a contractionary monetary shock. Given that higher MPLs amplify monetary shocks, the central bank increases its policy rate by less to avoid a larger decline in real variables. On impact, the central bank increases its policy rate by 18% *less* in the economy with more sensitive marginal costs and by 10% *more* in the economy with higher frictions. This occurs because frictions dampen monetary policy on impact. However, at longer horizons, the policy rate in both economies declines due to the amplification of monetary shocks.

Figure 3·8: Policy rate



3.4 Conclusion

This chapter studies the role of the marginal propensity to lend in the transmission of monetary policy to macroeconomic aggregates in a general equilibrium model. The main finding of the paper is that higher financial frictions that increase the aggregate MPL by 66% amplify the transmission of monetary policy to bank lending and investment, on impact, by 11% and 16%, respectively. Moreover, if the sensitivity of the marginal cost of wholesale funding to the policy rate also increases, the loan pass-through increases by 20%, which amplifies the response of bank lending and investment by 31% and 54%, respectively.

Financial frictions do not amplify the transmission of monetary policy to production in the short run. However, they do amplify the response of production to monetary shocks at longer horizons, due to the decline in investment. The amplification of monetary shocks is larger in economies with high sensitivity of their marginal costs. On impact, the response of total production after a 1% increase in the policy rate is amplified by 23% economies with 66% higher MPL and 20% higher loan pass-through. However, the central bank reduces its policy rate such that the response of production is identical to the baseline economy in the short run.

Appendix A

Supplementary Material for Chapter 1

A.1 Model

The maximization problem for a bank j is the following:

$$\max_{l_j, d_j, w_j} \quad i^l_j l_j + i b_j - i^d_j d_j - i w_j - \frac{\phi_j}{2} w_j^2$$

$$\text{s.t.} \quad \log l_j = -\varepsilon^l_j i^l_j + \gamma^l_j i + v^l_j$$

$$\log d_j = \varepsilon^d_j i^d_j - \gamma^d_j i + v^d_j$$

$$b_j + l_j = d_j + w_j$$

$$b_j \geq \bar{b}$$

$$w_j \geq 0$$

First order conditions

$$i^l_j - \frac{1}{\varepsilon^l_j} = \lambda_j$$

$$i^d_j + \frac{1}{\varepsilon^d_j} = \lambda_j$$

$$i + \mu^b_j = \lambda_j$$

$$i + \phi_j w_j = \lambda_j + \mu^w_j$$

Demand for wholesale funding

$$\phi_j w_j = \mu_j^b + \mu_j^w$$

Given $\phi_j > 0$: If $b_j > \bar{b}$, then $\mu_j^b(b - \bar{b}) = 0$ implies that $\mu_j^b = 0$. Then $\phi_j w_j = \mu_j^w$, and $\phi_j w_j^2 = \mu_j^w(w_j - 0) = 0$, which implies that we have $w_j = 0$

$$i_j^l - \frac{1}{\varepsilon_j^l} = i = i_j^d + \frac{1}{\varepsilon_j^d}$$

Then, if the liquidity constraint doesn't bind, changes in deposits are unrelated to lending changes. An increase in loan demand increases lending rates, which increases the demand for wholesale funding and the liquidity constraint binds, i.e. $b_j = \bar{b}$, and $w_j > 0$. Equilibrium quantities depend on the size of deposit and loan demands. In this case, we have $\lambda_j^{MPL} = MPL_j = 0$ if the liquidity constraint binds.

In the extreme case of $\phi_j = 0$, we have $i_j^l - \frac{1}{\varepsilon_j^l} = i = i_j^d + \frac{1}{\varepsilon_j^d}$ if the liquidity constraint (or borrowing constraint if we change the sign of b) doesn't bind. If the liquidity (borrowing) constraint binds, then we have $i_j^l - \frac{1}{\varepsilon_j^l} = i_j^d + \frac{1}{\varepsilon_j^d}$ and $\bar{b} + l_j = d_j$. In this case, we have $MPL_j = 1$, and MPL elasticity $\lambda_j^{MPL} = \frac{d_j}{l_j}$.

In the equilibrium with wholesale funding $w_j > 0$

$$i_j^l = i + \frac{1}{\varepsilon_j^l} + \phi_j w_j$$
$$i_j^d = i - \frac{1}{\varepsilon_j^d} + \phi_j w_j$$
$$b_j = \bar{b}$$

Then, using the balance sheet constraint:

$$i_j^l = i + \frac{1}{\varepsilon_j^l} + \phi_j(\bar{b} + l_j - d_j)$$
$$i_j^d = i - \frac{1}{\varepsilon_j^d} + \phi_j(\bar{b} + l_j - d_j)$$

The response of deposit rates, lending rates, deposits, and loans are given by the following system of equations:

$$\beta_j^{i^l} = 1 + \phi_j w_j \beta_j^w = 1 + \phi_j l_j \beta_j^l - \phi_j d_j \beta_j^d$$
$$\beta_j^{i^d} = 1 + \phi_j w_j \beta_j^w = 1 + \phi_j l_j \beta_j^l - \phi_j d_j \beta_j^d$$
$$\beta_j^l = -\varepsilon_j^l \beta_j^{i^l} + \gamma_j^l$$
$$\beta_j^d = \varepsilon_j^d \beta_j^{i^d} - \gamma_j^d$$

Then, we have the following

$$\beta_j^l - \gamma_j^l = -\varepsilon_j^l - \varepsilon_j^l \phi_j l_j \beta_j^l + \varepsilon_j^l \phi_j d_j \beta_j^d$$
$$\beta_j^d + \gamma_j^d = \varepsilon_j^d + \varepsilon_j^d \phi_j l_j \beta_j^l - \varepsilon_j^d \phi_j d_j \beta_j^d$$

Then,

$$\beta_j^l = \frac{\varepsilon_j^l \phi_j d_j}{1 + \varepsilon_j^l \phi_j l_j} \beta_j^d + \frac{1}{1 + \varepsilon_j^l \phi_j l_j} \left(\gamma_j^l - \varepsilon_j^l \right)$$
$$\beta_j^d = \frac{\varepsilon_j^d \phi_j l_j}{1 + \varepsilon_j^d \phi_j d_j} \beta_j^l - \frac{1}{1 + \varepsilon_j^d \phi_j d_j} \left(\gamma_j^d - \varepsilon_j^d \right)$$

We can express the response of loans as follows;

$$\beta_j^l = \lambda_j^{MPL} \beta_j^d + (1 - MPL_j)\Big(\gamma_j^l - \varepsilon_j^l\Big)$$

Appendix B

Supplementary Material for Chapter 2

B.1 Proof of results

Proof of Result 1:

(i) Equation (2.5) must hold for both types of banks and nominal interest rates are taken as given. Hence, it follows that

$$F_s(e_s) = F_b(e_b)$$

By Assumption 1, the result 1(i) immediately follows from above.

(ii) Similarly, equation (2.6) must hold for both types of banks:

$$(i^w - i^m) \int_{e_s}^{\overline{\omega}} \omega f_s(\omega) d\omega + \mu_s = (i^w - i^m) \int_{e_b}^{\overline{\omega}} \omega f_b(\omega) d\omega + \mu_b$$

By result 1(i) and Assumption 1

$$e_s > e_b \Rightarrow \int_{e_s}^{\overline{\omega}} \omega f_s(\omega) d\omega > \int_{e_b}^{\overline{\omega}} \omega f_b(\omega) d\omega \Rightarrow \mu_b > \mu_s$$

By nonnegativity of lagrange multiplier,

$$\mu_b > \mu_s \geq 0 \Rightarrow D_b = \phi N_b \quad \text{and} \quad D_s \leq \phi N_s$$

(iii) By market clearing conditions for deposits market.

$$D = \psi D_s + (1-\psi)D_b = \psi D_s + (1-\psi)\phi N_b$$

where D is the aggregate demand for deposits. Then, the optimal amount of deposits for small banks:

$$D_s = \frac{1}{\psi}\Big[D - (1-\psi)\phi N_b\Big]$$

(iv) Balance sheet for big banks,

$$L_b + M_b = D_b + N_b$$

Using the results from above,

$$L_b + e_b\phi N_b = (1+\phi)N_b$$

It follows that,

$$L_b = (1+\phi(1-e_b))N_b$$

In the case of small banks,

$$L_s + e_s D_s = N_s + D_s$$

Hence,

$$L_s = N_s + (1-e_s)D_s$$

Proof of Result 2:

(i) Use equation (2.5) and notice that under Assumption 2 the spread $i^w - i^m$ increases, which increases the cost of lending for all banks and the cost of deposits for small banks. Then, bank lending falls and the interest rate on loans increases, which increases the marginal benefit of deposits and the interest rate on deposits. A higher interest rate on reserves increases the reserve-deposit ratio, which reduces the cost of lending and the cost of deposits. Higher cost of lending reduces loans and increases reserves. Also higher interest rate on reserves increases reserve-deposit ratios.

$$\frac{\partial e_s}{\partial i^m} = \frac{f_b(e_b)}{f_s(e_b)} \frac{f_s(e_b)}{f_s(e_s)} \frac{\partial e_b}{\partial i^m} \tag{B.1}$$

In regime I, e_b is sufficiently small such that $\frac{f_b(e_b)}{f_s(e_b)} > 1$ and given that $f(.)$ is decreasing $\frac{f_s(e_b)}{f_s(e_s)} > 1$, which implies the result: $\frac{\partial e_s}{\partial i^m} > \frac{\partial e_b}{\partial i^m}$

(ii) Higher interest on reserves and higher discount window rate increases the lending rate. It also increases reserve-deposit ratios, which has a small impact on lending rates. Then, the interest rate on loans increases and total lending decreases. The interest rate on deposits also increases because higher interest rate on loans increases the marginal benefit of deposits.

(iii) Moreover, if the increase in the spread between the discount window rate and the interest on reserves is sufficiently high, interest rate on loans and deposits increases by more than 1% after a 1% increase in the interest on reserves. The spread between lending and deposit rates also increases because the cost of losing deposits also increases with higher discount window rates.

(iv) Result follows directly from Assumption 3.

(v) By result 1(iv), loans by big banks depend only on reserve-deposit ratio and loans by small banks depend additionally on total deposits. Then, the decline in lending is higher for small banks unless the increase in deposits is sufficiently high.

(vi) A higher withdrawal uncertainty for small banks induces a higher reserve-deposit ratio by equation (2.6) which reduces lending by small banks, increases i^l, and increases lending by big banks. Also reserve-deposit ratios for big banks will be lower. Then, a higher interest rate on reserves induces a larger decline in lending for small banks. A higher withdrawal uncertainty for big banks reduces the gap in the response between big and small banks because they need a higher reserve-deposit ratio.

Proof of Result 3:

(i) An increase in i increases i^l and $i^l - i^d$ by result 2. Then, by equation (2.18) lending decreases, and by equation (2.17) deposits by workers increase. In equilibrium, $L = D$, then inflation increases such that the real rate of deposits decreases and the real rate of loans increases. Therefore, deposits and loans decrease.

(ii) Follows from the decline in aggregate deposits and the increase in the reserve-deposit ratio.

(iii) As uncertainty about the withdrawal of deposits increases, the reserve-deposit ratio increases, which contracts lending and deposits and increases i^l in GE. Also, i^l is more sensitive to shocks because the spread between discount rate and interest on reserves rises. Hence, the decrease in output y_2 is higher which implies that the decline in labor demand, deposits, and lending is larger.

(v) Decline in lending is larger for small banks, then a higher concentration follows.

Proof of Result 4:

(i) Under a contractionary monetary policy shock, a larger share of large banks, induces a smaller decline in lending since large banks are less responsive to shocks (result 1) and a smaller increase in i^l. Moreover, this leads to a smaller increase in spread $i^l - i^m$, increasing the reserve-deposit ratio for small and big banks. This implies that big banks experience a larger decline in lending, and the response of small banks is even larger. Nevertheless, the aggregate response of lending is dampened.

(ii) From above, small banks have a larger increase in reserve-deposit ratio which implies that the reduction in lending is even more pronounced.

References

Angrist, J. D., Jorda, O., and Kuersteiner, G. M. (2018). Semiparametric estimates of monetary policy effects: String theory revisited. *Journal of Business & Economic Statistics*, 36(3):371–387.

Auclert, A. (2019). Monetary policy and the redistribution channel. *American Economic Review*, 109(6):2333–67.

Balloch, C. and Koby, Y. (2019). Low rates and bank loan supply: Theory and evidence from japan. *Unpublished manuscript Brown University*, 9(10).

Begenau, J. (2020). Capital requirements, risk choice, and liquidity provision in a business-cycle model. *Journal of Financial Economics*, 136(2):355–378.

Bellifemine, M., Jamilov, R., and Monacelli, T. (2022). Hbank: Monetary policy with heterogeneous banks. https://users.ox.ac.uk/~econ0628/HBANK_BJM.pdf

Bianchi, J. and Bigio, S. (2022). Banks, liquidity management, and monetary policy. *Econometrica*, 90(1):391–454.

Brunnermeier, M. K. and Sannikov, Y. (2016). The i theory of money. Technical report, National Bureau of Economic Research.

Christiano, L., Eichenbaum, M., and Evans, C. (2005). Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of Political Economy*, 113(1):1–45.

Coimbra, N. and Rey, H. (2023). Financial Cycles with Heterogeneous Intermediaries. *The Review of Economic Studies*, page rdad039.

Drechsler, I., Savov, A., and Schnabl, P. (2017). The Deposits Channel of Monetary Policy*. *The Quarterly Journal of Economics*, 132(4):1819–1876.

Drechsler, I., Savov, A., and Schnabl, P. (2021). Banking on deposits: Maturity transformation without interest rate risk. *The Journal of Finance*, 76(3):1091–1143.

Gertler, M. and Karadi, P. (2011). A model of unconventional monetary policy. *Journal of Monetary Economics*, 58(1):17–34. Carnegie-Rochester Conference Series on Public Policy: The Future of Central Banking April 16-17, 2010.

Gertler, M. and Kiyotaki, N. (2010). Chapter 11 - financial intermediation and credit policy in business cycle analysis. volume 3 of *Handbook of Monetary Economics*, pages 547–599. Elsevier.

Jamilov, R. (2021). A macroeconomic model with heterogeneous banks. *Available at SSRN 3732168*.

Jamilov, R. and Monacelli, T. (2023). Bewley banks. *Available at SSRN 3732172*.

Kaplan, G., Moll, B., and Violante, G. L. (2018). Monetary policy according to hank. *American Economic Review*, 108(3):697–743.

Kashyap, A. K. and Stein, J. C. (1995). The impact of monetary policy on bank balance sheets. *Carnegie-Rochester Conference Series on Public Policy*, 42:151–195.

Kashyap, A. K. and Stein, J. C. (2000). What do a million observations on banks say about the transmission of monetary policy? *American Economic Review*, 90(3):407–428.

Kekre, R. and Lenel, M. (2022). Monetary policy, redistribution, and risk premia. *Econometrica*, 90(5):2249–2282.

Khwaja, A. I. and Mian, A. (2008). Tracing the impact of bank liquidity shocks: Evidence from an emerging market. *American Economic Review*, 98(4):1413–42.

Kishan, R. P. and Opiela, T. P. (2000). Bank size, bank capital, and the bank lending channel. *Journal of Money, Credit and Banking*, 32(1):121–141.

Miranda-Agrippino, S. and Ricco, G. (2021). The transmission of monetary policy shocks. *American Economic Journal: Macroeconomics*, 13(3):74–107.

Nakamura, E. and Steinsson, J. (2018). High-Frequency Identification of Monetary Non-Neutrality: The Information Effect*. *The Quarterly Journal of Economics*, 133(3):1283–1330.

Polo, A. (2021). Imperfect pass-through to deposit rates and monetary policy transmission. https://www.bankofengland.co.uk/-/media/boe/files/working-paper/2021/imperfect-pass-through-to-deposit-rates-and-monetary-policy-transmission.pdf

Smets, F. and Wouters, R. (2007). Shocks and frictions in us business cycles: A bayesian dsge approach. *American Economic Review*, 97(3):586–606.

Stein, J. C. (1998). An adverse-selection model of bank asset and liability management with implications for the transmission of monetary policy. *The RAND Journal of Economics*, 29(3):466–486.

Ulate, M. (2021). Going negative at the zero lower bound: The effects of negative nominal interest rates. *American Economic Review*, 111(1):1–40.

Wang, O. (2018). Banks, low interest rates, and monetary policy transmission. *NYU Stern School of Business*.

Whited, T. M., Wu, Y., and Xiao, K. (2022). Will central bank digital currency disintermediate banks? *Available at SSRN 4112644*.

Williams, E. (2020). Costly external financing and monetary policy transmission: evidence from a natural experiment. *Available at SSRN 3080665*.

CURRICULUM VITAE

[redacted]