Welfare Cost of Reserves

Ali Haider Ismail[*] Diego Zúñiga [†]

October 21, 2025

Abstract

We develop a general equilibrium model to quantify the welfare costs of alternative reserve regimes. Banks in our model face stochastic deposit flows requiring settlement through reserves, creating a precautionary demand shaped by frictions in the over-the-counter interbank market. We establish a novel irrelevance theorem, providing conditions under which interbank frictions and reserve quantities have no welfare implications. With frictions, however, banks' liquidity management materially affects credit provision and welfare. Quantitatively, we find that increasing inflation to 10% while holding the real discount window rate constant generates welfare costs of 0.8% of consumption—comparable to Lucas (2000) despite our model excluding currency. The model reproduces the post-2008 tripling of reserve holdings and predicts that raising the real return on reserves initially reduces output as reserves crowd out productive loans, with this effect reversing only after deposit demand becomes satiated. Our results demonstrate that the welfare implications of returning to scarce reserves depend critically on implementation details.

[*]University of California, Los Angeles
[†]Bank of Canada

1 Introduction

Since the 2008 financial crisis, central banks worldwide have fundamentally transformed how they conduct monetary policy. The Federal Reserve expanded bank reserves from less than $500 billion to over $4 trillion at the height of Covid, transitioning from a "scarce reserves" regime—where reserves were minimal and the federal funds market was highly active with yearly trading volumes exceeding $20 trillion—to an "ample reserves" regime where banks hold substantial excess reserves but interbank trading has collapsed to less than $5 trillion each year. This regime shift has sparked an active policy debate: should the Federal Reserve return to the pre-2008 scarce reserves framework, maintain the current ample reserves system, or pursue some intermediate approach? Despite the profound implications for monetary policy implementation and financial stability, we lack a comprehensive framework for evaluating the welfare costs of these different reserve regimes.

The question of what the welfare cost is tied to the cost that bank's face for holding reserves. Intuitively, for each $1 of reserve holdings, it faces an opportunity case based on the difference in investing that reserve holding into a productive asset less the return on the reserve asset paid by the Fed. Consider the US where the eight largest banks held more than $3 trillion worth of high quality liquid asset in June 2025. The back of the envelope measure of the opportunity cost would be <blank>. But to fully quantify this effect, we need a model that can account for the costs and benefits banks face for holding reserves. One one hand, they forgo a more productive asset in the economy for an asset that pays an interest rate fixed by the CB. On the other they gain benefits of reserves in their use for settlement for interbank flows. The model thus would have to capture the effect of monetary policy on the interbank use of reserves, and through that the endogenous response on a bank's portfolio choices.

This paper develops a quantitative general equilibrium model to assess the welfare implications of alternative reserve regimes. Our key insight is that when banks' funding markets are frictionless, the welfare cost of reserves is independent of credit provision—a

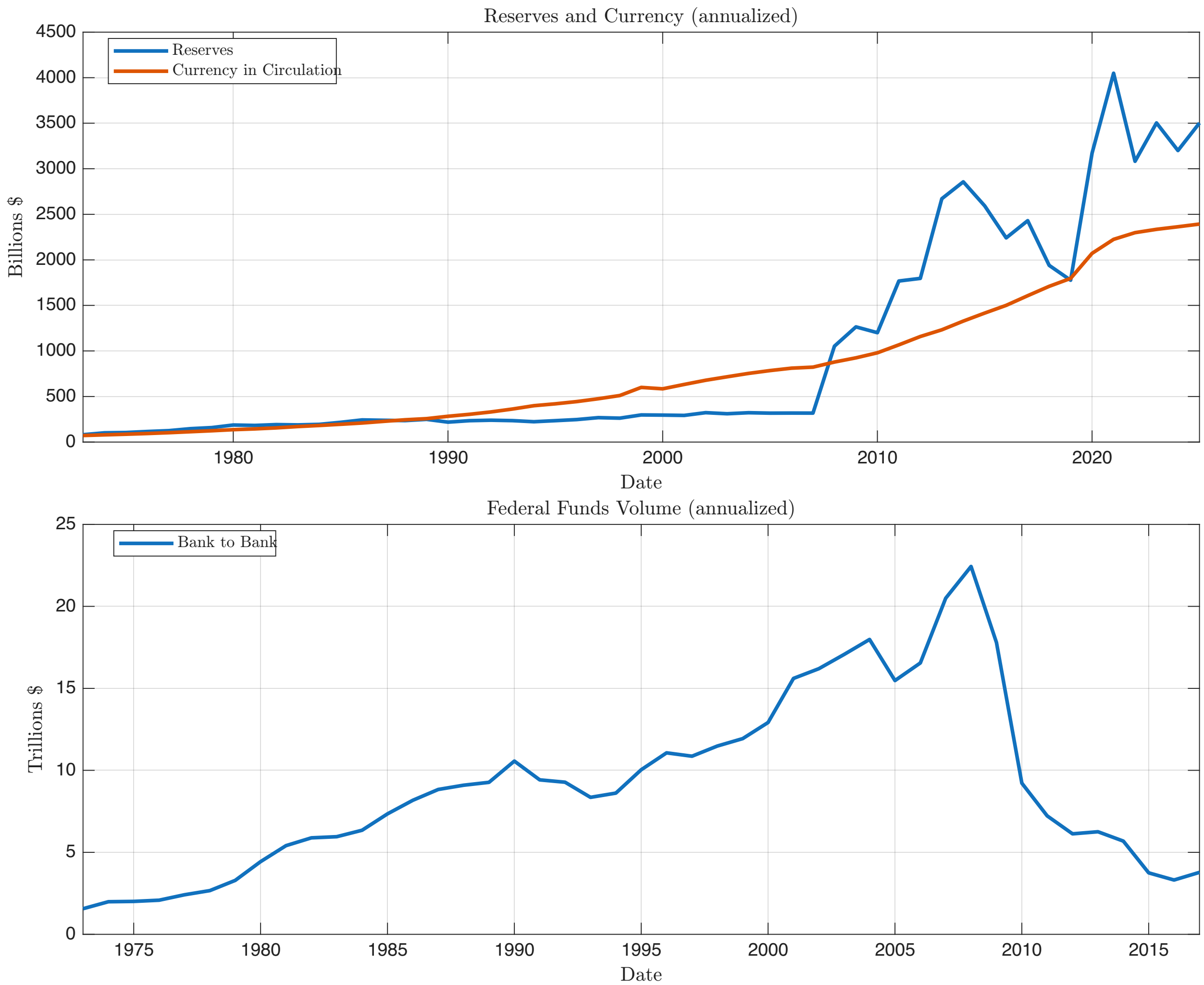

Figure 1: Money Supply and Interbank Volume in Different Regimes

standard result that justifies the traditional focus on household money demand alone. However, once we account for frictions in the interbank market, measuring the welfare cost requires understanding how banks' liquidity management affects the broader economy through the provision of credit and deposits.

We construct a model where banks intermediate savings to entrepreneurs who require loans to fund capital expenditures. Banks face stochastic deposit flows that must be settled using reserves, creating a precautionary motive for reserve holdings. Crucially, banks can borrow reserves in a frictional over-the-counter (OTC) interbank market or, as a last resort, from the central bank's discount window at a penalty rate. This framework allows us to capture the key trade-off banks face: holding more reserves provides insurance against costly liquidity shortfalls but reduces the funds available for productive lending.

Our approach addresses a significant gap in the literature on the welfare cost of inflation and monetary policy. Traditional studies in the spirit of Bailey 1956 and Lucas 2000 focus exclusively on households' money demand, abstracting from the banking sector entirely. While these papers provide valuable insights about the opportunity cost households bear from holding money, they miss a crucial component of the monetary system: banks' demand for reserves and how it affects credit provision. Before 2008, this omission might have been justified given the small quantity of reserves. However, with reserves now representing a significant fraction of bank assets and GDP, any welfare analysis that ignores the banking sector's liquidity management will necessarily be incomplete.

The model delivers several key theoretical insights. First, we establish conditions for what we call a "Interbank Market Irrelevance theorem"—circumstances under which the quantity of reserves, and therefore frictions in the interbank market, have no welfare implications. This result helps clarify when the traditional approach of focusing solely on household money demand is justified and when it fails. Second, we show that the welfare-maximizing policy depends critically on how the central bank implements the Friedman rule: achieving the same real return on reserves can have dramatically different welfare implications depending on whether it is accomplished through adjusting the interest on reserves or the discount window rate.

Our quantitative analysis reveals that the welfare costs of different reserve regimes are substantial and highly sensitive to implementation details. When we hold the real discount window rate constant and vary inflation from 0% to 10%, we find welfare costs of approximately 0.8% of steady-state consumption—comparable to Lucas (2000)'s estimates despite our model having no currency. Strikingly, we show that policies designed to implement the Friedman rule by raising the real return on reserves can initially *reduce* output as reserves crowd out productive loans. Only when deposit demand becomes satiated do additional reserves cease to displace lending. This non-monotonicity in the welfare-output relationship highlights the importance of general equilibrium effects that previous partial equilibrium analyses have missed.

The framework also reproduces key empirical patterns from the transition between reserve regimes. Our model generates the observed three-fold increase in the reserves-to-GDP ratio following the 2008 crisis when the central bank raises the real return on reserves. These patterns emerge endogenously from banks' optimal portfolio choices rather than from assumed structural breaks or policy changes.

Our analysis contributes to several strands of literature. We incorporate insights from Bianchi and Bigio (2022) on banks' liquidity management in OTC markets while extending their model to a GE setting suited to be comparable to previous work in the literature on welfare cost of inflation. Our work also connects to recent studies of monetary policy implementation under different operating frameworks (Lagos and Navarro (2023), Lopez-Salido and Vissing-Jorgensen (2023), and Afonso et al. (2022) by providing a unified framework for welfare evaluation.

1.1 Related Literature

One of the earliest papers to bring light to the opportunity cost that banks face for holding reserves begins with Tobin (1963), who first posited that reserve holdings and reserve requirements bear a cost to banks. The traditional approach to measuring the welfare cost of inflation, focuses on the households opportunity cost of holding money. This literature uses the consumer surplus methodology, as pioneered by Bailey (1956). Subsequent work by Cooley and Hansen (1989), Cooley and Hansen (1991), Lucas (2000), and Lucas and Nicolini (2015), refined these estimates through structural models incorporating different theoretical mechanisms for money demand. More recent contributions include Benati and Nicolini (2021), Serletis and Xu (2021), and Serletis and Xu (2025).

However, Craig and Rocheteau (2008) demonstrate that standard estimates of the welfare cost of inflation may be inaccurate in models with search frictions, as externalities highlighted in the Lagos and Wright (2005) framework create wedges between private and social marginal utilities of real balances. Similarly, Wen (2015) show that in a Bewley economy the area underneath the demand curve approach does not work for similar reasons. By contrast, Alvarez, Lippi, and Robatto (2019) demonstrate that the demand curve approach remains valid in a large class of monetary models like cash in advance, money in the utility, and shopping time. These papers underscore the importance of modeling financial frictions when measuring welfare costs. Dávila, Graves, and Parlatore (2024) provide sufficient statistics for measuring welfare changes of arbitrage gaps which can be also applied to monetary models.

A related strand of literature examines convenience yields on government bonds and near-money assets. Vissing-Jorgensen (2023) discuss how to measure welfare via convenience yields and what the optimal monetary regime is, though our model microfounds these measures structurally. Kurlat (2019) similarly examine how deposit spreads affect welfare in a model where banks have monopoly power in deposit provision. d'Avernas et al. (2025) examine the welfare cost of quantitative easing, complementing our analysis.

Our work complements models that incorporate bank balance sheet constraints as the main intermediary friction. Van den Heuvel (2008), Van Den Heuvel (2022), Wang (2025), and Begenau (2020) focus on how limited equity constraints and capital requirements are important. Our model instead focuses on the liquidity management problem of banks, and the opportunity costs that banks face for holding reserves. Stulz, Taboada, and van Dijk (2022) and Levine and Sarkar (2019) provide empirical evidence on why banks hold substantial amounts of liquid assets, motivating our focus on reserve management.

Payne and Sz\ Hoke (2025) and Williamson (2023) use similar in-flavor models to study different historical questions, while Li, Li, and Sun (2024) examines network structure in money markets with a focus on financial stability. Many of these models, especially the latter, microfound the liquidity shock in the same way - as random flows of deposits.

The remainder of the paper proceeds as follows. Section 2 presents our general equilibrium model of bank liquidity management, detailing the household, firm, and banking sectors along with the OTC interbank market structure. Section 3 characterizes the equilibrium and establishes our main theoretical results, including the conditions for reserve irrelevance. Section 4 presents our quantitative analysis, examining welfare costs under alternative steady states corresponding to different reserve regimes. Section 5 discusses extensions including bank default risk and market power. Section 6 concludes with implications for optimal monetary policy design and the ongoing debate about the appropriate level of reserves in the banking system.

2 GE Model of Bank's Liquidity Managment

The model aims to start with the growth model core and augments it with banking system. Banks intermediate savings in the form of deposits into loans which are used to fund capital holdings. Households motive to save using deposits comes from Sidrauski (1967) "money" in the utility framework. Their motive to borrow using bank loans comes from a binding constraint that capital be funded via long-term loans.

The main financial friction in the model is a cost to intermediation that comes from the market imperfection in the reserve interbank market. Namely, each bank faces stochastic shocks that represent inflows and outflows of deposits. Banks must settle these flows with reserves, inducing a distribution of reserves positions among banks. Banks borrow or lend reserves in a frictional interbank market. Because banks cannot perfectly insure themselves due to the market friction, this induces a cost to their intermediation activities. This cost is affected by monetary policy.

After setting the stage, we go into the problems of each of the agents in the economy. Our analysis of the equilibrium will show how the interbank market and the central bank's provision of reserves affect the issuance of deposits and loans. Then, we can show how the settlement risk is incorporated into measures of the welfare cost.

2.1 Environment

There are three types of agents in the economy: households, firms, and bankers. The measure of households and banks are normalized to one, while the measure of firms are indeterminate. The model is written in real terms with the numeraire being the real good.

Timing and Markets. Time t is discrete and goes forever. Each period t is divided into two stages. The first stage consists of a centralized, competitive market (CM) where households, firms, and banks will make consumption and portfolio decisions. The second stage consists of an interbank market (IM) where banks borrow and lend reserves among themselves to cover their deposit flows. Households and firms are inactive in the IM.

Assets. There are four types of assets in the economy: deposits d, capital k, loans l, and reserves m. All assets are one-period assets. Deposits are issued by banks and can be held by households. In this model, household's hold deposits due to the deposit in the utility [1]. Capital is created through investment and used for production by the firm. Banks intermediate the deposits from savers to loans to borrowers, backed by some fraction φ of capital, to fund capital holdings. Finally, there are central bank reserves which can only be held by bankers with their nominal supply M^g set by the central bank.

Production Technology. Firm's face a neoclassical production function $F(k_t, h_t) = A_t k_t^{1-\alpha} h_t^{\alpha}$. They rent capital and labor from households in the competitive factor markets in the CM. Capital investment is chosen by households which is then used in production in the next period. Capital depreciates at rate $\delta > 0$. The law of motion of capital is

$$k_{t+1} = i_t + (1 - \delta)k_t.$$

Reserves and Central Bank Policies. The central bank issues reserves that can only be held by bankers and must be used for interbank settlement. Reserves earn nominal interest, denoted i^m, to be paid in the next period's CM. The central bank also operates a

1. The appendix details a version of the model which uses Lagos and Wright (2005) to provide a more microfounded reason for monetary exchange using deposits, but the qualitative results remain the same.

discount-window lending facility at the end of the interbank market (IM) where banks can borrow reserves at a nominal interest rate i^d. We assume that the central bank provides reserves elastically in the CM, but commits to a level of inflation. Given this, the real interest rates on reserves and at the discount window are R^m and R^d. The central bank can also set reserve requirements, ϱ, for the bankers. Banks have perfect foresight over the policies, and take as given in period t the policies $R^m_{t+1}, R^w_{t+1}, \varrho$.

2.2 Households

Households discount the future via $\beta < 1$. Households save in the form of deposits where there willingness to save in deposits is governed the utility they gain from deposits. Household's utility function is therefore given by $U(c_t, d_t)$. We assume that $U(\cdot, \cdot)$ is concave, twice continuously differentiable on $\mathbb{R}^2_{++}$, strictly increasing in consumption $\frac{\partial U}{\partial c_t} > 0$ and weakly increasing in deposits $\frac{\partial U}{\partial d_t} \geq 0$. In particular, the forms we will focus on will have point where for $d > d^{sat}$, $\frac{\partial U}{\partial d_t} = 0$. This last assumption is crucial, as we will want a finite satiation point in deposits to make clear our welfare comparisons at the optimal policy [2].

Capital must be intermediated so households take loans out to fund some fraction φ of the next period's capital stock $l^h_{t+1} \geq \varphi k_{t+1}$. They provide a fixed endowment of one unit of labor and rent out capital to firms in competitive markets so take w the wage rate and R^k the return on capital as given.

To describe the activities we can borrow from the description given in Lucas (1980). Each household splits into two members: an entrepreneur/borrower and a shopper/saver. The entrepreneur supplies labor and capital to the firm, decides on new investment i_t, and borrows l_{t+1} to fund the next period's capital stock. Specifically we the following constraint[3]

$$l_{t+1} \geq \varphi \cdot k_{t+1} = \varphi \cdot i_t + \varphi \cdot (1 - \delta) k_t.$$

2. That is, with strictly increasing marginal utility there would exist no finite level of deposits where $\frac{\partial U}{\partial d} = 0$ and so households would hold infinite deposits at the optimal deposit rate

3. This allows for a satiation point in loans $l^{sat} = \varphi k^*$, where k^* is the level of capital in the neoclassical model. The constraint allows for $l > l^{sat}$ if loans are priced efficiently.

Loans are long term assets backed by the returns on capital. Some fraction φ of new investment i_t must be funded by loans. The entrepreneur also rolls over the (unmatured) existing debt on depreciated capital. This can also be understood in terms of corporate finance as a firms who have fixed debt to equity ratios equal to φ. Firms sell their output to the shopper/saver who purchases c_t units of consumption of the good, and the amount of savings d_{t+1}. At the end of the period, the two members reunite to aggregate wealth and consumption.

The problem of the household written in recursive form is therefore

Problem. Let $\mathcal{S}_t$ be a vector corresponding to the aggregate states in the economy. The variable $\mathcal{S}_t = \{d_t, a_t, k_t, l_t,\}$ represent the current composition of household assets at the start of period t. The value function of the household $V^h(\mathcal{S}_t)$ satisfies the following

$$V^h(\mathcal{S}_t) = \max_{c_t, d_{t+1}, l_{t+1}, k_{t+1}} u\left(c_t, d_t\right) + \beta V^h(\mathcal{S}_{t+1})$$

such that

$$c_t + d_{t+1} + R_t^l l_t + i_t \leq w_t \cdot 1 + R_t^k k_t + R_t^d d_t + l_{t+1} + div_t + T_t$$

$$k_{t+1} = i_t + (1-\delta)k_t$$

$$l_{t+1} \geq \varphi \cdot k_{t+1}$$

Note the div_t represents that the household earns the dividends from the bank[4].

2.3 Firms

Firms in this model are neoclassical. They rent capital k^f and labor h^f in the CM and face a production function $y = F(k^f, h^f)$ which is constant return to scale (CRS). Due to the CRS nature of the production function, the measure of firms is indeterminate. Their objective is

4. We should think that loans are made to different banks similar to Gertler and Kiyotaki 2010, though we abstract from their islands formulation for simplicity

given by

$$\max_{k_t^f, h_t^f} F(k_t^f, h_t^f) - R_t^k \cdot k_t^f - w_t \cdot h_t^f.$$

2.4 Banks

Banks last for one period, remitting their net worth to households, and every period new banks are set up. Banks finance loans l^b to households which are funded through deposits d^b.

We model the costs of a banks portfolio (see Freixas and Rochet (2008)) as coming from a microfoundation of banks liquidity management problem. Namely, that banks who issue deposits, face exogenous stochastic inflows and outflows, and must settle these flows with central bank reserves m[5]. Banks who are in deficit in reserves can borrow from those who are in surplus in an overnight frictional interbank market. The market friction prevents all banks in deficit from matching. In that case banks can borrow from the central bank's lender of last resort function via the discount window at a penalty rate. This induces a steeper marginal cost of being in deficit than the marginal benefit of being in surplus.

The key friction that banks face in this model the imperfection of the interbank market which we model based off Bianchi and Bigio (2022) (here in BB). This allows us to focus on how monetary policy changes the costs of intermediation and induces banks to hold more or less reserves. Due to the assumptions in BB, in a stationary equilibrium, the economy will reduce into a sequence of repeated static problems.

Banks in CM. Existing banks start period t with a portfolio of deposits, loans, and reserves d_t, l_t, m_t along with interbank loans f_t (negative indicating lending to other banks) and discount window borrowing from the central bank w_t borrowed in period $t-1$ IM. At the beginning of the CM, each of these bank pays interest on it's deposits, interbank loans, and discount window borrowing R_t^d, R_t^{ib}, R^w, earns interest on it's loans and reserves positions R_t^l, R^m.

5. That is we assume l are illiquid in the IM. We abstract from other liquid assets like US Treasuries in this treatment to focus on central bank reserves.

We assume that there is full commitment for banks to their liabilities so there is no default in deposits, interbank loans, nor discount window loans. Similar assumptions apply for the asset side of bank's portfolios, in that the households cannot default on loans and the central bank always pays interest on it's reserves.

Banks net worth is given

$$e_t = R^l_t \cdot l_t + R^m \cdot m_t - R^d_t \cdot d_t - R^{ib}_t f_t - R^w w_t.$$

which is then remitted to the household.

New banks enter and choose a portfolio $d_{t+1}, l_{t+1}, m_{t+1}$ to maximize their portfolio returns, taking into account the expected costs of going to the interbank market, detailed below. The market for deposits and loans are competitive so banks take R^l_{t+1}, R^d_{t+1} as given. Although there is a large literature arguing that banks have market power in deposits (Drechsler, Savov, and Schnabl 2017) we focus our analysis on the case of no market power as a baseline.

New banks enter with no capital, so their flow of funds constraint is given by $l_{t+1} + m_{t+1} = d_{t+1}$. Thus their assets must equal their liabilities, and their only source of funding is deposits. Denote the measure of banks as $j \in \mathcal{J}$ which is normalized to unity and kept constant so that the same amount of banks enter in each period. In this set up, we abstract from other sources of financing for banks, such as equity and non-demandable deposits and the frictions that may exist in these sources [6].

Banks in IM. At the beginning of the IM stage, banks face an exogenous[7] shock representing flows of deposits that must be settled with reserves. The shocks induces distributions in reserve positions, where banks short on reserves go to an OTC market subject to search frictions where they can settle their deficits by borrowing reserves. Due to matching frictions,

6. See Van den Heuvel (2008), Wang (2025) for examples of papers who focus on equity frictions

7. Note that in this treatment the shocks and their distributions are not affected by the household's problem. We think of these flows coming from an unmodeled payment process where deposits flow within the banking system.

not all surplus/deficit positions will match. An unmatched bank in deficit must borrow from the discount window at the discount window rate R^w, whereas an unmatched bank in surplus earns the interest on reserves R^m.

Formally, each bank j faces two i.i.d. shocks representing deposit flows. The inflow shock is denoted ξ^j. The bank receives an inflow of deposits $\xi^j \cdot D$, where D is the aggregate level of deposits across banks. The inflow shock ξ^j is a positive random variable with distribution Φ_ξ distributed i.i.d across banks. The outflow shock is denoted ω^j. The bank loses $\omega^j \cdot d^j$ of its deposits to other banks[8]. The outflow shock is a positive random variable with distribution Φ_ω also i.i.d across banks. Moreover, we assume that ω^j is bounded above by $\omega_H \leq 1$. If $\omega_H = 1$, the bank could lose all of its deposits and can only insure it can face outflows without using the discount window if it runs a narrow bank.[9] As shown later, the case of ω_H will be important in certain cases to ensure existence of an equilibrium where banks are satiated in reserves.

The bank's net reserve position at the start of the interbank market will then be given by

$$s^j = m^j_{t+1} + f\left(-\omega^j \cdot d^j_{t+1} + \xi^j \cdot D_{t+1} - \varrho\left(d^j_{t+1} - \omega^j \cdot d^j_{t+1} + \xi^j \cdot D_{t+1}\right)\right).$$

where x^j denotes the position of an individual bank, and $f(x) = \max\{x, \upsilon\}$ for some parameter υ to distinguish between settlement types detailed below. The first part, m^j_{t+1}, indicates the amount of reserves they chose in the CM in period t. Turning to the arguments of f, the next term $-\omega^j \cdot d^j_{t+1} + \xi^j \cdot D_{t+1}$ indicates the flows of deposits that are settled with reserves. Finally, $\varrho\left(d^j_{t+1} - \omega^j \cdot d^j_{t+1} + \xi^j \cdot D_{t+1}\right)$ indicates the fraction of end of period deposits that must be covered via reserves due to regulatory constraints. That is the end of period position in reserves must be at least as ϱ fraction of the end of period deposits. The

8. Note the difference that outflows are based on a bank's own position in deposits where inflows are based on the aggregate position of deposits. This can be microfounded via a more detailed matching process. See the appendix for a version of the model that is built off Lagos and Wright (2005).

9. This assumption is different from Bianchi and Bigio (2022) who assume only one shock to the deposit position of banks. With only one shock, a bank that increases its issuance of deposits can face larger outflows but also larger inflows. This can lead to the settlement cost of deposits being non-monotonic in deposit issuance. With separable shocks, deposits are always risky at the margin.

parameter ϱ can stand in for reserve requirements or more generally liquidity constraints (like the LCR).

We use the function $f(\cdot)$ to distinguish between two types of timing conventions that we interpret as different settlement types in the IM, fast vs slow. In turn, it will impact the distribution $\{s_j\}_j$ and therefore the rates offered to banks that we study. We denote fast (or net) settlement by $\upsilon = \infty$ to indicate that banks can use all of the inflows of reserves from deposits to settle same day outflows. We denote slow (or gross) settlement by $\upsilon = 0$ to indicate that banks can only use inflows to meet their regulatory requirements ϱ, but not to be used to settle for deposit outflows.

Banks face a non-negativity constraint of their reserve positions, so they settle their deficits by borrowing reserves in an OTC market subject to search frictions. The shocks induce a distribution in $\{s^j\}_{j\in\mathcal{g}}$, where $s^j > 0$ indicates a surplus position and $s^j < 0$ indicates a deficit position. Banks cannot end the period with $s^j < 0$, so these banks go to the interbank market to borrow reserves. On the other end of the market, banks with surplus, $s^j > 0$, can potentially earn a return by lending their reserves to the banks in deficit.

This OTC market is modeled as a two-sided decentralized market with a large family assumption; each bank is split into infinitesimally small traders who match traders from other banks. Specifically, we assume that traders match sequentially over a finite horizon with matching intensity λ and matching function G. At the end of the period, all the traders will join back up to their individual bank. This is akin to the assumption used in Atkeson, Eisfeldt, and Weill (2015).

The implication of the large family assumption is that trader's individual outcomes in the OTC market have no impact on the net worth of bank. This makes the model very tractable and the marginal benefits or costs of a bank's position depend only in the aggregate allocation. In essence, each bank is able to take the "average trade" in the OTC market. As in Bianchi and Bigio (2023), we can obtain closed form solutions for the endogenous rates that arise from this friction[10].

We define the aggregate positions and market tightness as

$$S^{+} \equiv \int_{0}^{1} \left[s^{j}\right]^{+} \, \mathrm{d}j$$
$$S^{-} \equiv -\int_{0}^{1} \left[s^{j}\right]^{-} \, \mathrm{d}j$$
$$\theta \equiv \frac{S^{-}}{S^{+}}$$

where we have dropped time subscripts for brevity. Aggregate surplus is defined as the sum over all banks who have surplus positions after the liquidity shock. Aggregate deficit is defined analogously. Market tightness is a measure of how many deficits positions there are relative to surplus. Naturally, the average rate at which deficit banks borrow reserves rises with θ, reflecting the greater supply of loanable reserves.

Lemma 1. *If banker's portfolio decisions are symmetric, then tightness in the interbank market depend only on the ratio of aggregate reserves to aggregate deposits, $M/D = \mu$.*

Proof. Here we show the proof for fast settlement, when $\upsilon = \infty$. The proof for slow settlement is largely similar, but more notation, so is left for the appendix. If decisions are symmetric, then $d^{j} = D$ and $m^{j} = M$ for all j. Then,

$$\begin{aligned} S^{+} &= \int_{0}^{1} \left[M + \xi^{j} \cdot D - \omega^{j} \cdot D - \varrho \left(D - \omega^{j} \cdot D + \xi^{j} \cdot D\right)\right]^{+} \, \mathrm{d}j \\ &= D\left(1-\varrho\right) \cdot \int \left[s^{*} + \xi - \omega\right]^{+} \, \mathrm{d}\Phi_{\xi,\omega}, \\ S^{-} &= \int_{0}^{1} \left[M + \xi^{j} \cdot D - \omega^{j} \cdot D - \varrho \left(D - \omega^{j} \cdot D + \xi^{j} \cdot D\right)\right]^{-} \, \mathrm{d}j \\ &= D\left(1-\varrho\right) \cdot \int \left[s^{*} + \xi - \omega\right]^{-} \, \mathrm{d}\Phi_{\xi,\omega}, \\ \theta &= \frac{\int \left[s^{*} + \xi - \omega\right]^{+} \, \mathrm{d}\Phi_{\xi,\omega}}{\int \left[s^{*} + \xi - \omega\right]^{-} \, \mathrm{d}\Phi_{\xi,\omega}} \end{aligned}$$

where we denote $s^{*} = \frac{\mu - \varrho}{1 - \varrho}$. □

10. A review of the derivations and closed form expressions are provided in the appendix.

This result is important because the bank's marginal payoffs will be a function of the market tightness. In the banker's portfolio problem, the banker will take as given the market tightness.

As a result of this trading process, the banker's payoffs from the interbank market can be written as a function, χ, which is piece-wise linear with a kink at zero

$$\chi(s, \theta) = \chi^{+}(\theta) \cdot [s]^{+} + \chi^{-}(\theta) \cdot [s]^{-}.$$

Here, $[x]^{+} = \max\{x, 0\}$ and $[x]^{-} = \min\{x, 0\}$, while $\chi^{+}(\theta)$ is the marginal benefit of being in surplus, and $\chi^{-}(\theta)$ is the marginal cost of being in deficit. Key to this model's tractability is that banks face a piecewise linear function, i.e. the magnitude of s does not affect the marginal return the bank faces. This assumption is well suited for markets where the assets being traded are very liquid, and so price elasticities are nearly zero.

The specific value for the marginal benefit and cost take into account that matching is frictional. In particular, an unmatched bank in deficit will need to borrow from the discount window; similarly, an unmatched bank in surplus will not be able to lend all of its position to deficit banks and so earns its outside option the interest on reserves. Specifically,

$$\begin{aligned}
\chi^{+} &= \Psi^{+}(\theta, \lambda) \left(R^{ib}(\theta) - R^{m} \right) \\
\chi^{-} &= \Psi^{-}(\theta, \lambda) \left(R^{ib}(\theta) - R^{m} \right) + \left(1 - \Psi^{-}(\theta, \lambda)\right) \left(R^{w} - R^{m}\right).
\end{aligned}$$

Here R^{ib} denotes the average interest rate in the interbank market. A bank in surplus can match a fraction $\Psi^{+}(\theta, \lambda)$ of its surplus to deficit banks and earns a net payoff of $R^{ib}(\theta) - R^{m}$ for each unit lent. An unmatched bank in surplus simply earns R^{m}, so it's net return is 0. A bank in deficit can borrow a fraction $\Psi^{-}(\theta, \lambda)$ of its deficit from surplus banks at a net cost $R^{ib}(\theta) - R^{m}$, but must cover the rest of its deficit at the discount window at a net cost $R^{w} - R^{m}$. We assume that the discount window is provided fully elastically so that settlement fully clears. Due to the large family assumption, its outcome only depends on the beginning of market tightness, θ, and parameters. Hence we obtain closed forms for the endogenous probabilities $\Psi^+(\theta, \lambda), \Psi^-(\theta, \lambda)$ and interbank rate $R^{ib}(\theta)$ which allow us to solve the model. Derivations of the interbank market payoffs are shown in Appendix C.

Since the banks lives only one period, their problem is given by maximizing next period's expected profit[11].

Problem 1.

$$\max_{l_{t+1}, m_{t+1}, d_{t+1}} \mathrm{E}_t \left[\Lambda_{t+1}\left(e_{t+1}\right)\right]$$

where

$$e_{t+1} = \overbrace{R^l_{t+1} l_{t+1} + R^m_{t+1} m_{t+1} - R^d_{t+1} d_{t+1}}^{\text{Exp. Portfolio Return}} \\ + \underbrace{\mathrm{E}_{\omega,\xi}\left[\chi(m_{t+1}, d_{t+1}, D_t, \theta_t) | D_t, \theta_t, \iota\right]}_{\text{Settlement Costs}}$$

Since the bank is owned by the household, they discount next period expected profits via the SDF of the household Λ_{t+1}. Banks choose their portfolios $l_{t+1}, m_{t+1}, d_{t+1}$ in the CM to maximize its profits which consists of two objects; the first being their expected portfolio return, the second its gains or losses from the interbank market (IM) represented by the expectation of χ over the measure of ω, ξ, taking aggregates $D_t, \theta_t, \iota_{t+1}$ as given. Monetary policy affects the payoffs in three ways. First R^m affects the direct return on reserves. Next, the payoffs in the interbank market is proportional to the corridor gap, $\iota \equiv R^w - R^m$. Finally ϱ influences the distribution of $\{s_j\}_j$ and θ_t by changing the threshold for which banks are in deficit or surplus.

2.5 Central Bank

The central bank in this model has access to several policy tools that it can use to influence the bank's returns. It has access to the nominal interest on the discount window i^w, nominal

11. Since inflows are fixed to the aggregate deposits, banks face decreasing returns to scale and will generate positive profit in equilibrium.

interest rate on reserves i^m, money supply M^g, discount window loans W and transfers T. In addition, it sets the reserve requirement ϱ. Its budget constraint is given by

$$R^m \frac{M_t^g}{P_{t-1}} + \frac{W_{t+1}}{P_t} = \frac{M_{t+1}^g}{P_t} + R^w \frac{W_t^g}{P_{t-1}} + T_t. \tag{1}$$

The central bank funds itself using current-period real transfers T_t, earnings from the discount window loans from last period W_t, and current period money issuance M_{t+1}^g. Its expenses are current period discount window loans W_{t+1}^g, and paying out interest on reserves M_t. Note that the reserves include borrowed reserves, so the net interest margin on discount window borrowing is $R^w - R^m$. As is convention in the rest of the model, payments are made at the beginning of the period, so the central bank pays out interest on reserves and earns interest on discount window loans simultaneously. Note that $\frac{M_{t+1}^g}{P_t}$ is the money supply inclusive of discount window lending. Define $\frac{\tilde{M}_{t+1}^g}{P_t}$ the money supply less discount window borrowing via $M_{t+1}^g \equiv \tilde{M}_{t+1}^g + W_{t+1}^g$. In other words, $\tilde{M}^g$ is the monetary policy decision in the CM of the problem where as M^g is determined at the end of the period after the IM is resolved.

2.6 Equilibrium Definition

Definition 1. Given a set of policy rates R^m, R^w and reserve requirements ϱ, a symmetric stationary equilibrium consists of value functions $\{V^h\}$, CM allocations $\{c, d^h, l^h, k, k^f, h^f l, m, d\}$, IM allocations $\{S^+, S^-, \theta, \Psi^+, \Psi^-\}$, CM prices and interest rates $\{R^d, R^k, R^l, \}$, IM prices $\{\chi^-, \chi^+, R^{ib}\}$ such that

- Value function solve the Bellman equations,
- Households choose c, d^h, l^h, k optimally given prices and value functions
- Firms choose k^f, h^f optimally given prices
- Banks choose l, m, d optimally given prices
- IM allocations and prices are a solution to the trading problem

- Markets for CM good, assets, and deposits clear.

For our analysis, T and W_t^g are set passively so that the central bank budget constraint is met and the discount window clears . The government policies set deterministically are $M_t^g, i_t^m, i_t^w, \varrho_t$. In the stationary equilium that we are interested in, we consider policies where M^g grows at a constant rate which via quantity-theory-like arguments pins down inflation. In turn the central bank can choose $R^m = \frac{1+i_m}{1+\pi}, R^w = \frac{1+i^w}{1+\pi}$.

2.7 Equilibrium Analysis

We solve the problems of each of the agents and show properties of the equilibrium.

Non-financial sector The first order conditions of the households problem and firm problem give rise to the "supply" of deposits and demand for loans that the bank takes as given. Taking first order conditions of the household's problem leads to

$$\begin{aligned}
[d_{t+1}]: \qquad & \mathrm{E}_t\left[\Lambda_{t+1}\left(R_t^f - R_{t+1}^d\right)\right] = \mathrm{E}_t\left[\Lambda_{t+1}\frac{u_d(c_{t+1}, d_{t+1})}{u_c(c_{t+1}, d_{t+1})}\right] \\
[k_{t+1}, l_{t+1}]: \qquad & \mathrm{E}_t\left[\Lambda_{t+1}\left(R_{t+1}^k + (1-\delta) - R_t^f\right)\right] = \varphi\, \mathrm{E}_t\left[\Lambda_{t+1}\left(R_{t+1}^l - R_t^f\right)\right]
\end{aligned}$$

where

$$\begin{aligned}
\Lambda_{t+1} &= \frac{\zeta_{t+1}}{\zeta_t}\beta\frac{u_c(c_{t+1}, d_{t+1})}{u_c(c_t, d_t)} \\
R_t^f &= \frac{1}{\mathrm{E}_t\left[\Lambda_{t+1}\right]}.
\end{aligned}$$

is the SDF from t to $t+1$ and the risk free rate. The first equation writes the demand for deposits based on the opportunity cost[12] of deposits $R^f - R^d$. The firm's problem is standard

12. Belongia and Ireland 2019 argue that this form is more amenable to having the model match the construction/theory behind Divisia money indices

leading to

$$
\begin{aligned}
R^k &= \frac{\partial F(k,1)}{\partial k} = (1-\alpha)Ak^{-\alpha} \\
w &= \frac{\partial F(k,1)}{\partial h} = \alpha A k^{1-\alpha}
\end{aligned}
$$

.

we have have employed that labor supply is normalized to unity.

The working capital friction manifests in a wedge in the real return on the marginal product of capital. In this model R^k will be set to a convex combination of R^f and R^l. That is when $\varphi > 0$ and $R^l > R^f$ then capital is more costly than in the neoclassical benchmark where $R^k = R^f$. In turn, the household holds less capital. We will see that this will be the channel/wedge through which monetary policy will have a real effect.

Banks The first order conditions of the banks will pin down a demand for deposits from the banking system and a demand for reserves that take into frictions in the interbank market.

$$
\begin{aligned}
\mathrm{E}_t\left[\Lambda_{t+1} R^l_{t+1}\right] &\geq \mathrm{E}_t\left[\Lambda_{t+1}\left(R^m_{t+1} + \frac{\partial\, \mathrm{E}_{\omega,\xi}[\chi(m_{t+1}, d_{t+1})]}{\partial m}\right)\right] \\
\mathrm{E}_t\left[\Lambda_{t+1} R^l_{t+1}\right] &= \mathrm{E}_t\left[\Lambda_{t+1}\left(R^d_{t+1} - \frac{\partial\, \mathrm{E}_{\omega,\xi}[\chi(m_{t+1}, d_{t+1})]}{\partial d_{t+1}}\right)\right]
\end{aligned}
$$

Notice that the first-order condition for reserves holds with equality if $m_{t+1} > 0$. In the following, we drop the time subscripts, and the household SDF while introducing some extra notation to make the exposition easier. Since $\mathrm{E}[\chi(m,d)] = \chi^+ \cdot [s]^+ + \chi^- \cdot [s]^-$, we can write these first-order conditions as

$$
\begin{aligned}
R^l_{t+1} - R^m_{t+1} &\geq \mathcal{L}^{\mathrm{m}}(m_{t+1}, d_{t+1}) \equiv \frac{\partial\, \mathrm{E}_{\omega,\xi}[\chi(m_{t+1}, d_{t+1})]}{\partial m_{t+1}}, \\
R^f_{t+1} - R^d &= \mathcal{L}^{\mathrm{d}}(m_{t+1}, d_{t+1}) \equiv -\frac{\partial\, \mathrm{E}_{\omega,\xi}[\chi(m_{t+1}, d_{t+1})]}{\partial d_{t+1}}
\end{aligned}
$$

Here, $\mathscr{L}^{\mathrm{m}}$ is the partial derivative of $\mathrm{E}[\chi]$ with respect to m, while $\mathscr{L}^{\mathrm{d}}$ is (the negative of) that with respect to d. We interpret $\mathscr{L}^{\mathrm{m}}$ and $\mathscr{L}^{\mathrm{d}}$ as the liquidity premia on reserves and deposits arising from the OTC friction in the interbank market.

The liquidity premia will take a particular form depending on the payoffs and distribution of surpluses.

$$Ł^{m} = \chi^{-} \Pr\left(\right).$$

Here $\Pr(s > 0)$ represents the probability of a bank being in surplus, which will depend on its choices m, d, aggregates D, as well as the distribution of the shocks ω, ξ. In the second equation, $\mathrm{E}_{\omega,\xi}\left[\omega | s < 0\right]$ represents the conditional expectation of the outflow shock, conditional on being in deficit. Since the deposit inflows are dependent on the aggregate position of which the bank takes as given, only the conditional expectation of the outflow shock shows up in the first order condition for deposits. The appendix details closed forms $\mathscr{L}^{m}, \mathscr{L}^{d}$ in terms of the interbank market functions and the distributions of ω, ξ.

Lemma 2. *The banker's portfolio (CM) problem is jointly concave in m, d.*

Proof. Here we show the argument for when $\varrho = 0$ and fast settlement. The argument for $\varrho > 0$ is similar. Substitute the budget constraint in the objective function of the banker. Then, the derivative with respect to m is given by $R^{m} + \mathscr{L}^{\mathrm{m}}(m, d) - R^{f}$, as shown in the first-order condition. The second derivative with respect to m is then given by

$$\frac{\partial^2}{\partial m^2} = \frac{\partial\, \mathscr{L}^{\mathrm{m}}(m, d)}{\partial m} = -(\chi^{-} - \chi^{+}) \int \delta(m + \xi \cdot D - \omega \cdot d)\, \mathrm{d}\Phi_{\omega,\xi} \leq 0,$$

where $\delta(x)$ is the Dirac delta function which equals one if $x = 0$. Similarly, the cross-derivative is

$$\frac{\partial^2}{\partial m \partial d} = \frac{\partial\, \mathscr{L}^{\mathrm{m}}(m, d)}{\partial d} = (\chi^{-} - \chi^{+}) \int \omega \cdot \delta(m + \xi \cdot D - \omega \cdot d)\, \mathrm{d}\Phi_{\omega,\xi} \geq 0.$$

The derivative with respect to d is given by $R^{f} - R^{d} - \mathscr{L}^{\mathrm{d}}(m, d)$, as shown in the first-order condition. The second-derivative with respect to d is

$$\frac{\partial^2}{\partial d^2} = -\frac{\partial \mathcal{L}^{\mathrm{d}}(m,d)}{\partial d} = -(\chi^- - \chi^+) \int \omega^2 \cdot \delta(m + \xi \cdot D - \omega \cdot d)\, \mathrm{d}\Phi_{\omega,\xi} \leq 0.$$

Notice that the expressions for the second derivative tell us that the problem is concave in m and d separately. To show that the problem is jointly concave we must show that the determinant of the Hessian matrix is positive. That is, we must show that

$$\left[\int \delta(m + \xi \cdot D - \omega \cdot d)\, \mathrm{d}\Phi_{\omega,\xi}\right] \cdot \left[\int \omega^2 \cdot \delta(m + \xi \cdot D - \omega \cdot d)\, \mathrm{d}\Phi_{\omega,\xi}\right] \\ \geq \left[\int \omega \cdot \delta(m + \xi \cdot D - \omega \cdot d)\, \mathrm{d}\Phi_{\omega,\xi}\right]^2,$$

but this is a simple application of the Cauchy-Schwartz inequality.[13] □

Given the fact that shocks are separable, this property is easy to establish and guarantees that we can rely on the first-order conditions for the analysis.

Corollary 1. *$\mathcal{L}^{\mathrm{m}}$ and $\mathcal{L}^{\mathrm{d}}$ are decreasing in m and increasing in d.*

This shows that the bank has a decreasing marginal benefit of acquiring reserves and an increasing marginal cost of issuing deposits.

Proposition 1. *The bank's optimal portfolio decision consists of holdings of reserves and deposits such that the first-order conditions hold. In particular, we have two cases: either the bank holds some reserves or none. If the bank holds some reserves, then the bank's holdings must solve*

$$R^l - R^m = \mathcal{L}^m(\mu, R^w, R^m),$$

$$R^l - R^d = \mathcal{L}^d(\mu, R^w, R^m),$$

13. The Cauchy-Schwartz inequality states $\mathrm{E}[A] \cdot \mathrm{E}[B] \geq \mathrm{E}[A \cdot B]^2$. Take $A = \sqrt{\delta(m + \xi \cdot D - \omega \cdot d)}$ and $B = \omega \cdot \sqrt{\delta(m + \xi \cdot D - \omega \cdot d)}$

If the bank holds no reserves, then we must have

$$R^l - R^m \geq \mathscr{L}^m(\mu, R^w, R^m),$$
$$R^l - R^d = \mathscr{L}^d(\mu, R^w, R^m),$$

The bank's decision to hold reserves trades off a benefit and a cost. The cost is the interest rate spread relative to the risk-free asset, $R^l - R^m$. The benefit is twofold, avoiding a deficit position with their costly borrowing of reserves and increasing surplus with their profitable lending of reserves. The decision to issue deposits trades off the benefit of earning the interest rate spread $R^l - R^d$ with the settlement cost. This settlement cost includes more borrowing when in deficit and less lending when in surplus.

It is easy to show then that the bank will hold less reserves and issue more deposits if the cost of discount-window borrowing is lower. Define $\iota \equiv R^w - R^m$ to represent the corridor gap. The proposition then explores what happens if the CB lowers ι, for a fixed R^m.

Proposition 2. *Given prices, interest rates, and matching probabilities in the IM, a bank will hold less reserves and issue more deposits if the cost of discount-window borrowing, ι, is lower.*

Proof. Again we show the proof for $\varrho = 0$ and fast settlement, with the other cases relegated to the appendix. Since the cross-derivative of the objective function with respect to m and d is positive, the objective function is supermodular in m, d. Therefore, we can show monotonicity results for m and d with respect to a parameter if the cross-derivative of the objective function with respect to the choice variable and the parameter always takes the same sign. (Single crossing property).

Toward showing the single-crossing property, notice that χ^- and χ^+ are increasing in ι. Therefore,

$$\frac{\partial^2}{\partial m \partial \iota} = \frac{\partial \mathscr{L}^{\mathrm{m}}}{\partial \iota} = \frac{\partial \chi^-}{\partial x} \Pr(s < 0) + \frac{\partial \chi^+}{\partial x} \Pr(s > 0)$$

and

$$\frac{\partial^2}{\partial d \partial x} = -\frac{\partial \mathcal{L}^{\mathrm{d}}}{\partial x} = -\frac{\partial \chi^-}{\partial x} \Pr(s<0)\,\mathrm{E}[\omega|s<0] - \frac{\partial \chi^+}{\partial x} \Pr(s>0)\,\mathrm{E}[\omega|s>0].$$

Therefore, if the sign of the derivatives of χ^-, χ^+ are the same, the single-crossing property holds. Any parameter that increases χ^- and χ^+ will induce an increase in reserves and a decrease in issuance of deposits. □

We can also show that banks will hold more reserves and issue more deposits if R^m increases.

Proposition 3. *Given prices, interest rates and matching probabilities in the IM, a bank will hold more reserves and issue more deposits if R^m is higher.*

Proof. Following the same type of argument from the proof of the previous proposition, we can write

$$\frac{\partial^2}{\partial m \partial R^m} = \frac{\partial R^m + \mathcal{L}^{\mathrm{m}}}{\partial R^m} = 1 - \Pr(s<0) - \Pr(s>0)\Psi^+ \geq 0,$$

where we see that R^m directly increases the benefit of holding reserves, but it decreases the benefit of reserves in settlement. However, the latter effect cannot dominate the former. Similarly, for deposits,

$$\frac{\partial^2}{\partial d \partial R^m} = \frac{\partial - \mathcal{L}^{\mathrm{d}}}{\partial R^m} = \Pr(s<0)\,\mathrm{E}[\omega|s<0] + \Pr(s>0)\,\mathrm{E}[\omega|s>0]\Psi^+ \leq 0,$$

which shows that deposits will increase because an increase in R^m lowers the cost of settlement.

□

In a symmetric equilibrium, banks will choose $d = D$. Notice that $\mathcal{L}^{\mathrm{m}}(M, D)$ and $\mathcal{L}^{\mathrm{d}}(M, D)$ evaluated at the aggregate quantities depend only on the reserve ratio $\mu = M/D$. Therefore, we can write $\mathcal{L}^{\mathrm{m}}(\mu), \mathcal{L}^{\mathrm{d}}(\mu)$. Thus, we can establish the following results

Proposition 4. *Let M, D be the aggregate holdings of reserves and deposits in equilibrium.*

Then, banks will hold no reserves, $M = 0$, if

$$R^l - R^m \geq \mathcal{L}^{\mathrm{m}}(0), \qquad R^l - R^d = \mathcal{L}^{\mathrm{d}}(0).$$

On the other hand, if banks hold some reserves, then

$$R^l - R^m = \mathcal{L}^{\mathrm{m}}(\mu), \qquad R^l - R^d = \mathcal{L}^{\mathrm{d}}(\mu).$$

In words, if the aggregate premium on reserves when banks don't hold reserves, $\mathcal{L}^{\mathrm{m}}(0)$ is too low, then banks will hold no reserves. Alternatively, if $\mathcal{L}^{\mathrm{m}}(0) > R^f - R^m$, then banks must hold some reserves in equilibrium. This shows that there is a cutoff rate on reserves below which banks hold no reserves. This cutoff depends on the friction in the interbank market.

3 Benchmarks without Liquidity Costs and Irrelevance

Using this model, we go through successive exercises to show how settlement risk that are settled in a friction interbank market are passed onto consumers. This risk informs us how to measure welfare.

3.1 Benchmarks without Liquidity Costs

AHI: Update these figures, to match the calibration below

Our first limiting case is when there are no interbank frictions. This can be represented by taking the interbank matching rate, $\lambda \to \infty$. In this case, we get the following behavior

$\theta < 1$

$$\Psi^+ = \theta \quad \Psi^- = 1$$

$$R^{ib} = R^m \quad \chi^+ = 0 \quad \chi^- = 0$$

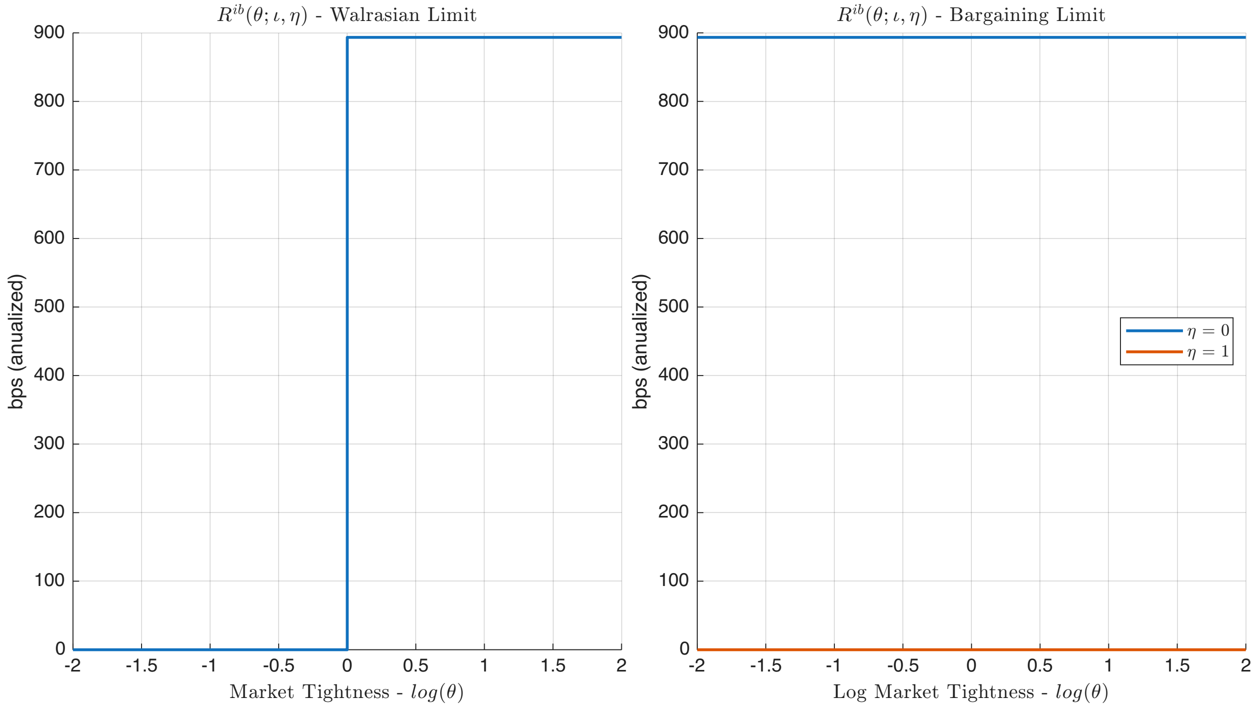

Figure 2: Limiting Cases for Interbank Rate

$\theta > 1$

$$\Psi^{+} = 1 \quad \Psi^{-} = \frac{1}{\theta}$$
$$R^{ib} = R^{w} \quad \chi^{+} = 0 \quad \chi^{-} = 0$$

The intuition here is that when $\theta < 1$ there is more surplus position than deficit, so the efficient price in the interbank market is the outside option for the lending bank R^m. On the other hand, if $\theta > 1$, there are more deficit positions than surplus, so the efficient price is for $R^{ib} = R^w$.

In either case we will get $\mathcal{L}^m = \mathcal{L}^d = 0$ and therefore there is no welfare cost arising from the settlement risk for banks.

Note however that this is not equivalent to one of the polar cases of the bargaining parameter η, which represents the share of surplus that the *borrowing* bank receives. When $\eta = 0$, the lending bank receives all the surplus and so sets the interbank rate to R^w *regardless* of θ. Similarly when $\eta = 1$, the borrowing bank receives all the surplus and sets $R^{ib} = R^m$ always.

Next we turn to the Friedman rule which is a common benchmark in monetary models. The intuition in the Friedman rule is since the social cost of providing liquidity is zero, the central bank should equalize the marginal benefit of reserves with the risk free rate in the economy. We show that in the interbank market, there is another form of efficiency that the central bank can implement to ensure that bank's do not face settlement risk. We show the model counterfactual of how central bank policy can implement this benchmark. This counterfactual has never been observed in the data, demonstrating the importance of a structural macro model in computing the welfare cost. The frictionless benchmark helps policy makers understand what central bank policy would minimize the welfare cost of money.

Proposition 5. *A necessary condition for efficiency is that $\mathcal{L}^d = \mathcal{L}^m = 0$. This is achieved either*

1. $i^w = i^m$

2. *or* $\Pr(s_t < 0) = 0$

Proof. To avoid ponzi schemes, rates must satisfy $R^l \geq R^f \geq R^d$. The efficient outcome is when the rates for deposits and loans are equal to the frictionless case, where $R^l = R^f = R^d$. This occurs only if $\mathcal{L}^d = 0$. Therefore $\mathcal{L}^m = 0$.

If $\iota = 0, \chi^+ = \chi^- = 0$, therefore $\mathcal{L}^{\mathrm{m}} = \mathcal{L}^{\mathrm{d}} = 0$. If $Pr(s < 0) = 0, S^- = 0, \chi^+ = \chi^- = 0$, and so again $Lm = 0, Ld = 0$. $\square$

The first condition represents that the marginal cost of funding a deficit position is the same as the marginal benefit of a surplus position. In other words, the the kink in χ arising from $\chi^- > \chi^+$ is removed. Note that the level of the rates do not matter here, only that they are equal to each other. To understand this condition, consider a bank in deficit. The bank with some probability is matched in the OTC interbank market, where it borrows at the interbank rate R^{ib} which due to the corridor rate being equal will be set so that $R^{ib} = R^w = R^m$. If the bank does not match in the interbank market, it funds its deficit position by going to the discount window. Hence, regardless of the matching process, it can always fund its deficit position at the marginal holding return of reserves. In other words, since the central bank can provide liquidity costlessly at the discount window, it can achieve efficiency by closing the corridor gap.

The second condition is met iff $\mu = \omega_H <= 1$ where bank's hold enough reserves to insure against the most extreme deposit outflow shock. If $\mu = \omega_H = 1$, this corresponds to a narrow bank. Banks in this case are satiated in their precautionary demand for reserves.

Corollary 2. *Efficiency requires*

1. *In a monetary equilibrium* $m, d > 0$, $R^f = R^d = R^m = \frac{1}{\beta}$ *and* $\mathcal{L}^m = \mathcal{L}^d = 0$

2. *In a nonmonetary equilibrium* $R^f = R^d = \frac{1}{\beta}, \mathcal{L}^m = \mathcal{L}^d = 0, R^w = R^m < \frac{1}{\beta}$

The first case corresponds to the Friedman rule in standard monetary models. Here the central bank has two degrees of freedom to set the real return on reserves to the risk free rate - either it deflates the economy or it increases the nominal interest on reserves i^m, or a combination of both. Due to proposition 5 we see that there are two cases for this equilibrium to hold. In the case that the central bank sets $i^w = i^m, \rightarrow R^w = R^m = 1/\beta$, there are multiple equilibria as bank's are indifferent from holding any mix of assets a, m.

The second case corresponds to the central bank eliminateing settlement risk for the banking system by providing reserves at the discount window equal to the return on reserves. Banks in this case will choose to not hold reserves in the CM. Interbank transfers are funded fully by banks in deficit going to the discount window. In this case the central bank would in effect be "subsidizing" the use of the discount window below the risk free rate since there is no cost for the central bank to supply discount window loans in this model.

3.2 Special Case: Irrelevance of Interbank Frictions

Here we show a novel contribution - a case where interbank frictions are irrelevant. The special case provided here of the model provides conditions in which the amount of reserves, and therefore monetary policy, have no impact on welfare. The result can be thought of as a way to nest the previous literature on the welfare cost of inflation Lucas 2000; Craig and Rocheteau 2008 in a model with banking.

In this set up we take a limiting case where the distribution of outflows ω is degenerate and $\varphi = 0$. In other words, bank face deterministic outflows and the demand curve for loans is flat. We show that in this case the effect of settlement risk is fully captures as the cost of acquiring enough reserves to cover the deterministic outflows.

Proposition 6. *In the case where Φ_ω is degenerate and we are in fast settlement, the liquidity premia take the following form*

$$\mathcal{L}^{\mathrm{m}}(\mu, \Theta) = \chi^{-}(\Theta) - \Phi_z(s^*)(\chi^{-}(\Theta) - \chi^{+}(\Theta)),$$
$$\mathcal{L}^{\mathrm{d}}(\mu, \Theta) = \omega \cdot \left\{\chi^{-}(\Theta) - \Phi_z(s^*)(\chi^{-}(\Theta) - \chi^{+}(\Theta))\right\} = \omega \cdot \mathcal{L}^{\mathrm{m}}(\mu, \Theta),$$

where $s^* = \frac{\tilde{m}}{\tilde{d}} = \mu$ represents the threshold for which a bank is in deficit or suprlus. Then $\Phi_z(s^*)$ represents the probability of a bank being in deficit. In a symmetric equilibrium, the threshold is solely determined by the equilibrium reserve ratio μ.

They result takes into account the assumption that bank's face a flat demand for loans, and so their supply of loans is not a margin of adjustment for their portfolio returns via affecting the equilibrium rate on loans. The conditions imply that the demand curve for reserves of the banking system is proportional to the demand curve for deposits of the banking system. This is the type of result that is used to justify why reserves are not part of M1 and higher monetary aggregates due to money multiplier type arguments that the quantity for deposits is some multiple of the quantity of reserves. The result here shows that this only holds in a specific special case in the model.

In fact, in this case, the quantity of reserves and interbank frictions do not for welfare. From the first-order conditions for m, d

$$R^f - R^d = \mathcal{L}^{\mathrm{d}} = \omega \cdot \mathcal{L}^{\mathrm{m}} = \omega \cdot (R^f - R^m).$$
$$\Rightarrow R^d = (1-\omega)R^f + \omega R^m.$$

In this special case, the rate on deposits is entirely determined by R^m. The other policy rate R^w and the search-and-matching parameters (bargaining power or matching rate) do not affect the interest rate on deposits or the quantity of deposits. Since welfare depends on the quantity of deposits, these parameters do not affect welfare, although they do affect the amount of reserves held by banks.

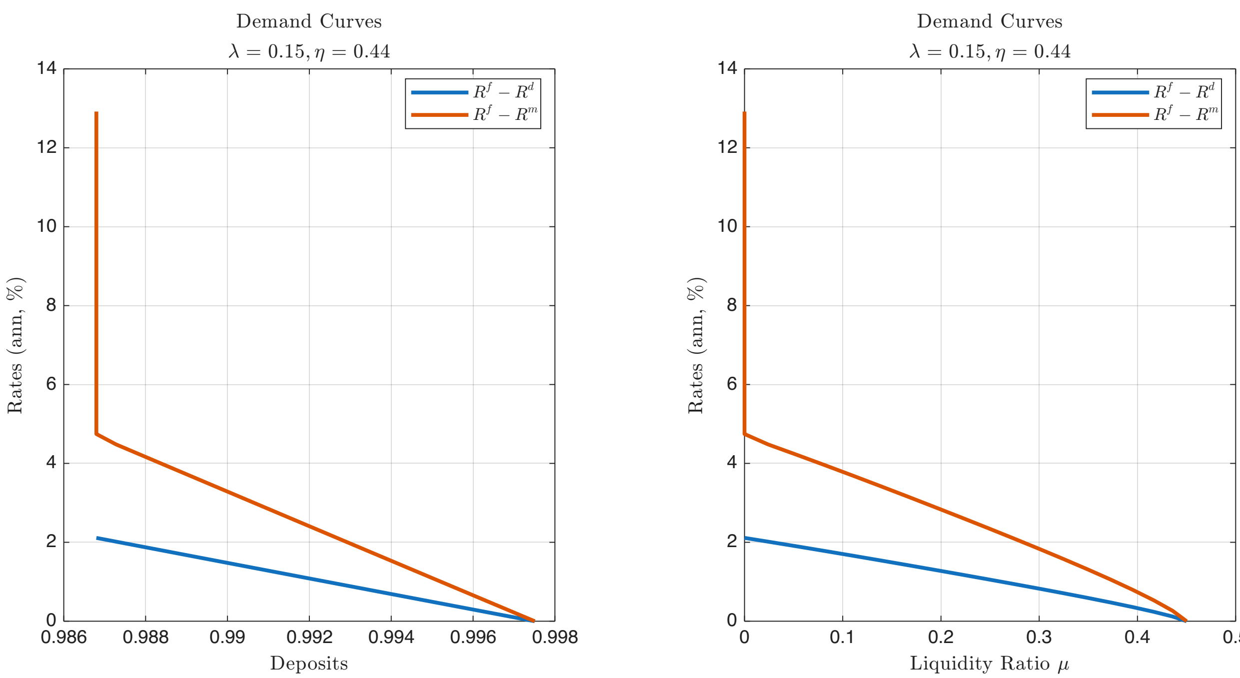

Figure 3: Demand Curves of Reserves and Deposits of Banks

Figure 3 plots the demand curves for reserves $(R^f - R^m)$ and deposits $(R^f - R^d)$ arising from the FOC of the bank. The left hand panel plots it as a function of the equilibrium deposits, while the right hand panel does it as a function of the equilibrium reserves. In either case, we can see that we would overstate the true welfare cost of inflation if we were to measure welfare using the opportunity cost of "money" as per Bailey's method using $R^f - R^m$. Since deposits pay interest, the opportunity cost of holding them is lower, relative to cash.

The vertical line in these plots in $R^f - R^m$ represent the non-monetary equilibria. Once the policy rate, R^m, goes low enough (interpreted equivalently as the central bank inflating the economy), the bank goes to a corner solution and holds no reserves. In this case the R^m, and therefore inflation π, is irrelevant for the welfare cost since $R^f - R^d$ is not changing.

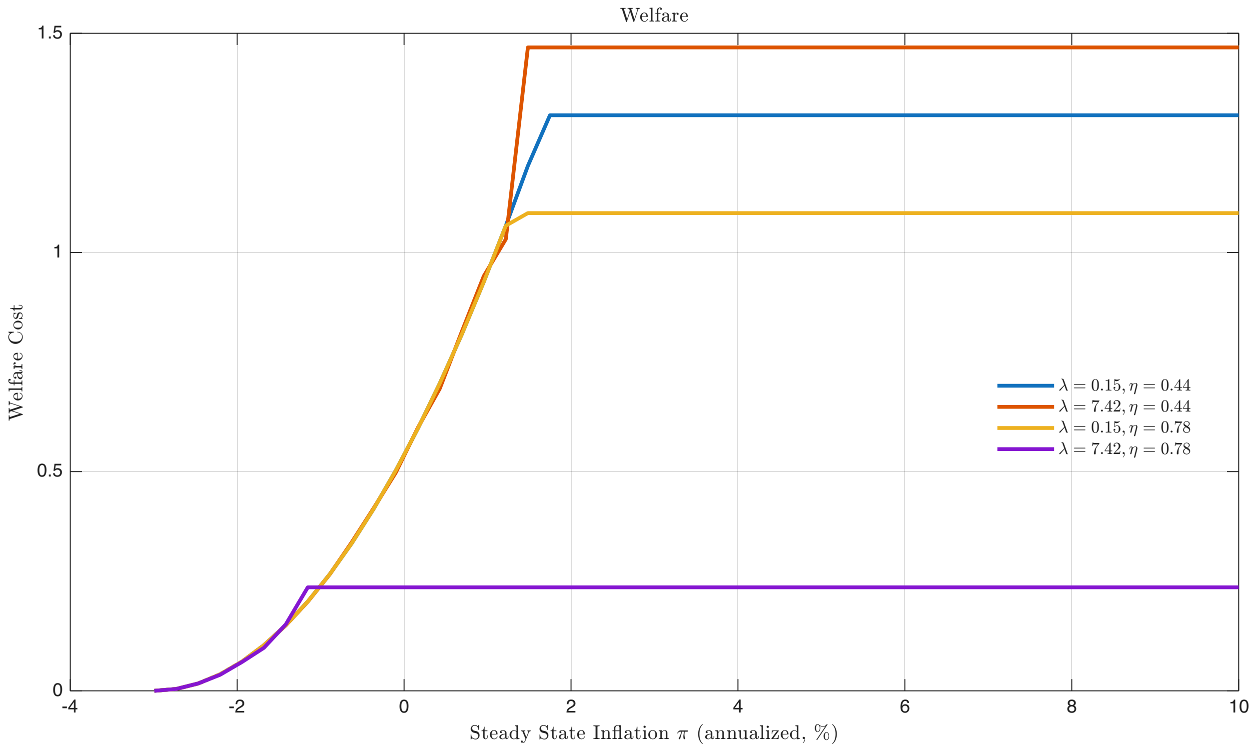

Figure 4: Welfare as Interbank Parameters Vary

Figure 4 shows the welfare cost in terms of the welfare measure presented above, under various interbank parameter combinations. Notice that the lines overlap where we are in a monetary equilibria. They diverge as changing the interbank parameters changes the marginal costs that the bank faces, and so changes the point at which they no longer hold reserves. Importantly, the parameter η which is the bargaining power of the bank in deficit, has a bigger impact. Even with a relatively high value of $\lambda = 7.42$, a change in η from 0.78 to 0.44 changes the cost that the bank faces drastically. As $\eta \to 0$, bargaining power goes to the lender, and the $R^{ib} \to R^w$. Hence the cost of being in deficit is higher, so bank's hold reserves for "longer".

We can interpret this special case of the model as the type of assumptions that would need to hold for the reserve regimes to not matter for the estimate of welfare. In other words, if these assumptions hold, it is enough to use the methods the previous literature on the welfare cost of infaltion Lucas 2000; Craig and Rocheteau 2008 by focusing on the household's demand for monetary assets. We can look past the banking system, as they are just an intermediary in this economy with no real impact, and focus on the household's holdings.

We turn into the next section into cases where the interbank frictions are relevant.

4 Quantifying Welfare Cost across Steady State

We estimate the equilibrium variables and welfare cost in steady state as we change the central bank policies in the model. We specialize functional forms to derive demand functions for deposits and loans and calibrate the economy to match portfolio variables and interbank market data. We provide several different scenarios representing different ways the central bank can implement the Friedman rule, each resulting in different equilibrium reserves regimes.

4.1 Quantiative Assumptions

For the quantitative exercise, we need to make some assumptions about utility. In this section we derive the demand and inverse demand curves for deposits and loans. We characterize their elasticities and how they map to primitive parameters. For these, we focus on steady state.

First take

$$u(c,d) = \frac{c^{1-\gamma^c}-1}{1-\gamma^c} - \frac{\min(d,\hat{d})}{\gamma^d}\left(\log(\min(d,\hat{d})) - \Theta^d - 1\right).$$

where $\hat{d}$ represents the point deposit satiation. For $d > \hat{d}$, $u_d = 0$ and deposits provide no liquidity service. The variable Θ^d represents a stochastic level-demand shifter for deposits that is uncorrelated with the SDF. We assume it has a stationary markov process.

We can solve for the demand functions for deposits via obtaining the inverse demand function from the first order condition

$$R^f - R^d = \begin{cases} \frac{c^{\gamma^c}}{\gamma^d}\left(\Theta^d - \log(d)\right) & \text{if } d < \hat{d} \\ 0 & \text{if } d \geq \hat{d} \end{cases}$$

$$R^d = \begin{cases} R^f - \frac{c^{\gamma^c}}{\gamma^d}\left(\Theta^d - \log(d)\right) & \text{if } d < \hat{d} \\ R^f & \text{if } d \geq \hat{d} \end{cases}$$

Solving for d we get

$$\log(d) = \begin{cases} \Theta^d - \frac{\gamma^d}{c^{\gamma^c}}\left(R^f - R^d\right) & \text{if } R^d < R^f \\ \left[\Theta^d, \infty\right) & \text{if } R^f = R^d \end{cases}$$

$$d = \begin{cases} \exp(\Theta^d) \cdot \exp\left(R^f - R^d\right)^{-\frac{\gamma^d}{c^{\gamma^c}}} & \text{if } R^d < R^f \\ \left[\exp(\Theta^d), \infty\right) & \text{if } R^f = R^d \end{cases}.$$

This satisfies the semi-log form used in previous literature (Lucas (2000) and Belongia and Ireland (2019)) but with $-\gamma^d/c^{\gamma^c}$ being the semi-elasticity of demand[14]. In partial equilibrium, where Θ^d, c are fixed, this exhibits a constant semi-elasticity. In general equilibrium, the elasticity reduces as $c \to c^*$.

14. Typically elasticities are measured with respect to net rates. Even though R^f, R^d denote gross rates in this notation, their difference is a net rate.

We can derive the demand curve for loans via $l = \varphi k$

$$
\begin{aligned}
R^k &= (1-\alpha)Ak^{-\alpha} \\
\beta(R^k - R^f) &= \varphi\beta(R^l - R^f) \\
R^k &= R^f + \varphi(R^l - R^f) \\
k &= ((1-\alpha)A)^{\frac{1}{\alpha}} \left(R^k\right)^{-\frac{1}{\alpha}} \\
&= ((1-\alpha)A)^{\frac{1}{\alpha}} \left(R^f + \varphi(R^l - R^f)\right)^{-\frac{1}{\alpha}} \\
l &= \begin{cases} \varphi\left((1-\alpha)A\right)^{\frac{1}{\alpha}} \left(R^f + \varphi(R^l - R^f)\right)^{-\frac{1}{\alpha}} & \text{if } R^l > R^f \\ [l^*, \infty) & \text{if } R^l = R^f \end{cases} .
\end{aligned}
$$

where l^* is the optimal amount of loans in the economy, see below. At $R^l = R^f$ the household is indifferent with borrowing more at the risk-free rate.

This implies that the inverse demand for loans is given by

$$
\begin{aligned}
R^l &= \begin{cases} \frac{1}{\varphi}\left[\left(\frac{\varphi}{l}\right)^{\alpha}((1-\alpha)A) - (1-\varphi)R^f\right] & \text{if } l < l^* \\ R^f & \text{if } l \geq l^* \end{cases} \\
R^l - R^f &= \begin{cases} \frac{1}{\varphi}\left[\left(\frac{\varphi}{l}\right)^{\alpha} \cdot (1-\alpha) \cdot A - R^f\right] & \text{if } l < l^* \\ 0 & \text{if } l \geq l^* \end{cases} .
\end{aligned}
$$

We can invert this to get the demand function for loans.

$$
\begin{aligned}
\bar{l} &= \varphi\left((1-\alpha) * A\right)^{\frac{1}{\alpha}} \\
\epsilon^l &= -\frac{1}{\alpha}
\end{aligned}
$$

Note that ϵ^l does not correspond to the loan elasticity. To get this we'd compute

$$\begin{aligned}\frac{\partial l}{\partial R^l} &= -\frac{1}{\alpha}\cdot\varphi\cdot l\left(R^f+\varphi(R^l-R^f)\right)^{-1}\\ \frac{\partial l}{\partial R^l}\cdot\frac{R^l}{l} &= -\frac{1}{\alpha}\cdot\varphi\cdot R^l\left(R^f+\varphi(R^l-R^f)\right)^{-1}.\end{aligned}$$

When $\varphi \in (0,1)$ the elasticity changes based on R^l. We can see that this elasticity is decreasing in R^l (increasing in absolute value). For $\varphi = 0$, loan demand is flat at $R^l = \frac{1}{\beta}$. For $\varphi = 1$, we get $\epsilon^l = -\frac{1}{\alpha}$ and we recover the elasticity exactly. The maximal elasticity is achieved when $R^l = \frac{1}{\beta}$ and is $-\frac{\varphi}{\alpha}$.

4.2 Calibration

Here we detail our method for calibrating certain parameters given data targets. For the steady states of Θ^d, A, we can target the deposits and loans demand shifters in equilibrium. We choose to target the average level of reserves and average level of deposits from 2000-2008. Given these values of m, d we can define

$$\begin{aligned}\mu &= \frac{m}{d}\\ l &= d-m\\ R^l &= R^m+\mathcal{L}^m(\mu)\\ R^d &= R^m+\mathcal{L}^m(\mu)-\mathcal{L}^d(\mu)\end{aligned}$$

We can invert the demand curve for l to get $\bar{l}$ (where we need to impose our assumption about $\mathscr{L}^g$) which also gets us A, c

$$\begin{aligned}
\bar{l} &= \frac{l}{((1-\varphi)\cdot R^f + \varphi \cdot R^l))^{\epsilon^l}} \\
A &= (\bar{l}/\varphi)^{\alpha} \frac{1}{(1-\alpha)} \\
c &= y - i = A\left(\frac{l}{\varphi}\right)^{1-\alpha} - \delta \cdot l/\varphi
\end{aligned}$$

We can do the same exercise for deposits

$$\begin{aligned}
\bar{d} &= d \exp(R^f - R^d)^{-\epsilon^d} \\
\Theta^d &= \log(\bar{d})
\end{aligned}$$

External Parameter Calibration

- Depreciation rate, $\delta = .12$
- Labor share $\alpha = .7$
- Risk Aversion $\gamma_c = 2$
- $\epsilon^l = \epsilon^d =$ which will be calibrated to the consumption in steady state since our specification has the elasticity change
- In the baseline case we set $\varphi = 1$ though we could calibrate this to the amount of loans in the economy

4.3 Steady State Computation

Estimating the model in steady state allows us to focus on the long-run behavior of the model and mimic the exercises done in Lucas (2000) and Craig and Rocheteau (2008). In each of the scenarios, we aim to trace out the response function of each of the equilibrium variables as the government policies, represented by the pair of real interbank rates set by the central bank (R^m, R^w) change. The response functions would trace out different steady states in deterministic economies where the central bank implemented a constant policy of (R^m, R^w). This allows us to abstract form higher-frequency frictions, like price stickiness, and focus on the long the run demand of reserves by the banking sector.

The steady state of the model can be summarized via the pair (μ, d). Namely the steady state of the model can be summarized via the following equations

$$R^l = R^m + \mathcal{L}^m(\mu, R^m, R^w) \tag{2}$$

$$R^l = R^d + \mathcal{L}^d(\mu, R^m, R^w) \tag{3}$$

$$l = l(R^l) \tag{4}$$

$$d = d(R^d) \tag{5}$$

$$l = d \cdot (1 - \mu) \tag{6}$$

This system can be rewritten as a function of two equations in two unknowns in (μ, d). The first two equations are the first order conditions of the banking system and represent the indifference conditions in loans and reserves and loans and deposits. The terms $\mathcal{L}^m, \mathcal{L}^d$ represent the liquidity frictions in the interbank market that will respond endogenously to central bank policies (R^m, R^w). Equation (6) is derived from the budget constraint $l + m = d$ of the bank whereas eqs. (4) and (5) come from the non-financial sector of the model and represent the demand functions for loans and deposits. Each of these demand functions feature a flat portion where demand is satiated, an important feature of the model to make sense of the demand of deposits and loans in the frictionless interbank market case.

To compute the steady, we employ a fixed point algorithm in (μ, d) to accommodate the kinks in the loan and deposit demand functions and corner cases. To compute the steady state we recognize that for a given candidate value of $\hat{\mu}$, we can compute

$$\hat{R}^l = R^m + \mathcal{L}^m(\hat{\mu}, R^m, R^w)$$
$$\hat{R}^d = R^m + \mathcal{L}^d(\hat{\mu}, R^m, R^w)$$
$$\text{s.t.}$$
$$\hat{R}^l \geq \frac{1}{\beta} \quad \& \quad \hat{R}^d \leq \frac{1}{\beta}.$$

where the constraints in the third line represent necessary conditions for an equilibrium. Given candidate rates $\hat{R}^l, \hat{R}^d$ we can use the demand functions to compute a new value for μ via

$$\mu^* = 1 - \frac{l}{d} = 1 - l(\hat{R}^l)/d(\hat{R}^d).$$

It is not necessary for $l(R^l), d(R^d)$ to be smooth differentiable functions, and so accommodates demand systems where there are kinks at the satiation point. Given characteristics of $\mathcal{L}^m, \mathcal{L}^d, l(R^l), d(R^d)$ this algorithm is well-behaved (shown delegated to the appendix). Moreover we test for corner cases where the bank holds no reserves by verifying $\frac{1}{\beta} > R^m + \mathcal{L}^m(\mu, R^m, R^w), \forall \mu \in [0, 1]$ to indicate that banks are in a corner case.

4.4 Steady State Policies

The first scenario shows that a central bank can implement an efficient outcome where banks hold reserves at the Friedman rule without flooding the real quantity of reserves. Figure 5 panel (c-d) demonstrates the demand curve's for reserves and deposits from the banking system (captured by the liquidity functions $\mathcal{L}^m, \mathcal{L}^d$) that shows that while banks will increase their demand for deposits without much change in their demand for reserves. Figure 6 shows the full equilibrium, including output and welfare cost, which are both monotonic towards the optimal policy. In this scenario, the central bank provides discount window borrowing enough to satiate bank's demand for liquidity and sets the return on reserves to the risk free

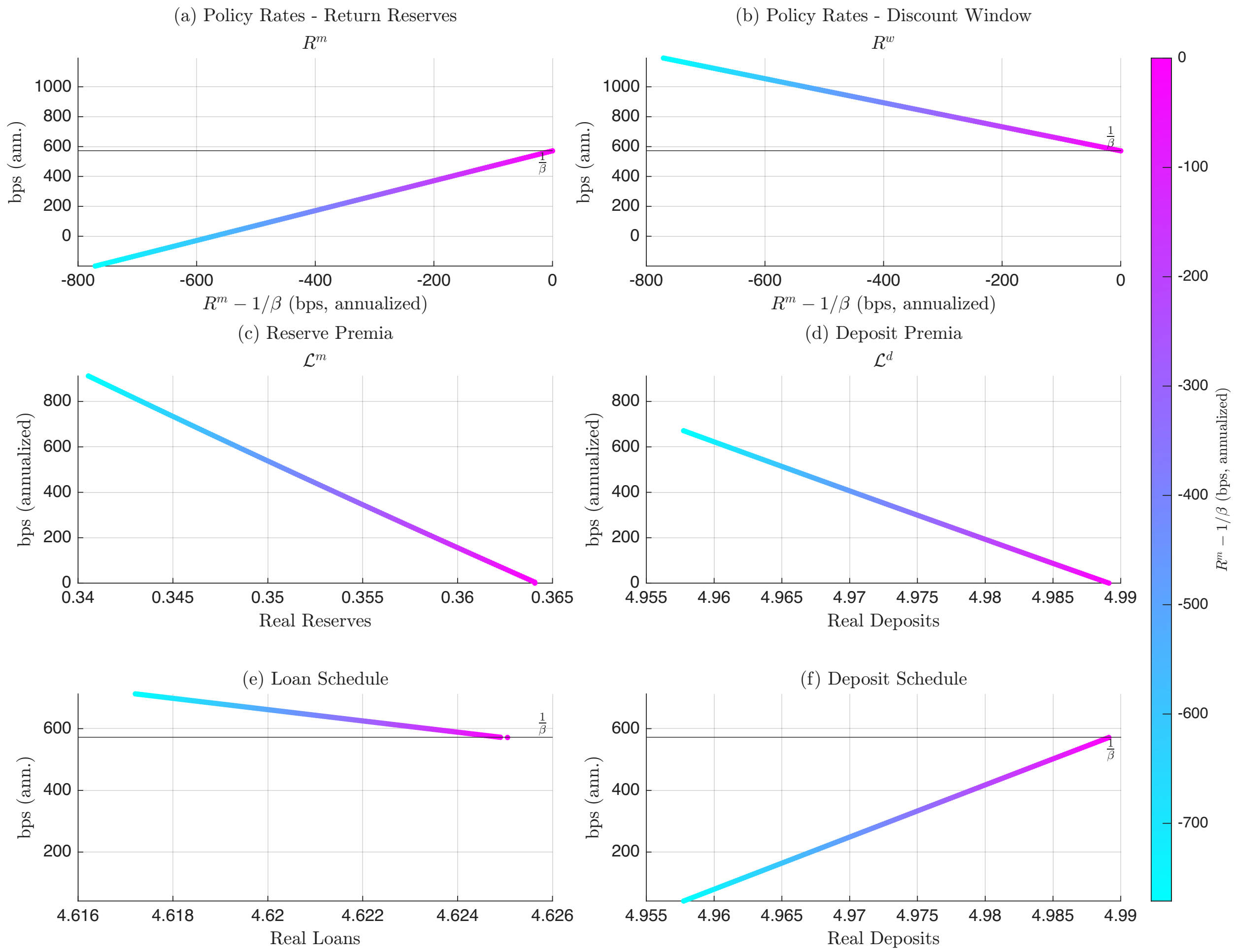

Figure 5: Policies such that $R^m \to \frac{1}{\beta}$ and $R^w \to \frac{1}{\beta}$, where each "dot" represents the steady state under a different policy.

rate $R^m \to \frac{1}{\beta}$ and $R^w \to \frac{1}{\beta}$. In turn, the liquidity frictions will vanish at the optimal policies, $\mathcal{L}^m \to 0, \mathcal{L}^d \to 0$, while bank will satisfy it's indifference equation for holding reserves. Panel (c) shows that bank's holdings of reserves do not change much in this equilibria. As $R^m \to \frac{1}{\beta}$, reserves become less costly to hold but since $R^w - R^m \to 0$, $\mathcal{L}^m \to 0$ and so reserves become less useful for insurance against the liquidity shocks. Instead they increase their balance sheet size by intermediating more deposits converting them into productive loans.

The next scenario is a regime where the central banks focuses on making it less costly to hold reserves and so bank's hold much more reserves at the optimal. Figure 7 demonstrates that banks will hold more reserves in equilibria, replicating the qualitative result in the data from 2000-2025 that banks triple there holdings of reserves in real terms. In this scenario

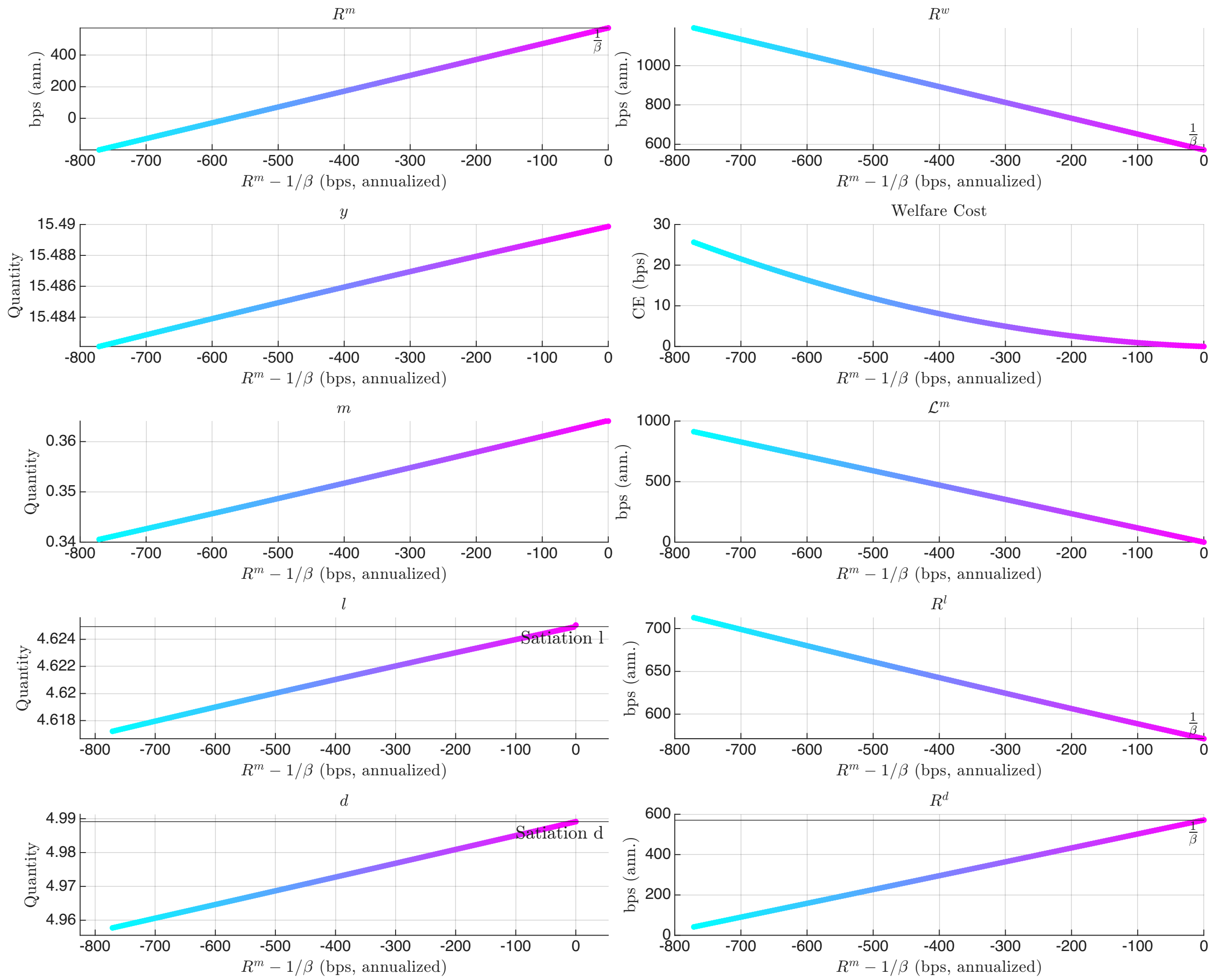

Figure 6: Policies such that $R^m \to \frac{1}{\beta}$ and $R^w \to \frac{1}{\beta}$, where each "dot" represents the steady state under a different policy.

the Friedman rule is implemented via the real return on money ($R^m \to \frac{1}{\beta}$) while R^w stays constant. Since liquidity is still costly $R^w - R^m > 0$ at the Friedman rule level of R^m, banks respond by holding more reserves, since the marginal cost of holding reserves $R^l - R^m$ is decreasing.

The striking fact in this scenario is that the initial response is that reserves crowd out loans as the central bank implements the Friedman rule. Figure 8 shows that output is decreasing in the region where reserves crowd out loans. In the steep portions of the demand curve for reserves (panel (c), fig. 7), the banking system holds more reserves while intermediating less loans. In turn output decreases as the policy is implemented. It is only as the demand curves

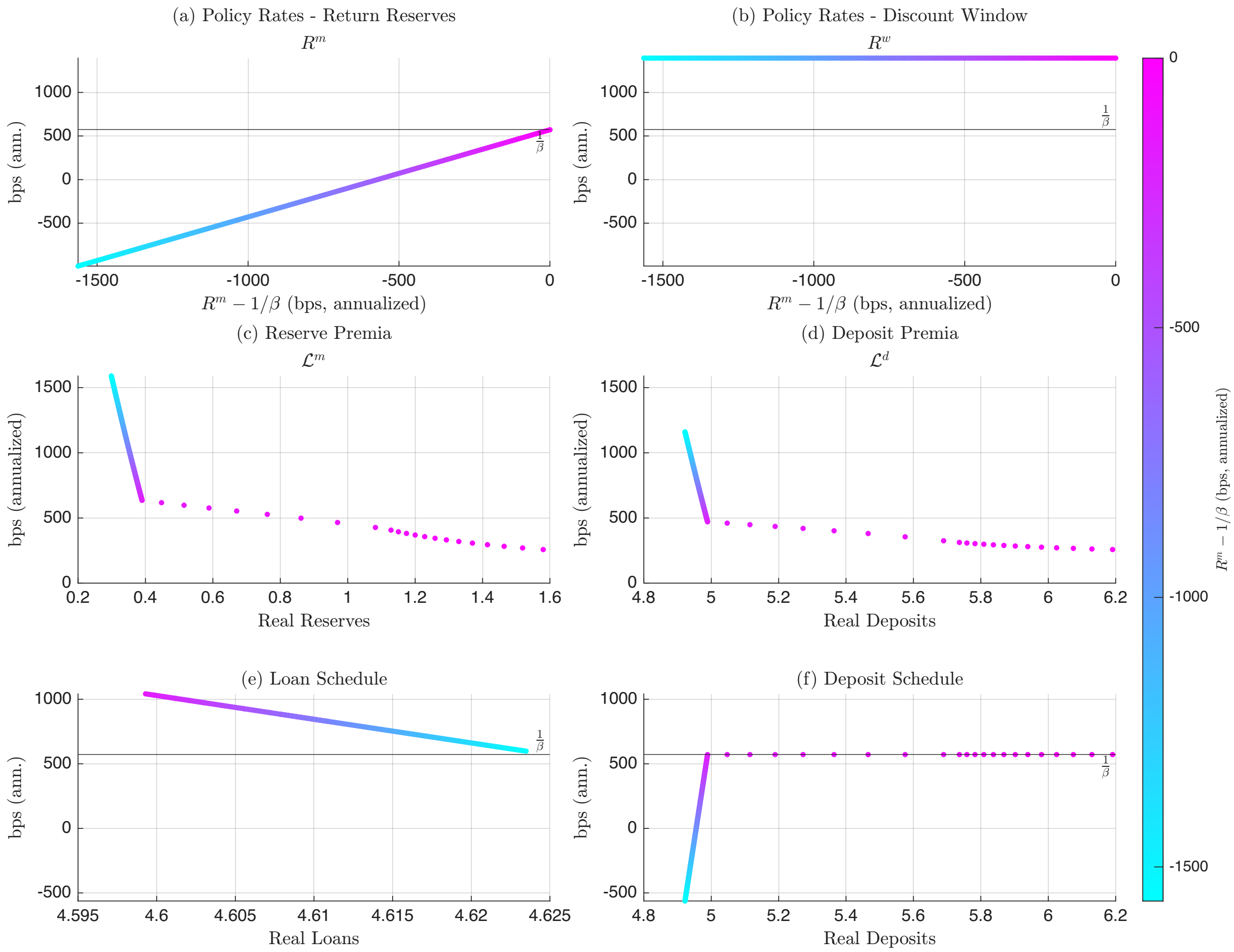

Figure 7: Policies such that $R^m \to \frac{1}{\beta}$ and R^w constant, where each "dot" represents the steady state under a different policy.

for reserves flattens, that loans start increasing again. Even at the Friedman rule, the bank does not hold the first best quantity of loans and so the welfare cost does not go to zero. In this example, the welfare cost in fig. 8 is still monotonically decreasing as the Friedman rule is implemented since bank's increase their intermediation of deposits. Here the relevant comparison the marginal value of a loan vs the marginal value of deposit which is determined by the form of the utility function $U(c, d)$.

The point where reserves start crowding in loans rather than crowding out is where households are satiated in their deposit demand. To demonstrate this, we plot in fig. 9 an equilibrium diagram of the values of (μ, l) that satisfy the first order conditions of the bank.

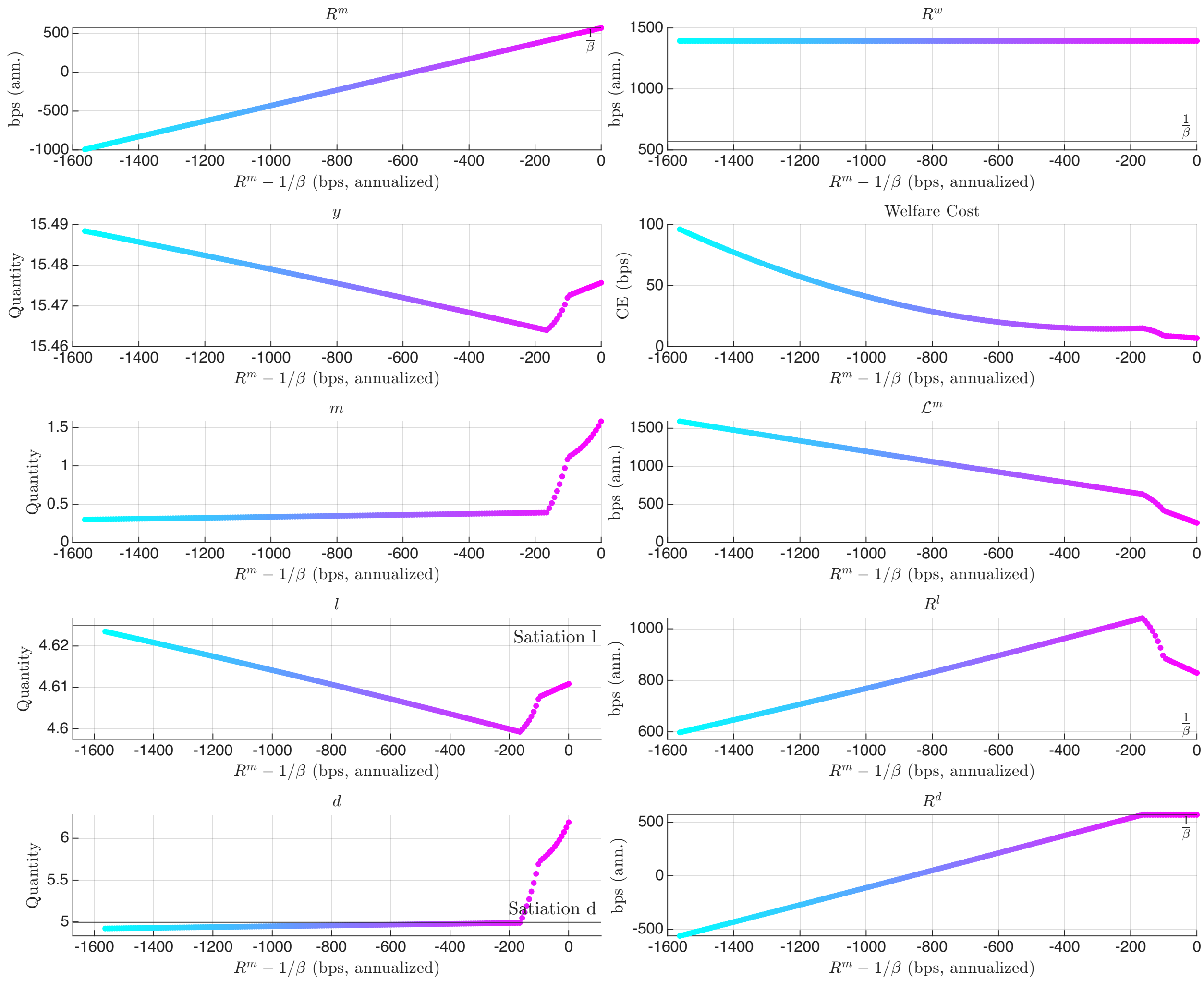

Figure 8: Policies such that $R^m \to \frac{1}{\beta}$ and R^w constant, where each "dot" represents the steady state under a different policy.

That is pairs (μ, l)

$$R^l(l) = R^m + \mathcal{L}^m(\mu, R^m, R^w) \tag{7}$$

$$R^d(l/(1-\mu)) = R^m + \mathcal{L}^d(\mu, R^m, R^w). \tag{8}$$

that satisfy each equation individually, where $R^l(l), R^d(d)$ are the inverse demand functions. The intersection (μ^*, l^*) would represent an equilibrium. The policies in panel (c-d) Figure 9 represent when the demand for deposits is first satiated in the response function and corresponds to the kink in the equilibrium diagram for eq. (8). As we move from policies in panel

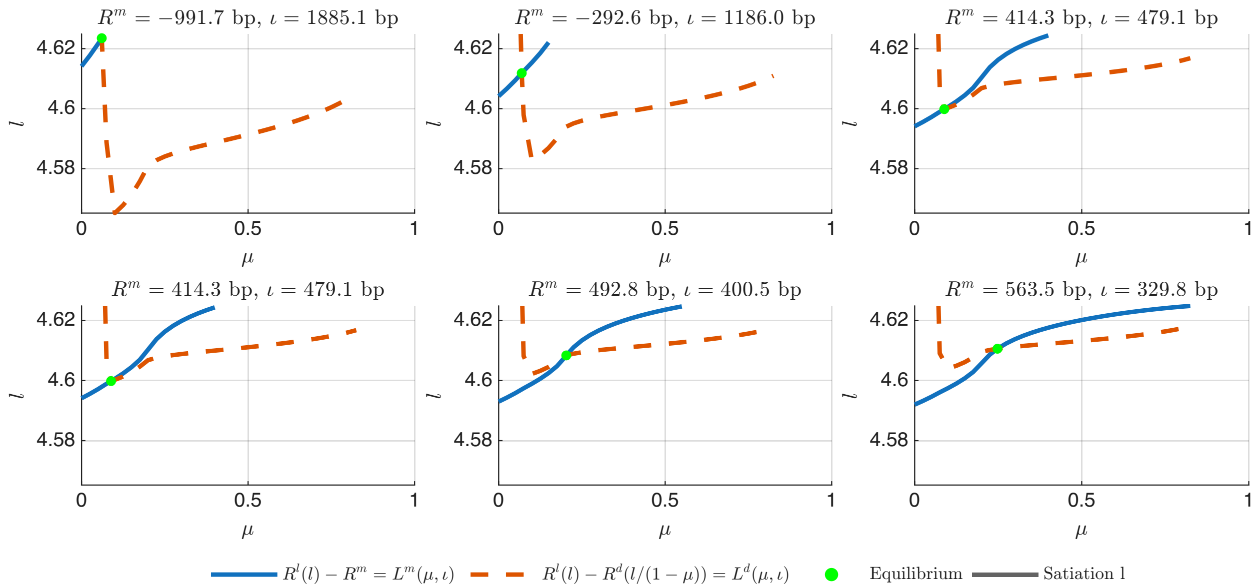

Figure 9: Values of (μ, l) that satisfy eqs. (7) and (8). Panel (c, d) represent the steady state where deposits are first satiated.

(a) to panel (c), the amount of loans is decreasing as the reserve ratio increases, indicating the portion of the previous diagram where loans are crowded out by reserves. As we move from panel (d) to panel (f), the amount of loans is increasing again, which correspond to the region where reserves crowd in loans.

Finally we show now a scenario where R^m, R^w move in tandem so that $\iota = R^w - R^m$ is constant which demonstrates that if the cost of holding reserves is too high, bank's will hold none. If R^m goes small enough, bank's find it worthwhile to get all of the reserves from the discount window. Due to the fixed level of $R^w - R^m$, when $R^m < \frac{1}{\beta}$ we can think that the CB is subsidizing the discount window, keeping it's use attractive even when R^m is very low.

In fact, fig. 11 shows that when reserves take a very small to no space on the balance sheet, banks are able to intermediate more loans than the economy needs, i.e. above it's satiation level. The output response function is flat as R^m increases initially, only decreasing once loans are below their satiation point - not when the economy transitions from a non-monetary equilibrium to a monetary equilibrium. This explains the extra kinks that output and the welfare cost exhibit.

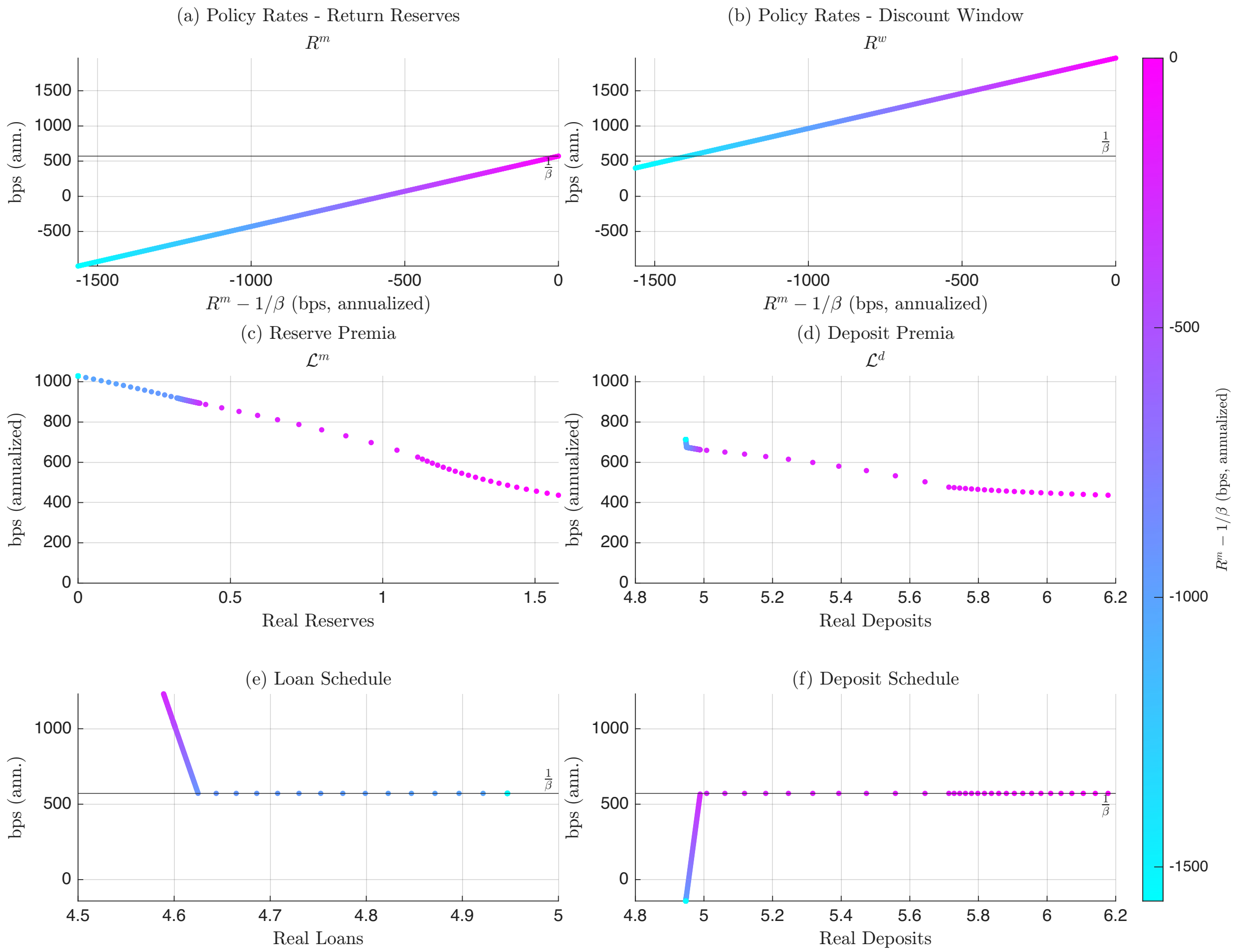

Figure 10: Policies such that $R^m \to \frac{1}{\beta}$ and $R^w - R^m$ constant, where each "dot" represents the steady state under a different policy.

5 Conclusion and Next Steps

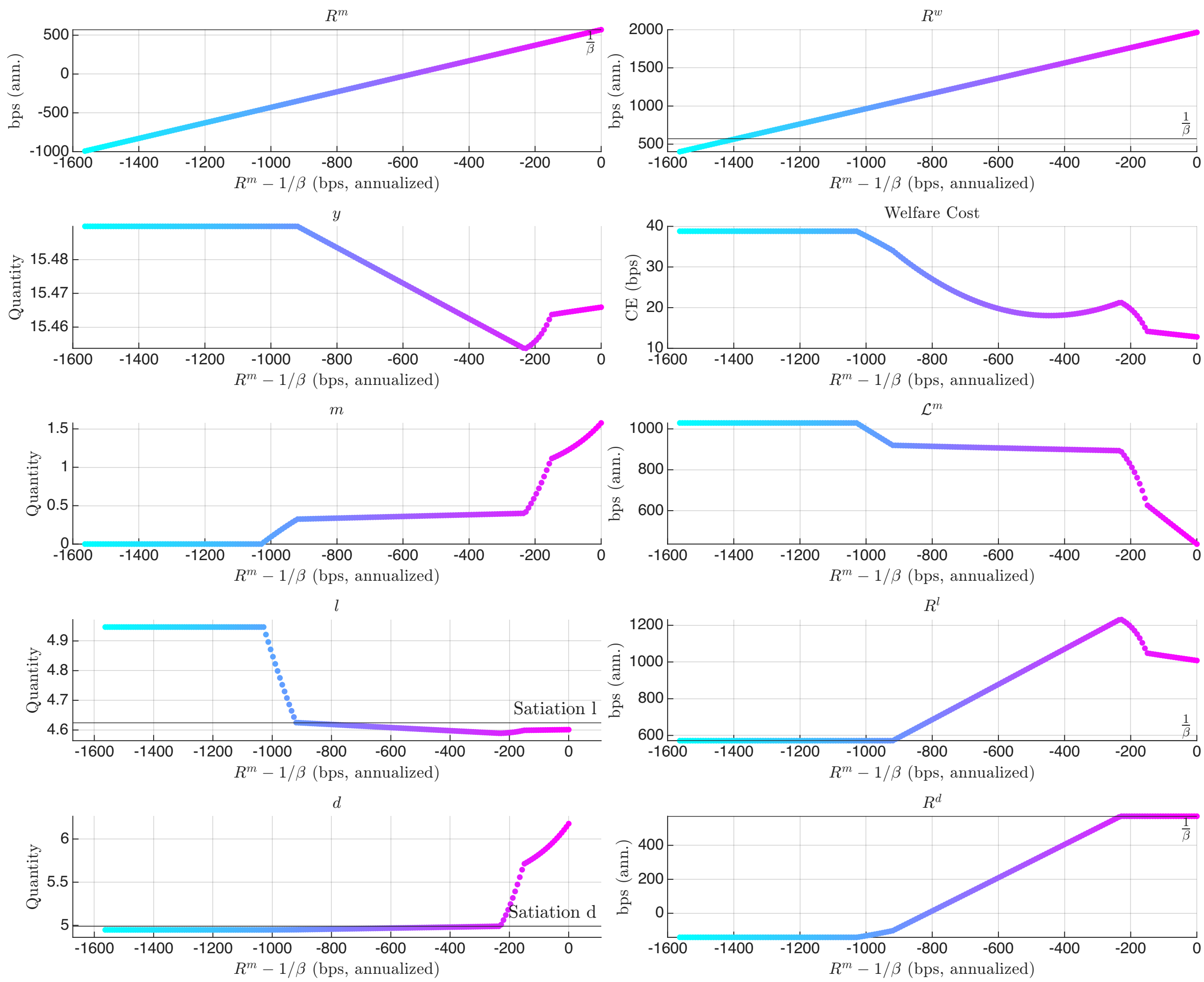

Figure 11: Policies such that $R^m \to \frac{1}{\beta}$ and $R^w - R^m$ constant, where each "dot" represents the steady state under a different policy.

References

Afonso, Gara, Domenico Giannone, Gabriele La Spada, and John C. Williams. 2022. *Scarce, Abundant, or Ample? A Time-Varying Model of the Reserve Demand Curve.* Working Paper 1019. Staff Reports.

Alvarez, Fernando, Francesco Lippi, and Roberto Robatto. 2019. “Cost of Inflation in Inventory Theoretical Models.” *Review of Economic Dynamics* 32 (April 1, 2019): 206–226.

Atkeson, Andrew G., Andrea L. Eisfeldt, and Pierre-Olivier Weill. 2015. “Entry and Exit in OTC Derivatives Markets.” *Econometrica* 83 (6): 2231–2292.

Bailey, Martin J. 1956. “The Welfare Cost of Inflationary Finance.” *Journal of Political Economy* 64 (2): 93–110. JSTOR: 1826826.

Begenau, Juliane. 2020. “Capital Requirements, Risk Choice, and Liquidity Provision in a Business-Cycle Model.” *Journal of Financial Economics* 136, no. 2 (May 1, 2020): 355–378.

Belongia, Michael T., and Peter N. Ireland. 2019. “The Demand for Divisia Money: Theory and Evidence.” *Journal of Macroeconomics* 61 (September 1, 2019): 103128.

Benati, Luca, and Juan Pablo Nicolini. 2021. *The Welfare Costs of Inflation.* 783. Federal Reserve Bank of Minneapolis, September 24, 2021.

Bianchi, Javier, and Saki Bigio. 2022. “Banks, Liquidity Management, and Monetary Policy.” *Econometrica* 90 (1): 391–454.

———. 2023. “Settlement Frictions and Portfolio Theory.” Pre-published, accessed April 10, 2024. http://www.sakibigio.net/s/Draftv8JETSubmission6.pdf.

Cooley, Thomas F., and Gary D. Hansen. 1989. "The Inflation Tax in a Real Business Cycle Model." *The American Economic Review,* 733–748.

———. 1991. "The Welfare Costs of Moderate Inflations." *Journal of Money, Credit and Banking* 23, no. 3 (August): 483. JSTOR: 1992683.

Craig, Ben, and Guillaume Rocheteau. 2008. "Inflation and Welfare: A Search Approach." *Journal of Money, Credit and Banking* 40 (1): 89–119.

D'Avernas, Adrien, Antoine Hubert De Fraisse, Liming Ning, and Quentin Vandeweyer. 2025. "The Fiscal Cost of Quantitative Easing." Pre-published, accessed July 18, 2025. https://doi.org/10.2139/ssrn.5314990. https://www.ssrn.com/abstract=5314990.

Dávila, Eduardo, Daniel Graves, and Cecilia Parlatore. 2024. "The Value of Arbitrage." *Journal of Political Economy* 132, no. 6 (June 1, 2024): 1947–1993.

Drechsler, Itamar, Alexi Savov, and Philipp Schnabl. 2017. "The Deposits Channel of Monetary Policy*." *The Quarterly Journal of Economics* 132, no. 4 (November 1, 2017): 1819–1876.

Freixas, Xavier, and Jean-Charles Rochet. 2008. *Microeconomics of Banking, Second Edition.* MIT Press, March 14, 2008. Google Books: vdrxCwAAQBAJ.

Gertler, Mark, and Nobuhiro Kiyotaki. 2010. "Chapter 11 - Financial Intermediation and Credit Policy in Business Cycle Analysis." In *Handbook of Monetary Economics,* edited by Benjamin M. Friedman and Michael Woodford, 3:547–599. Elsevier, January 1, 2010.

Kurlat, Pablo. 2019. "Deposit Spreads and the Welfare Cost of Inflation." *Journal of Monetary Economics,* SPECIAL CONFERENCE ISSUE: "Money Creation and Currency Competition" October 19-20, 2018 Sponsored by the Study Center Gerzensee and Swiss National Bank, 106 (October 1, 2019): 78–93.

Lagos, Ricardo, and Gastón Navarro. 2023. *Monetary Policy Operations: Theory, Evidence, and Tools for Quantitative Analysis.* National Bureau of Economic Research.

Lagos, Ricardo, and Randall Wright. 2005. "A Unified Framework for Monetary Theory and Policy Analysis." *Journal of Political Economy* 113, no. 3 (June): 463–484.

Levine, Jeffrey, and Asani Sarkar. 2019. "How Do Large Banks Manage Their Cash?" Liberty Street Economics, July 17, 2019, 7:00 a.m. (Z). Accessed May 6, 2025. https://libertystreeteconomics.newyorkfed.org/2019/07/how-do-large-banks-manage-their-cash/.

Li, Ye, Yi Li, and Huijun Sun. 2024. *The Network Structure of Money Multiplier.* Working paper, Columbia University, Federal Reserve Board, and University of ...

Lopez-Salido, David, and Annette Vissing-Jorgensen. 2023. "Reserve Demand, Interest Rate Control, and Quantitative Tightening." Preprint, February 27, 2023. SSRN Scholarly Paper. Accessed May 17, 2023. https://doi.org/10.2139/ssrn.4371999. https://papers.ssrn.com/abstract=4371999.

Lucas, Robert E. 2000. "Inflation and Welfare." *Econometrica* 68 (2): 247–274. JSTOR: 2999427.

Lucas, Robert E., and Juan Pablo Nicolini. 2015. "On the Stability of Money Demand." *Journal of Monetary Economics,* Carnegie-Rochester-NYU Conference Series on Public Policy "Monetary Policy: An Unprecedented Predicament" Held at the Tepper School of Business, Carnegie Mellon University, November 14-15, 2014, 73 (July 1, 2015): 48–65.

Payne, Jonathan, and Bálint Sz\ Hoke. 2025. "The Fragility of Government Funding Advantage" (July 17, 2025).

Serletis, Apostolos, and Libo Xu. 2021. "The Welfare Cost of Inflation." *Journal of Economic Dynamics and Control* 128 (July 1, 2021): 104144.

———. 2025. "The Welfare Cost of Inflation." Pre-published, April 30, 2025.

Stulz, René M., Alvaro G. Taboada, and Mathijs A. van Dijk. 2022. "Bank Liquid Assets, the Portfolio Motive, and Capital Requirements." Pre-published, August. Working Paper. Accessed May 7, 2025. https://doi.org/10.3386/w30340. National Bureau of Economic Research: 30340. https://www.nber.org/papers/w30340.

Tobin, James. 1963. "Commercial Banks as Creators of 'money."

Van den Heuvel, Skander J. 2008. "The Welfare Cost of Bank Capital Requirements." *Journal of Monetary Economics* 55, no. 2 (March 1, 2008): 298–320.

Van Den Heuvel, Skander J. 2022. "The Welfare Effects of Bank Liquidity and Capital Requirements." *Finance and Economics Discussion Series,* nos. 2022–072 (November): 1–65.

Vissing-Jorgensen, Annette. 2023. "Convenience Yields and Monetary Policy." Sixteenth Paolo Baffi Lecture on Money and Finance. Banca D'Italia, November.

Wang, Olivier. 2025. "Banks, Low Interest Rates, and Monetary Policy Transmission." *The Journal of Finance* 80 (3): 1379–1416.

Wen, Yi. 2015. "Money, Liquidity and Welfare." *European Economic Review* 76 (May 1, 2015): 1–24.

Williamson, Stephen. 2023. "Interest Rate Control and Interbank Markets" (July 27, 2023).

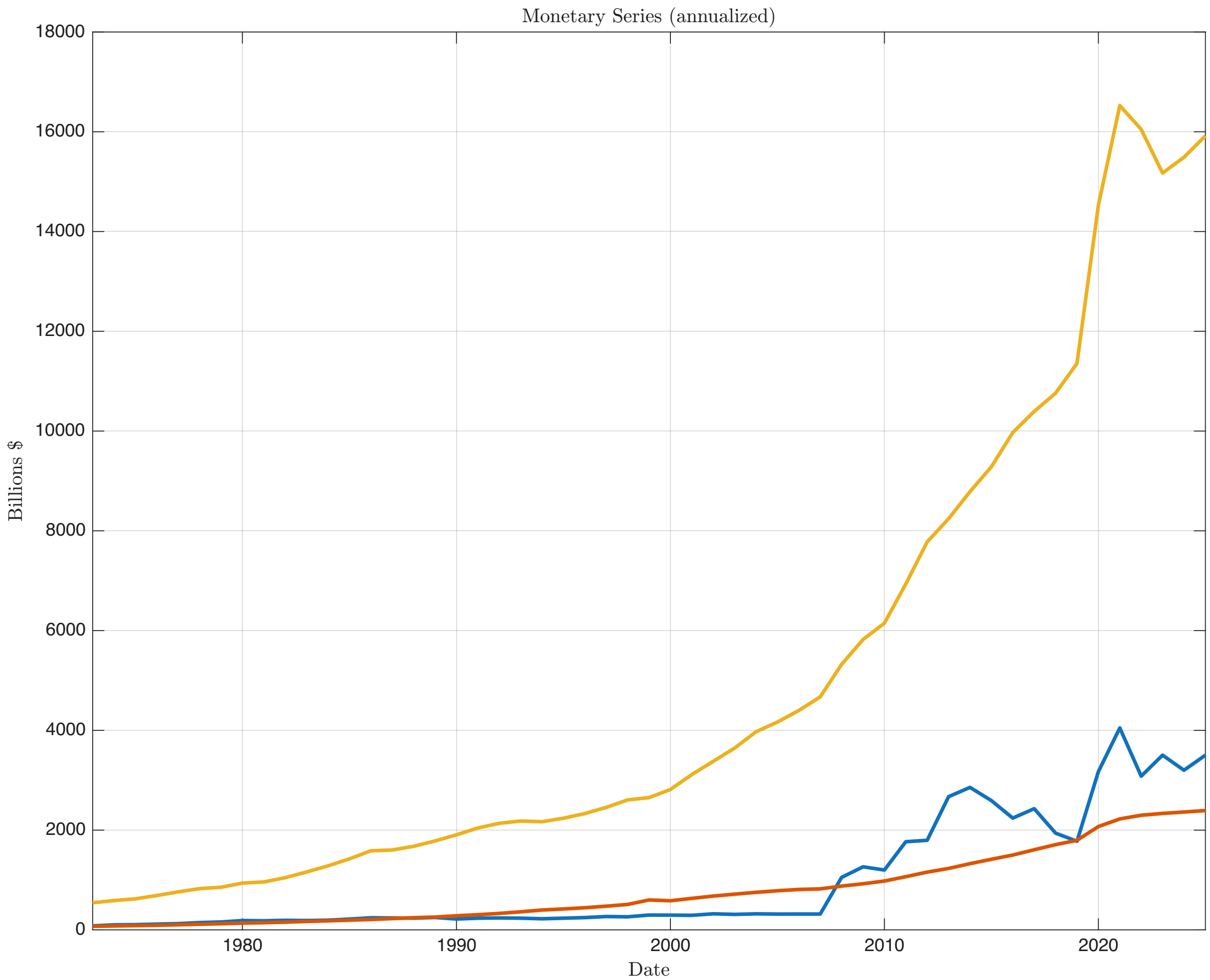

Figure 12: Deposit Growth

A Other Figures

B Proofs

We detail proofs for propositions in the main text.

B.1 Expanded Proof of lemma 1

Here we show the proof of lemma 1 for the case of slow settlement($\upsilon = 0$). The proof for fast settlement ($\upsilon = \infty$) is contained in the main text.

Lemma (1). *If banker's portfolio decisions are symmetric, then tightness in the interbank market depend only on the ratio of aggregate reserves to aggregate deposits,* $M/D = \mu$.

Proof. This does not change the formula for S^- since $s^j < 0$ implies that $\omega^j d^j + \xi^j D - \varrho\{(1-\omega^j)d^j + \xi^j D\} < 0$.

Again focus on a symmetric equilibria. Note that

$$-\varrho D - (1-\varrho)\omega D + (1-\varrho)\xi D < 0 \iff \xi - \omega < \frac{\varrho}{1-\varrho}.$$

Denote as $s^{**} \equiv \frac{\varrho}{1-\varrho}$. We will have

$$\begin{aligned} S^+ &= \int_{s^{**}}^{\infty} M \, \mathrm{d}\Phi_z(z) + \int_{s^*}^{s^{**}} M - \varrho D - (1-\varrho)\omega D + (1-\varrho)\xi D \, \mathrm{d}\Phi_z(z) \\ &= (1-\varrho)\, D \left[\frac{\mu}{1-\varrho} \left(1 - \Phi_z(s^{**}) - s^* \left(\Phi_z(s^{**}) - \Phi_z(s^*)\right) + \left(f_z(s^{**}) - f(s^*)\right)\right) \right]. \end{aligned}$$

The result follows from noting that s^*, s^{**} only depend on μ. □

C The Over-the-Counter Market for Reserves

(to be filled in)