Finance and Economics Discussion Series

Federal Reserve Board, Washington, D.C.
ISSN 1936-2854 (Print)
ISSN 2767-3898 (Online)

Beyond Reserves: The Federal Reserve's Balance Sheet and the Repo Market

Sriya Anbil, Alyssa Anderson, Ethan Cohen, Romina Ruprecht

2026-041

Please cite this paper as:
Anbil, Sriya, Alyssa Anderson, Ethan Cohen, and Romina Ruprecht (2026). "Beyond Reserves: The Federal Reserve's Balance Sheet and the Repo Market," Finance and Economics Discussion Series 2026-041. Washington: Board of Governors of the Federal Reserve System, https://doi.org/10.17016/FEDS.2026.041.

Beyond Reserves: The Federal Reserve's Balance Sheet and the Repo Market*

Sriya Anbil†[1], Alyssa Anderson[1], Ethan Cohen[2,3], and Romina Ruprecht[1]

[1]Board of Governors of the Federal Reserve System

[2]Federal Reserve Bank of Minneapolis

[3]University of Minnesota

Abstract

We present a new constraint on the size of the Fed's balance sheet: repo market capacity. Calibrating a structural model to the recent monetary tightening cycle, we show that repo market capacity—driven by money market fund liquidity supply—is the binding constraint on the Fed's balance sheet, not bank reserve demand, which was highlighted in the events of September 2019. We also demonstrate a novel complementarity between interest rate and balance sheet policies: higher policy rates expand repo capacity, allowing the central bank to operate with a smaller balance sheet.

Keywords: monetary policy implementation, quantitative tightening, reserves, overnight reverse repo facility, shadow banks

*The views expressed here are solely those of the authors and should not be interpreted as reflecting the views of the Board of Governors, its staff, or anyone associated with the Federal Reserve System. We are thankful for comments from Viral Acharya, Gara Afonso, Paola Boel [discussant], David Bowman, William Diamond [discussant], Wenxin Du, Thomas Eisenbach [discussant], Pedro Gomis-Porqueras, Christopher Gust, Kinda Hachem, Sylverie Herbert, Sebastian Infante, Ivan Ivanov, Benjamin Lester, Wenhao Li [discussant], Yiming Ma [discussant], Antoine Martin, Fernando Martin, Ralf Meisenzahl, Borghan Narajabad, Rafael Repullo [discussant], Bernd Schlusche, Zeynep Senyuz, Annette Vissing-Jorgensen, Ansgar Walther [discussant], Kairong Xiao [discussant], and seminar participants at the Summer Workshop of Money, Banking, Payments & Finance, the Midwest Macro Conference, the Swiss Winter Conference on Financial Intermediation, the SFS Cavalcade, the Women in System Economic Research Conference, WAPFIN, the EFA, the AFA, the Day-Ahead Conference on Financial Markets and Institutions, the ECB Money Markets Conference, the West Coast Search and Matching Workshop, the SNB Research Conference, the System Research Conference for Financial Intermediation, the Bank of England BEAR Conference, the Bundesbank, the Federal Reserve Banks of New York, Chicago and St. Louis, and the Federal Reserve Board. Ellie Newman provided excellent research assistance. All remaining errors are our own.

†Corresponding Author: 20th Street Constitution Ave. NW, Washington, DC 20551; E-mail: sriya.l.anbil@frb.gov

1 Introduction

The optimal size of the Federal Reserve's (Fed's) balance sheet needed to "implement monetary policy efficiently and effectively" in an ample reserves framework is highly uncertain.[1] Much of the literature views bank reserve demand as the primary determinant of this optimal size (Yang, 2026; Afonso et al., 2023b; Afonso et al., 2026; Acharya et al., 2023; Lagos and Navarro, 2023; and Lopez-Salido and Vissing-Jorgensen, 2025, among others). In this paper, we introduce a new constraint on the Fed's balance sheet—the supply of money held by non-banks, in particular money market funds (MMFs)—and show that this constraint can bind more tightly than bank reserve demand. Despite the central role of non-bank money supply in modern monetary policy implementation, its implications remain poorly understood. This paper aims to fill that gap.

We jointly model the demand for reserves by banks and the supply of money by non-banks and present two key results. First, we show that the supply of money held by non-banks, rather than reserve demand by banks, constrains the size of the Federal Reserve's balance sheet. Our results imply that if the Fed were to focus solely on bank reserve demand and ignore the role of non-banks in determining short-term liquidity conditions, it could lose control of the effective federal funds rate (EFFR). Such a loss of control would run counter to the FOMC's stated framework for policy implementation, potentially undermining central bank credibility and increasing volatility in funding markets (Gorton and Metrick, 2012; FOMC, 2019; He et al., 2022).

Second, we demonstrate a novel complementarity between the Fed's interest rate and balance sheet policy tools. In particular, the Fed can operate with a smaller balance sheet when interest rates are higher. This result implies that quantitative tightening (QT) can proceed further when the policy rate is elevated, all else equal.

We develop a tractable structural model to study how deposits flow between banks and non-banks as the Fed tightens monetary policy, either by raising its administered rates—the interest rate on reserve balances (IORB) and the interest rate on the Overnight Reverse Repurchase Facility (ON RRP)—or by reducing the size of its balance sheet. The model is designed to capture the financial plumbing that reallocates money across the U.S. financial system, which is essential for quantifying the effects of QT on money demand. We then calibrate the model using data from the recent monetary tightening cycle to

[1] See the January 2022 Principles for Reducing the Size of the Federal Reserve's Balance Sheet.

jointly determine bank reserve demand, the supply of money held by non-banks, and the size of the Fed's balance sheet that is consistent with an ample reserves framework.

During QT, the Fed reduces the size of its balance sheet by not replacing maturing Treasury securities.[2] This process has two key effects. First, the U.S. Treasury must issue new securities to replace those previously held by the Fed, which are initially bought by broker-dealers (dealers) and other investors in the primary market. These securities are often financed in the repurchase (repo) market. Second, as the Fed's asset holdings decline, its liabilities—including bank reserves—must also decline, mechanically reducing the supply of money in the financial system.

As a result, QT simultaneously increases the demand for repo borrowing while reducing the supply of liquid funding. This combination places upward pressure on repo rates. Our model provides a unified framework to capture both channels and quantify their implications for repo market conditions and the minimum balance sheet consistent with ample reserves.[3]

We show that the supply of money held by non-banks determines the capacity of the repo market, which in turn determines how much the Fed can shrink its balance sheet while maintaining control over short-term interest rates. In this framework, repo market capacity determines the quantity of securities that the Fed must hold. Because the Fed's assets and liabilities must balance, this quantity maps directly into the level of reserves.

In contrast, prior literature emphasizes that declines in aggregate reserves reduce bank liquidity and incentivize banks to pay higher rates for funding. As a result, reserve scarcity is typically viewed as the constraint on the size of the Fed's balance sheet. We show, however, that repo market capacity implies a larger minimum balance sheet than bank reserve demand. While conditions in the repo market do not affect banks' *demand* for reserves, they determine the *level* of reserves through the Fed's balance sheet identity.

Our model builds on the structural frameworks of Armenter and Lester (2017) and Xiao (2020), and is closely related to d'Avernas et al. (2025), while incorporating a richer institutional structure to reflect the functioning of U.S. money markets. In the model, Treasury securities that roll off the Fed's balance sheet are purchased in the primary market by dealers and other levered investors (such as hedge funds), who finance these

[2] See the May 2022 Plans for Reducing the Size of the Federal Reserve's Balance Sheet.

[3] The Fed currently uses an "ample reserves" monetary policy framework. See Afonso et al. (2023a) for a discussion of the ample reserves framework and how the Fed evolved to this framework.

positions in the repo market. MMFs allocate their portfolios across repo lending, Treasury bills, and the ON RRP facility, while banks allocate deposits between reserves and loans.

We show that when the Fed tightens monetary policy by raising interest rates, the model is consistent with the existing literature on both the bank deposit channel and the non-bank deposit channel. As the policy rate rises, deposits flow out of the banking system via the bank deposit channel (Bernanke and Blinder, 1988; Kashyap and Stein, 1995; Drechsler et al., 2017; Wang et al., 2022). These deposits are reallocated to MMFs through the non-bank deposit channel (Xiao, 2020; d'Avernas and Vandeweyer, 2024). Consistent with this evidence, households in our model shift deposits from banks to MMFs as the policy rate increases, since MMF yields adjust more fully to changes in IORB while bank deposit rates adjust more slowly.

MMFs then invest these additional deposits in the ON RRP facility rather than in the repo market or Treasury bills. This reflects the fact that a higher policy rate does not increase dealers' demand for repo financing or the supply of Treasury bills, limiting alternative investment opportunities. As a result, in these "abundant liquidity" equilibria, the repo rate equals the ON RRP rate, repo volumes remain largely unchanged, and ON RRP balances increase. These balances represent excess liquidity that can be redeployed into repo markets as needed. This implication is consistent with the empirical evidence in Afonso et al. (2022), who document that increases in MMF deposits during the 2022 tightening cycle were associated with substantial increases in ON RRP investment.

However, when the Fed tightens monetary policy by reducing the size of its balance sheet, which we model as dealers and investors absorbing an exogenous amount of securities that the Fed no longer purchases, dealers' demand for repo borrowing increases to finance these positions. When liquidity is abundant, MMFs reallocate funds from the ON RRP to the repo market, while repo rates remain anchored at the ON RRP rate.

As QT continues and demand for repo borrowing rises relative to the supply of funds that MMFs can provide, liquidity becomes less abundant. Once ON RRP take-up reaches zero, repo rates increase above the ON RRP rate. Higher repo rates allow MMFs to offer more attractive yields, leading to inflows of deposits from households, as MMF yields rise above bank deposit rates. In these "ample liquidity" equilibria, MMFs no longer invest at the ON RRP and instead allocate funds between repo lending and Treasury bills.

These dynamics highlight that reductions in the Fed's balance sheet shift liquidity from the ON RRP into the repo market, and that the capacity of MMFs to absorb dealers' repo financing needs is central to the transmission of QT. Importantly, MMFs expand repo lending primarily in response to balance sheet reduction, rather than to increases in the policy rate. Consistent with this mechanism, the ON RRP plays a key role in the Fed's current monetary policy implementation framework: ON RRP balances being zero is a necessary condition for the minimal size of the Fed's balance sheet, although it is not necessarily sufficient.

To quantify these mechanisms, we calibrate the structural parameters of the model to match key moments from the last monetary tightening cycle. Our sample begins on June 1, 2022, when the Federal Reserve initiated balance sheet run-off, and ends on August 31, 2024. While balance sheet reduction continued through December 2025, ending the sample earlier allows us to conduct out-of-sample tests in late 2024 to validate the model's fit.

The model matches the data well. During our calibration period, we estimate that if the Fed aimed to keep repo rates at the IORB rate, reserves could be reduced to approximately \$2.7 trillion. If reserves were reduced much below this level, interest rate control could be compromised, as repo rates would rise well above the IORB rate and the financial system would move out of an ample reserves regime. Our equilibrium estimate of ample reserves is consistent with Lopez-Salido and Vissing-Jorgensen (2025), who estimate that the Fed can reduce its balance sheet until reserves reach approximately \$2.76 trillion while maintaining the effective federal funds rate (EFFR) at the IORB rate.

Our calibration delivers two main results. First, the level of reserves consistent with an ample reserves framework implied by bank reserve demand is lower than that implied by non-bank money supply and the corresponding capacity of the repo market. If the Fed were to shrink its balance sheet to the level implied solely by bank reserve demand, our model predicts that repo rates would rise above the target range for the federal funds rate, compromising interest rate control and moving the system out of an ample reserves regime.

As an example of this, the model provides a mechanism that helps explain the spike in repo rates on September 16–17, 2019, following the Fed's first QT episode after the Global Financial Crisis. While Correa et al. (2025) and Copeland et al. (2025) emphasize the role of low aggregate reserves, our model shows that the key constraint operated through the repo market. As reserves declined and Treasury supply increased, dealers required additional repo financing, but the repo market lacked sufficient capacity to absorb this demand. In this sense, low reserve levels contributed to the spike in repo rates by tightening funding conditions, but were not, on their own, sufficient to generate the observed increase in rates.

Developments in 2025 have further reinforced the relevance of the mechanism highlighted in our model. As QT progressed in late 2025, repo markets became more constrained and repo rates increased, prompting the Fed to halt balance sheet reduction and resume reserve management purchases. As discussed in the December 2025 FOMC minutes, repo rates remained elevated and volatile, reflecting tighter funding conditions and continued Treasury issuance.[4] These developments are consistent with our central implication that policymakers should incorporate non-bank money supply into their assessment of short-term liability provision. Looking ahead, there has been growing discussion of possible ways to further reduce the size of the Fed's balance sheet.[5] Our mechanism is central to this discussion and to ensuring that financial markets can absorb the securities that the Fed no longer holds.

Our second main result highlights a novel complementarity between interest rate policy and balance sheet policy. While these tools are often viewed as substitutes in determining the stance of monetary policy, our results show that they can act as complements in its implementation. In particular, a higher policy rate allows the Fed to operate with a smaller balance sheet. As interest rates rise, households reallocate deposits from banks to MMFs, increasing MMF assets under management and, in turn, their capacity to supply funding in repo markets. This expanded repo market capacity allows the financial system to absorb a larger supply of Treasury securities as the Fed reduces its holdings.

This complementarity has important policy implications. If the Fed prefers to operate with a smaller balance sheet, our results suggest that raising the policy rate first would facilitate a smoother reduction in balance sheet size, all else equal. At the same time, the model is agnostic about the appropriate stance of monetary policy.

[4] As noted in the December 2025 FOMC minutes, "rates on Treasury repo remained relatively elevated and volatile... attributed to a decline in available liquidity and continued large Treasury debt issuance," and "higher repo rates... contributed to upward pressure on the spread between the federal funds rate and the interest rate on reserve balances."

[5] See, for example, Anderson et al. (2026), Duffie (2026), and Logan and Schulhofer-Wohl (2026).

Related Literature

Our paper is related to the growing literature on the transmission of quantitative tightening (QT) and quantitative easing (QE). The effects of QT remain relatively less well understood, in part because QT has only been implemented twice in the modern U.S. monetary policy era since 2008.[6] While a large literature studies balance sheet policy effects on asset prices (Krishnamurthy and Vissing-Jorgensen, 2011, Bauer and Rudebusch, 2014, Bhattarai et al., 2022, among others), the academic literature on QT implications remains limited. Diamond et al. (2024) study the effects of QE on bank reserves and show that reserves crowd out bank lending and, to a lesser extent, crowd in deposits. In contrast, we find that under QT, banks lose deposits and reduce lending, highlighting an asymmetry between balance sheet expansions and contractions. This lack of symmetry between QE and QT has also been emphasized in D'Amico and King (2013) and Smith and Valcarcel (2023). A key contribution of our framework is that it can be used to study how balance sheet policy affects the allocation of money between banks and non-banks. The macroeconomic effects of QT have been studied in Arazi (2025) and Kumhof and Salgado-Moreno (2024). While our model abstracts from long-term interest rates and term premia, it provides a framework to analyze the capacity of repo markets and the composition of short-term liabilities—reserves and ON RRP—under both balance sheet expansion and contraction.

The transmission of balance sheet policy depends critically on how Treasury securities are intermediated and financed in the repo market. A key distinction emphasized in this literature is between the primary market (where Treasury securities are initially absorbed) and the secondary market (where end investors purchase them). While cross-sectional holdings suggest that dealers hold only a small share of Treasuries, recent work shows that dealers and levered investors play a central role in absorbing Treasury issuance and financing these positions in repo markets before securities are redistributed to end investors (e.g., Fleming et al., 2024; Boyarchenko et al., 2021; Anbil et al., 2026b). We build on this by explicitly modeling settlement-stage intermediation of Treasury issuance and its implications for repo funding demand. This distinction between short-lived intermediation and long-run holdings is central to the mechanism we study.

This intermediation mechanism depends on MMFs as key repo funding providers.

[6] Small Treasury security sales were conducted in the 1950s to combat inflation (Carlson and Wheelock, 2014).

MMFs allocate across repo, Treasury bills, and ON RRP based on relative returns (Afonso et al., 2022; Anbil et al., 2021; Mojir and Anbil, 2025; Abdullah and Tase, 2025; Im et al., 2023). We incorporate this substitution margin explicitly, while distinguishing between bill markets and repo-financed intermediation of longer-term securities.

Our work also relates to the literature on the interaction between interest rate policy and balance sheet policy (Stein, 2012; Yellen, 2016; Carlson and Duygan-Bump, 2021). These tools are typically viewed as substitutes in determining the stance of monetary policy, and a large literature quantifies their relative effects through term premia (e.g., Sims and Wu, 2020, Crawley et al., 2022, among others). Interest rate policy affects longer-term rates through expectations of the path of short-term rates, while balance sheet policy operates primarily through term premia. Kiley (2014) suggests that expectations of short-term rates may be the dominant channel, implying a greater role for interest rate policy.[7] In contrast, we show that these tools are complementary in implementation: the effectiveness of balance sheet policy depends on the level of the policy rate. Higher policy rates increase repo market capacity, allowing a smaller central bank balance sheet, with implications for the size of the non-bank financial sector and, more broadly, for financial stability.

Finally, our structural model contributes to a growing literature that studies monetary policy implementation using micro-founded models of the banking system. Our framework is most closely related to d'Avernas et al. (2025), and builds on elements of Armenter and Lester (2017) and Xiao (2020). Recent work uses structural models to study bank market power (Wang et al., 2022; Li et al., 2026), bank lending (Diamond et al., 2024), demand for government bonds (Koijen et al., 2021), and liquidity premia (Krishnamurthy and Li, 2023). Our contribution is to incorporate the interaction between Treasury issuance, repo markets, and non-bank financial intermediaries into a unified framework for monetary policy implementation.

The paper proceeds as follows. Section 2 provides background information on the Fed's ample reserves framework and presents empirical observations that motivate our model. Section 3 describes the model. Section 4 presents the calibration, with subsections

[7]For an early application of a similar framework to study how balance sheet policies shape interactions between banks and non-banks across institutional settings, see Bluwstein et al. (2026). For an overview of quantitative tightening approaches across countries, see Du et al. (2024). More broadly, prior work has examined the role of policy tool uncertainty. Brainard (1967) shows that diversification across policy instruments may be optimal when their effectiveness is uncertain, while Williams (2013) argues that policy should rely more heavily on instruments with more predictable effects.

4.3, 4.4, and 4.5 discussing the results from the calibration, testing the model in-sample, and testing the model out-of-sample, respectively. Section 5 discusses the comparative statics of the model, Section 6 discusses the implications of certain model assumptions, and Section 7 concludes.

2 Institutional Background

2.1 The Fed's Ample Reserves Framework

Before the 2008 Global Financial Crisis (GFC), the Fed operated in a "scarce reserves" monetary policy implementation framework. The Fed would conduct small open market operations to adjust the supply of money held by banks, i.e. reserves, to target the price of reserves, i.e. the federal funds rate. Figure 1 illustrates the supply/demand curve for reserves. In a "scarce reserves framework", the amount of reserves (shown on the x-axis) is low enough to be on the downward-sloping portion of the demand curve.

During and after the GFC, the Fed considerably expanded its balance sheet by purchasing Treasury securities and agency mortgage-backed securities (MBS) through quantitative easing, which significantly increased the amount of reserves in the financial system. Since the amount of reserves was so large, the Fed was unable to use open market operations to target the federal funds rate. Reserves were on the flat portion of the demand curve so small changes in the amount of reserves no longer affected rates. Instead, the Fed started paying interest on its reserves—currently the IORB rate—in September 2008 to provide a floor to the federal funds rate. Banks can hold money (reserves) at the Fed and receive the IORB rate. Consequently, banks are not incentivized to lend at lower than the IORB rate because they can lend money to the Fed instead.

However, the IORB rate was not a sufficiently effective floor for the federal funds rate because non-banks that were ineligible to earn IORB would lend below the IORB rate (Anbil and Senyuz, 2022). This lending, in turn, put downward pressure on the federal funds rate, which typically traded below the IORB rate. As a result, the Fed introduced the ON RRP in September 2013 to provide a firmer floor to the federal funds rate. At the ON RRP, non-banks, including MMFs, can lend money to the Fed overnight and receive the ON RRP rate. While the ON RRP rate is set lower than the IORB rate to ensure an effective floor to the federal funds rate, the interest paid to MMFs at the ON RRP is economically identical to the interest paid to banks to hold reserves at the Fed, and can be thought of as the "interest rate for non-bank reserves."

In January 2019, the Fed officially adopted an "ample reserves framework" that uses administered rates (namely the IORB and ON RRP rates) to control the federal funds rate. In this framework, the maximum amount of reserves in the financial system considered "ample" is when reserves are near the flat portion of the demand curve shown in Figure 1 but not larger, denoted by the gray dashed line marked "$X". The vertical blue line marked as "Supply" indicates the amount of reserves, $2.9 trillion, in the financial system as of April 2026. At this point, reserves are considered within the ample range.

An important aspect of this new monetary policy implementation framework is that aggregate money provided by the Fed is not just reserves held by banks but also money invested by non-banks at the ON RRP. As the Fed's balance sheet expanded in the latest round of QE, banks were satiated with reserves and money instead flowed to non-banks, particularly MMFs, who invested it at the ON RRP. The ON RRP went from essentially zero in early 2021 to a peak of $2.55 trillion a day at the end of 2022. By the end of 2022, the Fed provided over $5 trillion of short-term liabilities to banks and non-banks.

2.2 Empirical Evidence from the 2022 to 2025 Tightening Cycle

In this section, we present empirical observations about the recent monetary tightening cycle to motivate the setup of the model and show that non-bank money supply is important.

In the ample reserves framework, the Fed can tighten monetary policy in two ways. It can raise its policy rate via its administered rates—the IORB and ON RRP rates—thereby raising the federal funds rate, or it can remove short-term liabilities from the financial system by reducing the size of its balance sheet through QT. The Fed's short-term liabilities, which include reserves and non-bank reserves at the ON RRP, are backed by the Treasury securities and agency MBS that the Fed purchased during QE. During QT, the Fed lets these securities mature and roll off its balance sheet, mechanically reducing the short-term liabilities it initially created, either from reserves or non-bank reserves. As more securities mature, the amount of short-term liabilities will decline further, until eventually an amount of reserves would be reached where the federal funds rate starts to respond to changes in reserves. In this case, the amount of reserves in the financial system is considered "ample" and is just past the flat portion of the demand curve shown in Figure 1. The vertical dashed line $X represents the minimum amount of reserves consistent with maintaining the ample reserves framework and therefore staying near the flat portion of the demand curve.

In March 2022, the Fed began raising the target range for the federal funds rate by raising its administered rates in order to combat inflation (FOMC, 2022a). Soon after, on June 1, 2022, the Fed began QT (FOMC, 2022b). Figure 2 displays Treasury securities, reserves, and non-bank reserves (investment at the ON RRP) on the Fed's balance sheet between January 1, 2022 and December 31, 2023. During this period, $800 billion of Treasury securities rolled off the balance sheet. While the total amount of short-term liabilities declined mechanically by the same amount, we observe that reserves were stable while non-bank reserves, that is the investment at the ON RRP, declined substantially.

As these securities roll off the Fed's balance sheet, they must be held by the private sector instead, which has direct implications for Treasury intermediation and repo markets. A key institutional feature of the Treasury market is the distinction between the primary and secondary markets, as well as the timing of intermediation. Newly issued Treasury securities are absorbed in the primary market by dealers and levered investors and are typically financed in repo markets at settlement. These positions are subsequently redistributed to end investors such as mutual funds, pension funds, and foreign institutions. As a result, cross-sectional holdings of Treasuries understate the importance of intermediary balance sheets at issuance. This short-lived but economically meaningful intermediation plays a central role in repo funding demand and is a key feature of the environment we model.

Figure 4 illustrates this mechanism by showing the share of newly-issued Treasury coupon securities financed in repo markets at settlement. A substantial fraction of issuance is repo-financed, highlighting the importance of intermediary balance sheets at the point of issuance. Moreover, this mechanism is primarily active for longer-dated Treasury securities, as bills are financed in repo markets only to a small extent (Figure 5).

This intermediation mechanism has direct implications for QT. As the Fed allows its longer-dated Treasury securities holdings to mature, the U.S. Treasury sells new securities to the private market via primary dealers rather than to the Fed.[8] As these securities

[8]When the Federal Reserve maintains a constant level of Treasury holdings, maturing securities are

roll off the Fed's balance sheet and must be absorbed by the private sector instead, dealer financing needs increase correspondingly. Figure 3 shows that both primary dealer positions of Treasury securities and overnight Treasury repo volumes to finance those securities increased with QT. Indeed, repo volumes were much higher than Treasury security positions indicating that primary dealers rely heavily on repo financing for their Treasury positions. In addition to dealers, levered investors such as hedge funds also participate in Treasury auctions and finance their positions in repo markets, further contributing to demand for repo funding. Figure 6 shows that hedge funds finance a substantial share of newly issued Treasury securities in the repo market, consistent with their role as levered intermediaries.

An important feature of the repo market is the identity of the marginal suppliers of funding. While both banks and MMFs lend in repo markets, MMFs are the dominant providers of funding. Figure 7 shows that MMFs consistently supply over $1 trillion in repo funding on a daily basis. By contrast, bank participation is limited: although banks are net lenders, their lending volumes are small relative to MMFs and increase primarily during periods of elevated repo rates, such as Treasury settlement dates (Figure 8). These patterns indicate that MMFs are the primary suppliers of repo funding in normal times, while banks act as residual lenders when funding conditions tighten. This distinction is central for our model, which treats MMFs as the primary providers of repo funding and captures repo market capacity through the availability of non-bank liquidity.

3 The Model

This section presents a stylized model of Treasury issuance and repo market intermediation. The model captures four key features emphasized in Section 2: (i) newly issued Treasury securities are absorbed by dealers and levered investors, (ii) these positions are financed in repo markets using collateralized borrowing, (iii) money market funds allocate liquidity across repo, Treasury bills, and the ON RRP facility, and (iv) QT does not necessarily lead to an immediate reduction in bank reserves but can first manifest in a decline in non-bank reserve investment at the ON RRP. We use this framework to study how changes in the central bank's balance sheet affect repo funding conditions.

replaced through "add-ons" at Treasury auctions. See FRBNY's FAQs on Treasury rollovers for additional details.

3.1 Environment

The model builds upon Armenter and Lester (2017) and Xiao (2020). We consider a two-period economy with six types of agents: primary dealers (dealers), a monopoly bank, MMFs, investors, households, and firms. We interpret one period in this model as events occurring over the course of a typical Treasury coupon security settlement period, which is several days. The main text focuses on dealers, MMFs, investors, and the bank, while firms and households—whose role is to generate deposit demand and support a monetary equilibrium—are described in Appendices A.2 and A.3, respectively. In addition, there is a central bank that implements monetary policy and a government that issues two types of assets: government bonds and government bills.[9]

The timing of events is as follows: At the beginning of period $t = 1$, the government issues government bonds B^L and government bills B^S. The central bank sets the interest rate on reserve balances (IORB), denoted R, the rate on the ON RRP facility, denoted r, and its holdings of government bonds, $b^{L,CB}$, thereby determining the size of its balance sheet. Thus, the policy instruments are $R, r, b^{L,CB}$. The composition of central bank liabilities—reserves and ON RRP balances—is determined endogenously in equilibrium. We assume that the central bank determines the overall size of its liabilities, while the composition between reserves and ON RRP balances adjusts endogenously to meet demand from banks and MMFs, subject to the balance sheet constraint.

The model captures the settlement-stage intermediation of Treasury issuance emphasized in Section 2. Dealers and investors absorb newly-issued bonds in the primary market, holding $b^{L,d}$ and $b^{L,i}$, respectively. Investors (which we interpret as hedge funds) finance their bond purchases by borrowing from dealers at rate ι, using government bonds as collateral.[10] Dealers, in turn, finance both their own bond purchases and their lending to investors by borrowing in the repo market at the equilibrium rate ρ. As a result, Treasury issuance generates demand for repo funding through both dealer and investor balance sheets. Dealers earn a spread $\iota - \rho$ on intermediated funds and are subject to

[9]In the model, both bonds and bills mature after one period but differ in who holds them and how they are financed. For simplicity, we abstract from bill holdings by the central bank. In practice, the Federal Reserve held only a small share of Treasury bills during our sample (approximately 5% of total Treasury holdings). Bills primarily serve as an investment option for MMFs, and abstracting from central bank bill holdings does not affect the repo-financing mechanism, which operates through longer-term securities.

[10]Dealers' borrowing can be interpreted as repo transactions. For simplicity, we do not explicitly model the centrally cleared segment of the repo market and instead treat the spread between ι and ρ as fixed and exogenous.

balance sheet costs that increase with their bond holdings. These costs capture regulatory and organizational constraints, such as the supplementary leverage ratio, that make large balance sheets costly. Finally, we assume that the spread $\iota - \rho$ is exogenous.

The monopoly bank raises deposits from households and allocates funds between reserves, which earn the IORB rate R, and loans ℓ, which yield an exogenous return i_ℓ. Following Ennis (2018), we assume loan issuance is costly, with convex cost function $\chi(\ell)$, where $\chi(0) = 0$, $\chi'(\cdot) > 0$ and $\chi''(\cdot) > 0$. The costs can be motivated, for example, through monitoring costs Holmstrom and Tirole (1997). We assume the spread $i_\ell - R$ is positive and constant in policy rate R. A constant lending spread could be motivated by assuming a perfectly competitive lending market as in Martin et al. (2016), for example. Finally, the bank is subject to reserve requirements and faces linear balance sheet costs k^b, which can be interpreted as FDIC fees, for example.

MMFs receive deposits d^m and allocate funds across three assets: (i) ON RRP at rate r, (ii) repo lending at rate ρ, and (iii) government bills purchased at price p^S. We denote MMF investments in ON RRP and repo as d^{RRP} and z^m, respectively. MMFs pay depositors an interest rate i_{d^m} and incur linear balance sheet costs k^m.[11] This structure captures the empirical substitution between repo lending and bill holdings documented in Section 2.2.

Investors purchase government bonds in the primary market and finance these positions through collateralized borrowing from dealers, capturing the role of levered investors such as hedge funds in Treasury markets.

Dealers, the bank, MMFs, and investors maximize profits, π^j:

$$\mathbb{E}\left[\sum_{t=1}^{4} \pi_t^j\right], \tag{1}$$

for $j = \{dealers, bank, MMF, investors\}$. In period $t = 2$, all positions mature: loans are repaid and asset returns are realized.

Finally, households receive an endowment of money, which can be allocated between bank deposits and MMF deposits, and consume a consumption good. Firms produce this consumption good. Their maximization problems are provided in Appendices A.3 and

[11] We introduce balance sheet costs to generate a wedge between the interest rate paid to depositors and the return on MMFs' investments, consistent with the empirical moments targeted in the calibration.

A.2, respectively. The full model—including the optimal allocation of households and firms—is used for the quantitative analysis in Section 4.

3.2 Agents' Problems

Dealers. In period $t = 1$, dealers participate in the primary government bond market, where they purchase bonds $b^{L,d}$ at price p^L. In addition, dealers lend to investors, denoted z^i, at the exogenous rate ι. To finance both their own bond purchases and their lending to investors, dealers borrow in the repo market on a collateralized basis at rate ρ.

Dealers are subject to balance sheet costs that depend on their bond holdings $b^{L,d}$. Specifically, they incur a cost $c(b^{L,d})$, where $c(0) = 0$, $c'(\cdot) > 0$, and $c''(\cdot) > 0$. These costs capture regulatory and organizational constraints that make large balance sheets costly.

The dealers' problem is:

$$\max_{b^{L,d}, z^d, z^i, b^{L,i}} (1 - p^L) b^{L,d} - \rho z^d + \iota z^i - c(b^{L,d}) \tag{2}$$

subject to

$$\begin{aligned} b^{L,i} - z^i(1+\iota) &\geq 0, \\ b^{L,d} + b^{L,i} - z^d(1+\rho) &\geq 0. \end{aligned}$$

The first constraint ensures that lending to investors is fully collateralized and has Lagrange multiplier $\lambda_{d,i}$. The second constraint captures that the dealer's repo borrowing must be backed by collateral, consisting of both the dealer's own bond holdings and bonds used to collateralize loans to investors, and has Lagrange multiplier λ_d.

The first-order conditions are:

$$b^{L,d}: \quad (1 - p^L) - c'(b^{L,d}) + \lambda_d = 0, \tag{3}$$

$$z^d: \quad -\rho - \lambda_d(1+\rho) = 0, \tag{4}$$

$$z^i: \quad \iota - \lambda_{d,i}(1+\iota) = 0. \tag{5}$$

This structure captures the settlement-stage intermediation of Treasury issuance described in Section 2, where both dealers and levered investors absorb newly-issued securities and finance these positions in the repo market.

MMFs. MMFs receive deposits d^m from households in period $t = 1$ at market clearing rate i_{d^m}. They allocate these funds across three assets: (i) repo lending z^m at rate ρ, (ii) deposits at the ON RRP facility d^{RRP} at rate r, and (iii) purchases of government bills b^S at price p^S. MMFs incur linear balance sheet costs k^m.[12]

The MMF problem is:

$$\max_{z^m, d^{RRP}, b^S, d^m} \rho z^m + r d^{RRP} + (1 - p^S) b^S - d^m (i_{d^m} + k^m) \tag{6}$$

subject to

$$d^m - z^m - d^{RRP} - p^S b^S \geq 0.$$

This constraint represents the MMF balance sheet identity and has Lagrange multiplier λ_m. The first-order conditions are:

$$z^m : \quad \rho - \lambda_m \leq 0, \tag{7}$$

$$d^{RRP} : \quad r - \lambda_m \leq 0, \tag{8}$$

$$b^S : \quad (1 - p^S) - \lambda_m p^S \leq 0, \tag{9}$$

$$d^m : \quad -(i_{d^m} + k^m) + \lambda_m \leq 0, \tag{10}$$

with complementary slackness holding.

Monopoly bank. At the beginning of period $t = 1$, taking administered rates R and r as given, the monopoly bank sets the deposit rate i_{d^b} to attract deposits $d^b(i_{d^b})$. It allocates these funds between reserves m^r and loans ℓ, so that $d^b = m^r + \ell$.

Loans yield return i_ℓ and incur a monitoring cost $\chi(\ell)$, where $\chi(0) = 0$, $\chi'(\cdot) > 0$, and $\chi''(\cdot) > 0$. The bank faces linear balance sheet costs k^b and must hold a fraction δ of deposits as reserves.

[12]We introduce balance sheet costs to generate a wedge between the interest rate paid to depositors and the average return on MMFs' investments, consistent with the data used in the calibration.

The bank solves:

$$\max_{\ell, i_{d^b}} (i_\ell - R)\ell - \chi(\ell) + d^b(i_{d^b})(R - k^b - i_{d^b}) \tag{11}$$

subject to

$$(1-\delta)d^b(i_{d^b}) \geq \ell, \quad i_{d^b} \geq 0.$$

Let λ_ℓ and $\lambda_{i_{d^b}}$ denote the Lagrange multipliers. The first-order conditions are:

$$\ell: \quad (i_\ell - R) - \chi'(\ell) - \lambda_\ell = 0, \tag{12}$$

$$i_{d^b}: \quad d^{b\prime}(i_{d^b})(R - k^b - i_{d^b}) - d^b(i_{d^b}) + \lambda_\ell(1-\delta)d^{b\prime}(i_{d^b}) + \lambda_{i_{d^b}} = 0. \tag{13}$$

Investors. We interpret investors as levered institutions, such as hedge funds, that participate in Treasury auctions and finance their positions through repo markets. Investors purchase government bonds $b^{L,i}$ in the primary market in period $t = 1$ and finance these purchases through collateralized borrowing from dealers, denoted z^i, at rate ι. In period $t = 2$, they repay their loans and receive the return on their bond holdings.

The investor problem is:

$$\max_{b^{L,i}, z^i} (1 - p^L)b^{L,i} - \iota z^i \tag{14}$$

subject to

$$b^{L,i} - z^i(1+\iota) \geq 0.$$

The constraint ensures borrowing is fully collateralized and has multiplier λ_i. In equilibrium, this implies:

$$p^L = \frac{1}{1+\iota}. \tag{15}$$

3.3 Equilibrium

Definition 1. *A competitive equilibrium in this economy is*

1. *An allocation* $\Gamma = \left\{\Gamma^D, \Gamma^{MMF}, \Gamma^B, \Gamma^I, \Gamma^{HH}, \Gamma^F, \Gamma^{CB}\right\}$ *where*

- $\Gamma^D = \left\{b^{L,d}, z_d^i, z^d\right\}$ *is a dealer allocation*
- $\Gamma^{MMF} = \left\{z^m, d^{RRP}, b^S, d^m\right\}$ *is a MMF allocation*
- $\Gamma^B = \left\{d^b, m^r, \ell\right\}$ *is a bank allocation*
- $\Gamma^I = \left\{b^{L,i}, z^i\right\}$ *is an investor allocation*
- $\Gamma^{HH} = \left\{q_1, q_2, d_h^b, d_h^m\right\}$ *is a household allocation*
- $\Gamma^F = \{q_1^s, q_2^s\}$ *is a firm allocation*
- $\Gamma^{CB} = \left\{m_{CB}^r, d_{CB}^{RRP}\right\}$ *is the central bank's liquidity provision*

2. *A price system* $\mathcal{P} = \left\{p^S, p^L, i_{d^b}, i_{d^m}, \rho, \iota, i_\ell, p_1, p_2\right\}$
3. *A central bank monetary policy* $\mathcal{M} = \left\{R, r, b^{L,CB}\right\}$
4. *A government policy* $\mathcal{G} = \{g, \tau_h\}$

such that given the price system, the government policy, and the primitive endowments of the environment $\mathcal{E} = \left\{B^L, B^S, m\right\}$*:*

1. *Agents solve their respective problems: i.e. Equations* (2), (6), (11), (14), (A.4), (A.8)
2. *A consolidated government budget constraint balances, i.e. Equation* (A.2)
3. *All markets clear:*

$$\begin{aligned}
&[\textit{Goods}] && q_1^s = q_1, \quad q_2^s = q_2 + g \\
&[\textit{Deposits}] && d^b = d_h^b, \quad d^m = d_h^m \\
&[\textit{Secured Lending}] && z^i = z_d^i, \quad z^d = z^m \\
&[\textit{Government Debt}] && B^s = b^s, \quad B^L = b^{L,CB} + b^{L,d} + b^{L,i} \\
&[\textit{Central Bank Liquidity Provision}] && d^{RRP} = d_{CB}^{RRP}, \quad m^r = m_{CB}^r
\end{aligned}$$

3.4 Characterization of Equilibrium

The government bond market. Dealers demand funding in the repo market to finance both their purchases of government bonds and their lending to investors. Since dealers do not have an endowment, their repo borrowing satisfies

$$z^d = \frac{b^{L,d} + b^{L,i}}{1+\rho}.$$

Similarly, investors do not have an endowment and must fully finance their bond holdings through collateralized borrowing from dealers:

$$z^i = \frac{b^{L,i}}{1+\iota}.$$

Using the first-order conditions (3) and (4), the optimal quantity of bonds held to maturity by dealers is given by

$$b^{L,d} = c'^{-1}\left((1-p^L) - \frac{\rho}{1+\rho}\right). \tag{16}$$

Using (15), this expression can be rewritten as

$$b^{L,d} = c'^{-1}\left(\frac{\iota}{1+\iota} - \frac{\rho}{1+\rho}\right). \tag{17}$$

This expression highlights the role of dealer balance sheets in absorbing Treasury issuance. The central bank determines how many bonds must be held by the private sector through its balance sheet choice, since market clearing requires $B^L = b^{L,CB} + b^{L,d} + b^{L,i}$.

If $\iota = \rho$, dealers act purely as cash intermediaries between MMFs and investors and optimally demand no bonds, since $b^{L,d} = c'^{-1}(0) = 0$. When $\iota > \rho$, holding bonds becomes profitable, and dealers optimally hold bonds until the marginal return from holding them equals the marginal balance sheet cost, as characterized by (17).

Repo market, Treasury bill market, and ON RRP take-up. MMFs allocate their deposits across repo lending, Treasury bills, and the ON RRP facility. Given linear preferences, MMFs invest only in the asset offering the highest return.

In equilibrium, when both repo lending and bill holdings are strictly positive, returns must be equalized:

$$1+\rho = \frac{1}{p^S}.$$

ON RRP take-up is given by

$$d^{RRP} = \begin{cases} d^m - z^m - p^S b^S & \text{if } \rho = r, \\ 0 & \text{otherwise.} \end{cases} \tag{18}$$

Market clearing in the bills market implies $b^S = B^S$, while market clearing in the repo market implies

$$z^m = \frac{b^{L,d} + b^{L,i}}{1+\rho}. \tag{19}$$

This condition captures the core mechanism of the model: Treasury issuance increases the demand for repo financing through dealer and investor balance sheets, which must be met by MMFs' supply of funding.

The bank lending and deposit markets. This section characterizes how deposits are allocated between the bank and MMFs, and how the bank allocates its balance sheet between reserves and lending.

From Equation (12), the optimal loan quantity satisfies

$$\ell = \begin{cases} \ell^* & \text{if } (1-\delta)d^b - \chi'^{-1}(i_\ell - R) \geq 0, \\ (1-\delta)d^b & \text{otherwise,} \end{cases} \tag{20}$$

where $\ell^* = \chi'^{-1}(i_\ell - R)$.

If the reserve requirement does not bind, banks choose lending such that the marginal return equals the marginal cost of issuing a loan. If the constraint binds, banks hold the required quantity of reserves and allocate the remaining deposits to loans.

Let $\eta(i_{d^b}) = (i_{d^b}/d^b) \cdot d^{b'}(i_{d^b})$ denote the household's demand elasticity of bank deposits with respect to the deposit rate. Then, from Equation (13), the optimal deposit rate $i_{d^b} > 0$ satisfies:

$$i_{d^b} = \left[\frac{\eta(i_{d^b})}{1+\eta(i_{d^b})}\right] \left(R - k^b + [(i_\ell - R) - \chi'(\ell)](1-\delta)\right) \tag{21}$$

Separately, the zero-profit condition for MMFs in the deposit market implies that the equilibrium MMF deposit rate satisfies:

$$\rho z^m + r d^{RRP} + (1 - p^S) b^S = d^m (i_{d^m} + k^m) \tag{22}$$

The MMF deposit rate i_{d^m} is therefore equal to the weighted average return on MMFs' assets, net of balance sheet costs k^m.

Finally, market clearing in the deposit markets and the households' optimization problem imply that total deposits equal the household currency endowment:

$$d^m + d^b = m.$$

3.5 Liquidity Regimes

Before deriving results on the level of the central bank's balance sheet consistent with an ample liquidity framework, we define what "ample" liquidity means for both banks and non-banks.

Liquidity regimes based on the repo market. We define two thresholds for the size of the central bank's balance sheet. These thresholds determine three liquidity regimes in the repo market: (i) abundant liquidity when $b^{CB} \geq \widetilde{b^{CB}}$, (ii) scarce liquidity when $b^{CB} < \underline{b^{CB}}$, and (iii) ample liquidity in between.

Since the central bank only holds government bonds (and not bills), we drop the L-superscript for notational simplicity.

Lemma 2. *Define the total amount of bond holdings held by the private sector as $b^p = b^{L,d} + b^{L,i}$. Market clearing in the bond market therefore satisfies $B = b^{CB} + b^p$. There exists a unique threshold*

$$\widetilde{b^{CB}} = B - \widetilde{b^p}, \tag{23}$$

where

$$\widetilde{b^p} = (1 + r)\left(\widetilde{d^m} - \frac{B^S}{1 + r}\right)$$

such that:

1. *If $b^{CB} > \widetilde{b^{CB}}$: the equilibrium has $\rho = r$, $p^S = 1/(1 + r)$, and $d^{RRP} > 0$*

2. *If* $b^{CB} = \widetilde{b^{CB}}$*: the equilibrium has* $\rho = r$*,* $p^S = 1/(1+r)$*, and* $d^{RRP} = 0$

3. *If* $b^{CB} < \widetilde{b^{CB}}$*: the equilibrium has* $\rho > r$*,* $p^S = 1/(1+\rho)$*, and* $d^{RRP} = 0$

where $\widetilde{d^m}$ *is the threshold level of MMF deposits defined as*

$$m - p_1 u'^{-1}\left(\frac{1+\widetilde{i_{d^m}}}{1+i_{d^b}}\right) = \widetilde{d^m}$$

under the threshold interest rate $\widetilde{i_{d^m}} \equiv r - k^m$*.*

Proof: See Appendix B.

Intuitively, when the central bank holds a sufficiently large amount of government bonds, the repo rate is pinned down by the ON RRP rate and abundant liquidity is absorbed through the ON RRP. When the central bank's balance sheet is small enough, ON RRP balances are zero and the repo rate is above the ON RRP rate.

Corollary 1. *Define* $\bar{\varepsilon}$ *as a spread relative to the ON RRP rate* r*, where the preferences of the central bank are such that it considers* $\rho = r + \bar{\varepsilon}$ *the highest level of repo rates consistent with ample liquidity in the repo market. There exists a unique threshold*

$$\underline{b^{CB}} = B - \underline{b^p}, \tag{24}$$

where

$$\underline{b^p} = (1 + r + \bar{\varepsilon})\left(\underline{d^m} - \frac{B^S}{1+r+\bar{\varepsilon}}\right)$$

such that:

1. *If* $b^{CB} > \underline{b^{CB}}$*: the equilibrium has* $r + \bar{\varepsilon} > \rho$

2. *If* $b^{CB} \leq \underline{b^{CB}}$*: the equilibrium has* $r + \bar{\varepsilon} \leq \rho$

where $\underline{d^m}$ *is the threshold level of MMF deposits defined as*

$$m - p_1 u'^{-1}\left(\frac{1+\underline{i_{d^m}}}{1+i_{d^b}}\right) = \underline{d^m}$$

under the threshold interest rate $\underline{i_{d^m}} \equiv r + \bar{\varepsilon} - k^m$*.*

Proof: Follows directly from the proof of Lemma 2.

This corollary formalizes the notion of "ample" liquidity used by policymakers, where repo rates may rise modestly above the ON RRP rate without indicating a breakdown in market functioning, and where ON RRP take-up equals 0. That is, liquidity in the repo market is ample if the repo rate is modestly above the ON RRP rate, while liquidity in the repo market is scarce if the repo rate is higher.

Overall, the liquidity regimes in the repo market are summarized as follows. Starting from an *abundant* liquidity regime, when the central bank reduces its government bond holdings, demand for repo financing by dealers and investors increases as they need to absorb a larger quantity of government bonds. With abundant liquidity, MMFs hold positive ON RRP balances, so increased demand in repo is met by moving funds out of the ON RRP and into repo. MMFs remain indifferent between repo and the ON RRP as the rates are equal. Once ON RRP balances have reached zero, repo rates increase beyond the ON RRP rate, and the *ample* range begins. In this range, the yield of MMF deposits increases to attract more funding from households, which allows MMFs to meet the increased demand in repo. The substitution out of bank deposits into MMFs also implies that bank reserve holdings are shrinking. At some point, repo rates increase beyond the upper threshold for the ample range and liquidity conditions become *scarce*.

Liquidity regimes based on bank reserve demand. Define the level of reserves consistent with the bank's reserve demand from its liquidity requirements as $\widetilde{m}^r = \delta d^b$. Banks hold abundant reserves if $\ell = \chi'^{-1}(i_\ell - R)$ and $m^r \geq \widetilde{m}^r$. Banks hold scarce reserves if $\ell < \chi'^{-1}(i_\ell - R)$ and $m^r = \widetilde{m}^r$; that is, their reserve requirement is binding.

3.6 Central Bank's Balance Sheet Size and the Policy Rate

Proposition 3. *The critical thresholds* $\widetilde{b^{CB}}$ *and* $\underline{b^{CB}}$ *are decreasing in the policy rate* R, r *for a constant spread* $s = R - r$ *if*

$$\frac{\mathrm{d}i_{d^m}/\mathrm{d}r}{\mathrm{d}i_{d^b}/\mathrm{d}r} > u'(q) + u''(q) \cdot q. \tag{25}$$

Proof: see Appendix B.

Proposition 3 shows that the thresholds defining the range of ample liquidity in the repo market decrease as the policy rate increases, provided that pass-through to the MMF deposit rate is sufficiently strong relative to pass-through to the bank deposit rate.

The intuition is straightforward. When the policy rate rises and the MMF deposit rate responds more strongly than the bank deposit rate, households substitute out of bank deposits and into MMFs. This reallocation reduces bank deposits and increases MMF assets under management. As a result, MMFs have greater capacity to supply funding in the repo market. With a larger supply of repo funding, the system can accommodate higher levels of Treasury issuance without putting upward pressure on the repo rate. Consequently, the central bank can operate with a smaller balance sheet while maintaining an ample liquidity environment.

3.7 Repo Market Liquidity vs. Bank's Reserve Demand

As discussed above, the model features two notions of liquidity: bank reserve holdings and liquidity in the repo market. If the central bank aims to maintain an ample liquidity environment in which neither reserves nor repo funding are scarce, the relevant constraint on the size of its balance sheet is determined by whichever condition binds first.

Proposition 4 shows that, for a sufficiently low level of policy rates, repo market liquidity is the binding constraint.

Proposition 4. *For a constant $s = i_\ell - R$ spread, there exists a level of the policy rate R' such that for policy rates $R \leq R'$, the capacity of the repo market is the binding constraint for the central bank's balance sheet size. R' satisfies*

$$\frac{\chi'^{-1}(i_\ell - R')}{1-\delta} = p_1' u'^{-1}\left(\frac{1 + r + \bar{\varepsilon} - k^m}{1 + i'_{d^b}}\right), \tag{26}$$

where the right hand side is bank deposits at $R = R'$, $r = R' - s$, and $\rho = r + \bar{\varepsilon}$.

Proof: see Appendix B.

Proposition 4 shows that when the policy rate is sufficiently low ($R \leq R'$), the capacity of the repo market determines the minimum size of the central bank's balance sheet consistent with an ample liquidity environment. In this case, $\underline{b^{CB}}$ represents the smallest balance sheet size that avoids scarcity in the repo market.

If the central bank reduces its balance sheet below $\underline{b}^{CB}$, the repo rate rises above the level consistent with ample liquidity, i.e. $\rho > r + \bar{\varepsilon}$. This reflects the fact that repo funding becomes scarce relative to demand from dealers and investors financing Treasury securities. While the model does not explicitly include an interbank market, persistent increases in repo rates are likely to spill over into unsecured funding markets. Such dynamics were observed during the September 2019 episode and again in fall of 2025, when elevated repo rates contributed to increases in the effective federal funds rate. This result highlights that repo market conditions, rather than bank reserve demand, can be the primary determinant of the minimum balance sheet size consistent with ample liquidity.

4 Calibration

Before discussing comparative statics, we describe our calibration approach. We calibrate the structural parameters of the model to match aggregate moments in the data from June 2022 through August 2024. We begin the sample in June 2022, when the Federal Reserve formally initiated QT, and end in August 2024 to allow for several months of out-of-sample validation during a period with both QT and changes in the policy rate. We end our sample prior to the January to July 2025 Treasury debt issuance suspension period (DISP), as such episodes can affect reserve balances and ON RRP usage through channels outside the model.

4.1 Agents and Data Sources

Agents. We map model agents directly to their empirical counterparts. The monopoly bank, dealers, MMFs, and the central bank correspond to commercial banks, primary dealers, government MMFs, and the Federal Reserve, respectively. We focus on primary dealers, as they are required to participate in Treasury auctions and therefore most closely match the dealers in our model (FRBNY, 2016).[13]

[13]We focus on government MMFs rather than prime MMFs because their investment composition more closely aligns with the MMFs in our model (Anderson and Kandrac, 2018). In the model, MMFs invest in Treasury bills, private repo, and ON RRP. These assets account for approximately 77% of government MMF portfolios over our calibration period, with the remainder consisting primarily of agency securities and agency repo.

We interpret investors as hedge funds that finance Treasury positions by borrowing from primary dealers in the centrally cleared or bilateral repo markets. This mapping is consistent with the role of levered investors in Treasury market intermediation documented in Section 2. Households are modeled as residual agents that allocate their cash endowment between commercial banks and government MMFs.

Data. The set of independently calibrated parameters is reported in Table 1. All values are averages over the calibration period, measured at the frequency of the underlying data. For a complete description of data sources, see Appendix Table D.1.

We map the model's administered rates R and r to the Federal Reserve's IORB and ON RRP offering rates, respectively. The central bank's holdings of government debt $b^{L,cb}$ are mapped to the Federal Reserve's holdings of U.S. Treasury securities.

For the supply of government bills B^S, we use Treasury bills held by government money market funds. For the supply of government bonds B^L, we focus on coupon-bearing Treasury securities (bonds and notes), excluding Treasury bills. This distinction is deliberate: our mechanism operates through repo financing, which is empirically concentrated in coupon Treasuries, while Treasury bills play a more limited role in repo markets. We map the model's market clearing condition $B^L = b^{L,cb} + b^{L,d} + b^{L,i}$ to the sum of Federal Reserve Treasury security holdings, primary dealer Treasury coupon positions, and investor holdings of Treasury securities.[14] This mapping reflects the flow-based nature of Treasury market intermediation emphasized in Section 2, where newly issued securities are initially absorbed by dealers and levered investors before being redistributed to other sectors.

For the outside loan return of banks i_ℓ, we use the average rates on new loan products by commercial banks.[15] The dealer lending rate ι is mapped to the volume-weighted average rate at which dealers lend to hedge funds in FICC sponsored service (see Anbil et al. (2026b) for more details). Finally, the required reserve ratio δ is calibrated to the average reserves-to-deposits ratio observed prior to the September 2019 repo episode.[16]

[14] We measure investor holdings of Treasury securities as the residual between the cumulative decline in Federal Reserve Treasury holdings since their June 2022 peak and primary dealer Treasury positions, attributing the remaining absorption to non-dealer investors.

[15] The loan rate is constructed as the average rate on new longer-term loan products offered by commercial banks (e.g., auto loans, fixed-rate mortgages, and home equity loans), using data from RateWatch. See Appendix D for details.

[16] Estimates in the literature place ample reserves at approximately 13–14% of total bank assets. In our

4.2 Targets and Parameters

Parameters. Given the independent parameters described above, and conditional on functional form assumptions, the model contains seven additional parameters that are jointly calibrated. These parameters are: the curvature of the household utility function α, the linear balance sheet costs of MMFs k^m and banks k^b, the bank's loan monitoring cost parameter β, the dealer's balance sheet cost parameter γ, the household fiat currency endowment m, and the price of currency in terms of period-$t = 2$ consumption ϕ. For a description of the functional form assumptions, see Appendix C.1.

Targets. To discipline these seven parameters, we target seven empirical moments, all computed as averages over the calibration period. These moments are: the commercial bank deposit rate i_{d^b}, the government MMF deposit rate i_{d^m}, the quantity of commercial bank deposits d^b, the quantity of government MMF deposits d^m, ON RRP take-up by government MMFs d^{RRP}, aggregate reserves m^r, and dealer holdings of Treasury securities $b^{L,d}$. We measure bank deposits using "other deposits" from the H.8 release, which primarily capture transaction balances.[17]

These targets are chosen to capture the key margins in the model: the allocation of household savings between banks and MMFs, the pricing of deposits across these intermediaries, the level of central bank liabilities (reserves and ON RRP), and the balance sheet role of dealers in absorbing Treasury supply. In particular, matching dealer holdings and MMF balance sheet allocations ensures that the model captures the financing of Treasury issuance through repo markets.

4.3 Results

As described above, we jointly calibrate seven parameters,

$$\Xi = \left\{\alpha, k^m, k^b, m, \gamma, \beta, \phi\right\}.$$

framework, deposits equal assets due to the balance sheet identity $d^b = m_R + \ell$, so calibrating $\delta = 13\%$ aligns with these estimates (Afonso et al., 2026).

[17]The H.8 release decomposes deposits into "large time deposits" and "other deposits." The latter include demand, checkable, and savings deposits, while large time deposits behave more like longer-term savings instruments. Over our sample, "other deposits" account for approximately 89% of total deposits (Lopez-Salido and Vissing-Jorgensen, 2025).

Let $S^i(\Xi)$ denote the vector of seven moments implied by the model (i = model) and observed in the data (i = data). We choose Ξ to solve

$$\begin{aligned} \min_{\Xi} \quad & \left(S^{data} - S^{model}(\Xi)\right)^2 \\ \text{s.t.} \quad & EC(\Xi) = 0, \end{aligned}$$

where $EC(\Xi)$ denotes the system of equilibrium conditions evaluated at Ξ.[18]

Panel A of Table 2 reports the fit of the model to the targeted moments, while Panel B reports the calibrated parameter values. The model matches the targeted moments closely, capturing key features of deposit pricing, balance sheet allocations, and central bank liabilities. Importantly, the model matches moments related to MMF balance sheet allocations and dealer Treasury holdings, which are central for capturing the repo-based transmission mechanism emphasized in the paper.

Discussion of calibrated parameters. The calibrated parameters reported in Panel B of Table 2 are economically plausible and broadly consistent with values used or implied in the literature.

We begin with the parameters m and ϕ, which are specific to the static structure of the model. The parameter m captures the aggregate nominal endowment of money held by households and is pinned down by the observed quantities of bank deposits d^b and MMF assets under management d^m, consistent with the household budget constraint. The parameter ϕ represents the price of money in terms of the period-$t = 2$ consumption good, $\phi = 1/p_2$. In a dynamic setting, ϕ would be linked to inflation dynamics; in our static environment, it serves as a normalization mapping nominal quantities into real allocations.

We next consider the balance sheet cost parameters k^b and k^m for banks and MMFs, respectively. These parameters generate wedges between asset returns and deposit rates and capture operational, regulatory, and intermediation costs. We estimate $k^b = 1.47\%$ for banks and $k^m = 0.18\%$ for MMFs. These magnitudes are consistent with empirical evidence. For example, Wang et al. (2022) document an average net non-interest expense of 1.23% of assets for commercial banks, and Xiao (2020) document an average operating

[18] We solve the model using a nonlinear least-squares routine (`lsqnonlin` in Matlab). The calibration is equivalent to a GMM procedure with an identity weighting matrix. See Appendix C.2 for details.

cost net of management costs of 0.16% of assets under management for MMFs.

Turning to household preferences, we estimate $\alpha = 0.0024$, which governs the curvature of utility over contemporaneous consumption. In our model, α is identified jointly by the bank deposit rate and the elasticity of deposit demand. Because we directly match the bank deposit rate, a natural validation is whether the implied deposit elasticity η is consistent with the literature. Panel B of Table 3 reports a model-implied elasticity of 0.27, which lies well within the range of existing estimates: Kurlat (2019) estimate 0.2, Chiu and Hill (2018) estimate 0.3, and Egan et al. (2017) estimate 0.56.

Finally, we consider the curvature parameters governing intermediation costs. The dealer balance sheet cost parameter γ implies, at the calibration equilibrium quantity, a dealer balance sheet marginal cost of approximately 7 basis points. This is consistent with empirical estimates of dealer balance sheet costs of 5–6 basis points (Chabot et al., 2024). The bank loan cost parameter β governs the cost function $\chi(\cdot)$. While direct estimates of loan monitoring costs are scarce, the magnitude of costs that our parameters imply are economically reasonable. At the unconstrained optimum, the model implies an average loan cost of 1.47% and an efficiency ratio of 42%. These values are broadly consistent with existing estimates: Corbae and D'Erasmo (2021) estimate an average loan cost of 1.91% for commercial banks, and the FDIC Banking Profile reports an average efficiency ratio of approximately 57% over our calibration period.

4.4 In-Sample Test

As a first assessment, we evaluate whether the calibrated model can replicate the moments targeted in the calibration. The results of this in-sample test are reported in Table 3.

The model matches the targeted moments very closely. Panel A of Table 3 shows that bank deposit rates, MMF deposit rates, bank deposits, and reserves are matched exactly, while ON RRP take-up and MMF deposits are within \$1 billion and \$5 billion of their corresponding data values, respectively.

Importantly, the model also performs well along dimensions not directly targeted in the calibration. Panel B shows that the model closely reproduces the Tri-party General Collateral Rate (TGCR), bank lending, and the price elasticity of bank deposits. These results indicate that the model captures key relationships between deposit markets, repo markets, and balance sheet allocations beyond the moments used for estimation.

4.5 Out-of-Sample Tests

We next evaluate the model's ability to match aggregate moments in out-of-sample environments. We consider three scenarios: repo market conditions following the (i) September and (ii) November 2024 FOMC rate cuts, and (iii) the September 2019 repo market stress episode.

September 2024. At the September 2024 FOMC meeting, the Committee reduced its policy rate by 50 basis points, with the IORB rate going from 5.40% to 4.90%, while maintaining a 10 basis point spread to the ON RRP rate. During this period, the SOMA Treasury portfolio averaged $4.39 trillion in Treasury holdings, government MMFs held approximately $2.52 trillion in Treasury bills, and aggregate deposits and MMF assets totaled $19.61 trillion. These inputs are summarized in Panel A of Table 4.

The model closely matches observed outcomes shown in Panel B. Predicted MMF deposits d^m, bank deposits d^b, and reserves m^r are within \$70 billion, \$140 billion, and \$50 billion of the data, respectively, while ON RRP take-up d^{RRP} is underestimated by approximately \$30 billion.

Deposit rates are also reasonably well captured. The model predicts bank deposit rates i_{d^b} that are 37 basis points below observed values and MMF deposit rates i_{d^m} that are 10 basis points lower. While the model estimates bank deposits lower than what is observed in the data, it still matches the quantity of bank deposits closely.

For untargeted moments, the model performs particularly well. The predicted repo rate ρ deviates by only 3 basis points from the average TGCR, and predicted repo lending by MMFs z^m is close to observed levels (within \$160 billion). The model also closely matches non-reserve bank assets ℓ, with a difference of only \$200 billion.

November 2024. At the November 2024 FOMC meeting, the Committee reduced its policy rate by an additional 25 basis points, moving the IORB rate to 4.65% and again maintaining a 10 basis point spread to the ON RRP rate. During this period, the SOMA Treasury portfolio averaged $4.35 trillion, MMFs held $2.64 trillion in Treasury bills, and total deposits and MMF assets reached $19.76 trillion. These inputs are summarized in Table 5 Panel A.

The model again aligns closely with the data. MMF deposits are slightly lower in the model (by approximately $50 billion), while bank deposits, ON RRP take-up, and reserves are somewhat higher (by $110 billion, $10 billion, and $20 billion, respectively).

MMF deposit rates are well matched, differing by only 11 basis points, while bank deposit rates are underestimated by 58 basis points.

For untargeted moments, the model continues to perform strongly. The predicted repo rate is within 2 basis points of the observed TGCR, and predicted repo lending by MMFs is close to observed levels (within $60 billion). Non-reserve bank assets are also reasonably well matched, with a difference of $270 billion.

A key feature of both out-of-sample periods is that the TGCR slightly exceeded the ON RRP rate, by approximately 2–3 basis points on average, while ON RRP take-up remained positive. In our model, MMFs are highly interest-rate sensitive, which implies that ON RRP take-up cannot be positive when repo rates are larger than the ON RRP rate in equilibrium. One possible explanation for this discrepancy is the presence of frictions in the repo market, such as counterparty risk limits, which constrain the ability of MMFs and dealers to expand exposures.[19] Because the model abstracts from such frictions, it predicts that the repo market remains in an abundant liquidity regime with $\rho = r$. Given that policymakers continued to characterize liquidity conditions as abundant during this period, we interpret the model's prediction as broadly consistent with observed market conditions.

September 2019. Finally, we assess the model's ability to capture scarce liquidity conditions by testing it on the September 2019 repo market stress episode. The results are reported in Table 6.

In September 2019, the IORB and ON RRP rates were 2.1% and 2.0%, respectively. The SOMA portfolio held approximately $2.1 trillion in Treasury bonds, and MMFs held

[19]Perli (2024) provides a detailed discussion of these frictions, emphasizing that counterparty risk limits can constrain the ability of dealers and MMFs to expand exposures to one another. As a result, balances may persist at the ON RRP facility even when alternative instruments, such as repo, offer higher returns. Logan (2024) similarly highlights counterparty risk limits as an important source of frictions in repo markets.

$880 billion in Treasury bills. To reflect the smaller scale of balance sheets relative to our calibration period, we adjust total Treasury bond supply to $2.64 trillion and aggregate deposits and MMF assets to $12.7 trillion as shown in Panel A.

The model matches ON RRP take-up and reserves closely, with deviations of only $2.3 billion and $10 billion, respectively. It modestly overestimates bank deposits (by $170 billion) and underestimates MMF assets (by $140 billion). The model predicts a near-zero bank deposit rate, about 16 basis points below the data, while MMF deposit rates are overestimated.[20]

The model captures a substantial portion of the spike in repo rates, predicting a repo rate of 4.61%, approximately 64 basis points below the observed TGCR on September 17. This suggests that, within the model, liquidity conditions had moved out of the ample range and into scarcity. This interpretation is consistent with the Federal Reserve's subsequent intervention to inject liquidity into funding markets by beginning reserve management purchases.

Because the model is static, it does not generate high-frequency volatility. Instead, the results should be interpreted as capturing the underlying tightening in funding conditions that contributed to the observed spike in repo rates.

These results highlight the mechanism emphasized in the model. The tightening of funding conditions was driven not by banks' reserve demand per se, but by the limited capacity of the repo market to intermediate Treasury securities. As the supply of Treasuries increased and reserves declined, dealers and levered investors required additional repo financing to absorb issuance, while the supply of repo funding from MMFs became constrained. This imbalance led to upward pressure on repo rates, even in the presence of still relatively high aggregate reserves.

Overall, these results highlight that the model captures how the size and composition of the Federal Reserve's balance sheet affect money market conditions through repo market capacity, rather than solely through bank reserve scarcity. This mechanism provides a unified framework for understanding both the September 2019 episode and the effects of quantitative tightening more broadly.

[20]The discrepancy in the MMF deposit rate reflects the frequency of the underlying data. MMF rates are measured at a weekly frequency, while the repo spike in September 2019 was short-lived and therefore did not fully pass through to reported MMF deposit rates.

5 Comparative Statics

In this section, we study the comparative statics of tightening monetary policy using the calibrated model.

5.1 Increasing the Policy Rate

First, we analyze the effects of an increase in the policy rate R, while keeping the spread $s = R - r$ constant and assuming abundant liquidity in the repo market and excess reserves in the banking sector.

An increase in the policy rate transmits to both bank and MMF deposit rates, as shown in Figure 9a. Furthermore, if condition (25) holds, the pass-through of the policy rate to MMF deposit rates exceeds that to bank deposit rates, leading to an increase in the spread $i_{d^m} - i_{d^b}$, as illustrated in Figure 9b. As a result, households reallocate funds away from bank deposits and into MMFs, as shown in Figure 9c.

On the central bank's balance sheet, this substitution from bank deposits into MMF deposits leads to a decline in reserves and an increase in ON RRP take-up, as shown in Figure 10. The intuition is straightforward. As MMF deposits become more attractive, households shift their portfolios toward MMFs and away from banks, increasing MMF assets under management. Holding the central bank's balance sheet fixed implies no change in repo funding demand from dealers. Together with fixed Treasury bill supply, this implies that MMFs have no new investment options so allocate the additional funds to the ON RRP facility.

At the same time, the balance sheet of the monopoly bank contracts as it receives fewer deposits. Since the optimal level of lending remains unchanged and banks hold excess reserves, the bank reduces its reserve holdings one-for-one with the decline in deposits. If the reserve constraint binds, both reserves and loans decline.[21]

[21] Lower deposits reduce required reserves, allowing banks to partially offset the decline through lower reserve holdings, with the remaining adjustment reflected in reduced lending. Because lending is modeled in partial equilibrium, the decline in reserves need not be one-for-one with the reduction in central bank assets. In a full general equilibrium setting, lower lending would reduce deposits system-wide, restoring a one-for-one relationship between asset reductions and reserves.

5.2 Reducing the Size of the Central Bank's Balance Sheet

Next, we analyze the effects of tightening monetary policy through a reduction in the size of the central bank's balance sheet.

Starting from an abundant liquidity regime, with $b^{CB} \geq \widetilde{b^{CB}}$ as defined in Section 3.5, a reduction in the central bank's balance sheet has the following effects.

First, a decline in central bank holdings of government bonds implies that the private sector—dealers and investors—must absorb a larger share of Treasury securities. This increases the demand for repo financing, as these agents borrow to fund their positions. At the same time, when ON RRP balances are positive, MMFs can reallocate funds from the ON RRP facility into the repo market. As a result, the repo market continues to clear at the ON RRP rate. This is illustrated on the right-hand sides of Figures 11 and 12, respectively. In particular, as the central bank's balance sheet declines, ON RRP balances fall as MMFs shift funds into repo lending (Figure 11). As long as ON RRP balances remain positive, the repo rate remains equal to the ON RRP rate (Figure 12). Thus, in the abundant liquidity regime, a reduction in the central bank's balance sheet primarily affects quantities—ON RRP balances and repo lending volumes—rather than prices.

Once the first critical threshold, $\widetilde{b^{CB}}$, is reached (at the dashed line in Figure 11), ON RRP balances are exhausted. At this point, MMFs can no longer substitute funds from ON RRP into the repo market and repo rates begin to increase. As a result, MMF demand for deposits increase, leading to an increase in the MMF deposit rate and therefore increasing the spread $i_{d^m} - i_{d^b}$. The increase in the spread induces households to shift their deposits from banks to MMFs, increasing available funds for MMFs to lend in repo. Figure 12 shows how repo rates rise once the central bank's balance sheet falls below $\widetilde{b^{CB}}$, while Figure 13 illustrates the corresponding shift from bank deposits to MMF deposits.

Figure 14 shows how reserve holdings evolve. Initially, in the abundant liquidity regime, a reduction in the central bank's assets is matched one-for-one by a decline in ON RRP balances, leaving reserves unchanged. Once the threshold $\widetilde{b^{CB}}$ is reached (at the dashed line), ON RRP balances are depleted and liquidity becomes ample. Further balance sheet reduction induces a reallocation from bank deposits to MMF deposits, shrinking the banking sector's balance sheet. Since the reserve constraint is non-binding, the bank reduces reserves one-for-one with the decline in deposits. Thus, beyond this threshold, reductions in central bank assets translate directly into declines in reserves.

5.3 Liquidity Thresholds and the Capacity of the Repo Market

In this subsection, we compare the constraint on the central bank's balance sheet arising from repo market conditions with that arising from banks' reserve demand. Figure 15 plots the critical thresholds in the repo market $\widetilde{b^{CB}}$ and $\underline{b^{CB}}$ as their corresponding level of reserves held by banks in equilibrium. The red dotted line corresponds to reserves held by banks when $b^{CB} = \widetilde{b^{CB}}$, the threshold between abundant and ample liquidity in the repo market, and the blue dashed line corresponds reserves held by banks when $b^{CB} = \underline{b^{CB}}$, the threshold between ample and scarce liquidity. The yellow line represents the reserve constraint for the banking sector. As discussed in Section 3.5, scarce reserves arise when this constraint binds.

Figure 15 shows that the constraint implied by repo market conditions is substantially tighter than that implied by bank reserve demand. This indicates that repo market capacity is the binding constraint on the size of the central bank's balance sheet. At IORB of 4.5% (black vertical dashed line), the model implies that reserves between approximately \$2.7 and \$3.3 trillion are consistent with an ample liquidity regime.

This result is also reflected in Table 7, which shows the model-implied level of the repo spread $\rho - R$ for given levels of reserves m^r. When repo rates are such that $\rho = r + \bar{\varepsilon} = R$, our assumed lower bound of the ample regime, this corresponds to approximately \$2.7 trillion in reserves.[22] Table 7 further shows that reducing reserves to approximately \$2.1 trillion—close to the minimum level implied by the bank's reserve constraint at $R = 4.5\%$—would increase repo rates to roughly 10 basis points above IORB. Further reductions would push repo rates above the top of the target range for the federal funds rate. As highlighted by recent events and discussed in Anbil et al. (2026a), increases in repo rates can transmit to the federal funds rate, underscoring the importance of repo market conditions for interest rate control.

Taken together, these results show that, in an environment similar to our calibration

[22] If the central bank tolerates repo rates above the IORB rate, the lower bound of the ample regime would be smaller. Quantitatively, the calibration also depends on the level of Treasury bill supply. Because bill issuance has increased since our sample period, the implied lower bound on reserves as of end-2025 would be somewhat higher.

period, the binding constraint on the central bank's balance sheet is determined by repo market capacity rather than by banks' reserve demand. As Treasury supply is shifted to the private sector, dealers and levered investors require additional repo financing, and the ability of MMFs to supply funding becomes the key margin determining equilibrium outcomes. Once ON RRP balances are exhausted, repo rates rise, even in the presence of substantial aggregate reserves. In contrast, bank reserve demand adjusts endogenously through deposit reallocation and does not impose a binding constraint. This mechanism provides a unified explanation for how balance sheet reductions transmit to money market conditions and highlights the central role of repo markets in the implementation of monetary policy.

5.4 Interest Rate and Balance Sheet Policy as Complements

Finally, we discuss how balance sheet policy and interest rate policy interact. As shown formally in Proposition 3 and illustrated in Figure 15, both repo market thresholds, $\widetilde{b^{CB}}$ and $\underline{b^{CB}}$, decrease with the policy rate R. This implies that, if the central bank aims to maintain an ample liquidity regime—that is, a balance sheet size such that $\widetilde{b^{CB}} \leq b^{L,CB} \leq \underline{b^{CB}}$—the range of balance sheet sizes consistent with ample liquidity declines as the policy rate increases. In other words, the central bank can operate with a smaller balance sheet when the policy rate is higher.

The intuition is as follows. Higher policy rates increase the spread between MMF deposit rates and bank deposit rates, making MMF deposits relatively more attractive. As a result, households allocate a larger share of their funds to MMFs, increasing MMF assets under management. This directly expands the capacity of MMFs to lend in the repo market. More repo funding supply allows the system to accommodate higher Treasury financing needs without putting upward pressure on repo rates. Consequently, the central bank can maintain a smaller balance sheet consistent with ample liquidity when the policy rate is higher. This generates a novel complementarity between interest rate policy and balance sheet policy in the implementation of monetary policy.

6 Discussion of Simplifying Model Assumptions

In this section, we discuss several simplifying model assumptions and their implications.

6.1 Who Can Hold Government Securities

In the model, only the central bank, dealers, and investors hold government bonds, while only MMFs hold government bills. We now discuss how relaxing these assumptions would affect the results.

Treasury bills. Whether allowing agents other than MMFs to hold Treasury bills would expand repo market capacity depends on the underlying scenario. In the current model, Treasury bill supply is exogenous and bills are not held by the central bank. If other agents such as households or banks held some of this fixed bill supply instead of MMFs, MMFs would hold smaller bill portfolios but would correspondingly receive fewer deposits, potentially leaving their capacity to lend in the repo market unchanged.

However, if the central bank held Treasury bills and reduced these holdings as part of its balance sheet policy, allowing other agents to hold bills could expand repo capacity. As the central bank's balance sheet shrinks, reserves and bank deposits decline correspondingly. If households use these deposits to absorb the bills leaving the Fed's balance sheet, MMFs would hold smaller bill portfolios while holding stable deposits. This would free up additional funds for MMFs to lend in the repo market, allowing the central bank to reduce its balance sheet further while maintaining ample liquidity conditions.

Treasury bonds. We consider the effects of allowing either banks or households to hold Treasury bonds. If banks could hold Treasury bonds, the share of bonds held by dealers and investors would decline. Banks would fund these holdings by reducing excess reserves rather than through the repo market, thereby reducing overall demand for repo financing. With unchanged MMF deposits but lower repo demand from dealers, repo market capacity would effectively expand, allowing the central bank to operate with a smaller balance sheet. However, this effect is likely modest in practice. Dealers and investment funds maintain a dominant presence at primary auctions of Treasury securities, and Cordes and Ferris (2024) show that banks have not materially increased their Treasury securities holdings during periods of balance sheet reduction.

Allowing households to hold government bonds would not materially affect our results. While households can purchase Treasury securities, their participation in primary auctions is minimal. Instead, households acquire Treasuries either in the secondary market or indirectly through intermediaries such as mutual funds and hedge funds. Since our mechanism operates through the repo financing of newly-issued securities in the primary market, household holdings acquired through other channels would not affect the model's predictions.

Taken together, these extensions suggest that allowing additional agents to hold Treasury securities may expand repo market capacity and enable the central bank to operate with a smaller balance sheet. Consequently, our estimates of the minimum balance sheet size and associated reserve levels should be interpreted as upper bounds.

6.2 Who Can Lend in the Repo Market

For simplicity, we assume only MMFs can lend in the repo market. A reasonable alternative may be to allow banks to lend in repo. Allowing banks to participate in repo would not change the main mechanism of the model but could affect the overall capacity of the repo market.

Empirically, banks do participate in repo markets, but their role is limited. While banks are net lenders, their lending volumes are quite small relative to MMFs and are concentrated in periods when repo rates are elevated, such as Treasury settlement days (Figure 8). By contrast, MMFs provide large and stable volumes of lending on a daily basis (Figure 7). This pattern is largely consistent with the model: repo funding demand is driven by dealers and levered investors financing Treasury securities, while the supply of funding is primarily determined by MMFs.

In the model, banks choose between holding reserves, which earn IORB, and making loans. If banks were allowed to lend in repo, they would do so only when repo rates exceed IORB. Banks would therefore act as a secondary or residual source of funding, stepping in only when repo rates are sufficiently high. In our calibration, we assume the upper bound of the repo rate range is $\bar{\rho} = R$, implying that in ample liquidity, repo rates are less than IORB, so banks would have no incentive to lend even if permitted.

Under a more relaxed definition of ample liquidity, allowing banks to lend in repo may not directly expand aggregate repo market capacity, as bank lending would largely substitute for MMF lending. The total amount available to lend in repo is dependent on the household endowment, which is fixed. However, bank participation could moderate the rate of increase needed to attract additional funding, potentially allowing for more balance sheet reduction for a given repo rate target. We abstract from this channel for tractability and because it does not alter the main mechanism, which emphasizes the dominant role of MMFs in repo intermediation.

6.3 The Lending Technology

We model bank lending as an exogenous technology, where the monopoly bank faces perfectly elastic loan demand at an exogenous rate i_ℓ. We assume a positive spread $\mathcal{S} \equiv i_\ell - R > 0$, which ensures that banks have an incentive to lend. This assumption is consistent with empirical evidence (Wang et al., 2022; Diamond et al., 2024). For simplicity, we also assume that this spread is constant with changes in the policy rate R; that is, we assume the exogenous lending rate moves one-for-one with the policy rate. We now discuss how relaxing this assumption would affect our main results.

Main results remain robust. Our core findings about repo market capacity do not depend on the constant spread assumption. First, our definition of ample liquidity in the repo market is unchanged, as it depends on dealer demand for repo financing and MMF supply of funding, neither of which is directly affected by bank lending behavior. Second, Proposition 3—which shows that repo market capacity increases with the policy rate—continues to hold because it operates through the deposit rate pass-through channel described in condition (25), not through bank lending.

Where the assumption matters. Relaxing the constant spread assumption primarily affects the banking sector's internal allocation between lending and reserves. This has two implications. First, it may change the level of reserves at which the reserve constraint becomes binding, though it does not alter the constraint itself. Second, it may affect whether there exists a critical policy rate R' above which bank reserve demand becomes the binding constraint on the Fed's balance sheet (Proposition 4). The proof in Appendix E shows how the existence and location of R' depend on the relationship between i_ℓ and R.

The mechanism is straightforward. If the loan rate increases faster than R (i.e., $\mathcal{S}' > \mathcal{S}$), bank lending becomes more attractive as the policy rate rises. Banks therefore hold fewer excess reserves and the reserve constraint binds once the policy rate becomes sufficiently high. Conversely, if the loan rate increases more slowly than R (i.e., $\mathcal{S}' < \mathcal{S}$), lending becomes less attractive and banks hold more reserves, meaning the constraint does not bind for reasonable levels of the policy rate.

Robustness. Despite these shifts in when the reserve constraint binds, Appendix E shows that the qualitative results in Figure 15 may continue to hold for reasonable parameter values. In particular, there remains a wide region where reserve levels associated with the repo market thresholds $\widetilde{b^{CB}}$ and $\underline{b^{CB}}$ exceed minimum reserve requirements, confirming that repo market capacity is the binding constraint for the Fed's balance sheet in the policy-relevant range.

6.4 Frictions in the Repo Market

We assume that the repo market is perfectly competitive. In practice, several frictions may limit the capacity of MMFs to supply repo funding. First, persistent lending relationships between MMFs and dealers can create counterparty-specific limits on repo exposures (Mojir and Anbil, 2025; Anbil et al., 2023). Second, even when repo rates exceed the ON RRP rate, counterparty risk limits may prevent MMFs from fully reallocating funds from the ON RRP to the repo market.

This second friction was evident in late 2024 and early 2025, when repo rates began rising while substantial ON RRP balances remained. In our frictionless model, MMFs would have fully drained the ON RRP once repo rates exceeded the ON RRP rate. The persistence of positive ON RRP balances alongside elevated repo rates suggests that counterparty limits constrained MMFs' ability to expand repo lending, consistent with observations by Perli (2024) and Logan (2024).

We abstract from these frictions for tractability. However, to the extent that they limit MMFs' ability to supply funding, incorporating them would reduce repo market capacity. This would imply that $\widetilde{b^{CB}}$—the threshold between abundant and ample liquidity—is larger than our estimates suggest. Consequently, our estimates of the minimum balance sheet size should be interpreted as a lower bound: accounting for repo market frictions would require the central bank to maintain a larger balance sheet to remain in an ample liquidity regime.

6.5 The Spread between Administered Rates

In the model, we maintain a constant spread $s = R - r$ between the IORB and ON RRP rates for simplicity. In practice, the central bank can adjust these rates independently, thereby changing the spread. The Fed has made such "technical adjustments" in the past. We now discuss how changes in the spread affect the thresholds that define the ample liquidity regime.

Effect on the abundant/ample threshold. Narrowing the policy spread (either by raising r or lowering R) increases the gap between MMF and bank deposit rates, $i_{d^m} - i_{d^b}$. Intuitively, this makes MMF deposits more attractive to households, leading them to shift funds toward MMFs. As MMF assets under management increase, so does repo market capacity. As a result, the private sector can absorb more Treasury supply before exhausting ON RRP balances, allowing the Fed to operate with a smaller balance sheet while maintaining abundant liquidity. Formally, the threshold $\widetilde{b^{CB}}$ declines. Conversely, widening the spread compresses $i_{d^m} - i_{d^b}$, reduces repo capacity, and increases $\widetilde{b^{CB}}$.

Effect on the ample/scarce threshold. The effect on the threshold $\underline{b^{CB}}$ depends on how the central bank defines the upper bound of ample liquidity, $\bar{\rho}$. Recall that $\bar{\rho}$ represents the maximum repo rate consistent with "ample" conditions. If policymakers define this bound as a fixed spread over the ON RRP rate (i.e., $\bar{\rho} = r + \bar{\varepsilon}$ with constant $\bar{\varepsilon}$), then changes in R alone do not affect $\underline{b^{CB}}$. However, if the upper bound is instead tied to the IORB rate (e.g., $\bar{\rho} = R$), then increasing R lowers $\underline{b^{CB}}$ by raising the tolerance for elevated repo rates.

Net effect on the ample range. Both channels described above affect the width of the ample regime—the range of Fed balance sheet sizes between $\widetilde{b^{CB}}$ and $\underline{b^{CB}}$.[23]

To fix ideas, consider an increase in R while holding r fixed, which widens the policy spread. This has two effects. First, it raises the abundant/ample threshold $\widetilde{b^{CB}}$ by reducing repo capacity. Second, it lowers the ample/scarce threshold $\underline{b^{CB}}$ by increasing the central bank's tolerance for higher repo rates. Together, these effects expand the

[23] This summary assumes that the upper bound on the repo rate $\bar{\rho}$ is tied to the IORB rate. If instead $\bar{\rho}$ is defined relative to the ON RRP rate, the width of the ample range is unchanged and only its level shifts with changes in repo capacity.

range of balance sheet sizes consistent with ample liquidity. Conversely, lowering R while holding r fixed narrows the policy spread, which lowers $\widetilde{b^{CB}}$ and raises $\underline{b^{CB}}$, shrinking the ample range from both sides.

The practical implication is that maintaining a wider spread between administered rates expands the set of balance sheet sizes consistent with ample liquidity, though at the cost of reduced repo market capacity at any given balance sheet size.

In practice, for most of the recent tightening cycle, the Fed has maintained a constant spread, with only one technical adjustment occurring in December 2024, suggesting a preference for simplicity in communication over the additional flexibility that varying spreads might provide.

7 Conclusion

In this paper, we present a new constraint on the central bank's balance sheet: repo market capacity. While existing literature focuses on reserve demand as the binding constraint for the central bank's balance sheet, we show that the supply of money held by non-banks is the binding constraint. We develop a structural model that features both the demand for reserves by banks and non-bank liquidity supply, calibrated to the most recent monetary tightening cycle. The model yields two key results. First, repo market capacity—not bank reserve demand—is the binding constraint on the size of the Fed's balance sheet. Second, there is a novel complementarity between the Fed's interest rate and balance sheet policies: higher interest rates allow for a smaller balance sheet.

The mechanism operates through three stages that reveal why repo markets, not reserves, bind first. As the Fed shrinks its balance sheet from abundant liquidity, dealers and levered investors must absorb more Treasuries and finance them through repo markets. Initially, MMFs shift funds from ON RRP to repo lending while reserves remain stable—demonstrating that aggregate reserve levels can stay high even as funding conditions tighten. Once ON RRP balances reach zero and liquidity becomes ample, repo rates rise, attracting deposits from banks to MMFs, which depletes reserves. Eventually, repo rates exceed a threshold where liquidity becomes scarce. Our calibration confirms that this repo market capacity constraint binds well before reserves become scarce.

This finding has important policy implications. The Fed's minimum balance sheet size depends on policymakers' tolerance for repo market volatility, which can spill over to the federal funds market. More broadly, our framework suggests that focusing exclusively on reserve levels may lead to balance sheet reductions that go too far and compromise interest rate control. Our model provides a tool for determining the minimum supply of short-term liabilities needed to maintain interest rate control and highlights the critical role of non-bank financial intermediaries in modern monetary policy implementation. These insights are particularly timely given emerging discussions about strategies to further reduce the Fed's balance sheet. Our results underscore that such efforts must account for repo market capacity alongside traditional reserve metrics to avoid compromising interest rate control.

References

Abdullah, Samin, and Manjola Tase, 2025, Policy rate uncertainty and money market funds (MMF) portfolio allocations, *Finance and Economics Discussion Series* 2025-063, Washington: Board of Governors of the Federal Reserve System.

Acharya, Viral V, Rahul S Chauhan, Raghuram Rajan, and Sascha Steffen, 2023, Liquidity dependence and the waxing and waning of central bank balance sheets, Working Paper 31050, National Bureau of Economic Research.

Afonso, Gara, Marco Cipriani, and Gabriele La Spada, 2022, Banks' balance-sheet costs, monetary policy, and the ON RRP, *Federal Reserve Bank of New York Staff Report* 1041.

Afonso, Gara, Domenico Giannone, Gabriele La Spada, and John C Williams, 2026, Scarce, abundant, or ample? A time-varying model of the reserve demand curve, *Journal of Finance* forthcoming.

Afonso, Gara, Kyungmin Kim, Ed Nosal, Simon Potter, and Sam Schulhofer-Wohl, 2023a, Monetary policy implementation with an ample supply of reserves, *Federal Reserve Bank of New York Staff Report* 910.

Afonso, Gara, Gabriele La Spada, Thomas M. Mertens, and John C. Williams, 2023b, The optimal supply of central bank reserves under uncertainty, *Federal Reserve Bank of New York Staff Report* 1077.

Anbil, Sriya, Alyssa Anderson, Lucy Cordes, and Romina Ruprecht, 2026a, Repo markets and the Fed's balance sheet: Implications for monetary policy implementation, *FEDS Notes* .

Anbil, Sriya, Alyssa Anderson, Marco Macchiavelli, and Andy Xia, 2026b, Treasury financing by hedge funds: A tale of three channels, *SSRN* .

Anbil, Sriya, Alyssa Anderson, and Zeynep Senyuz, 2021, Are repo markets fragile? Evidence from September 2019, *Finance and Economics Discussion Series* 2021-028, Washington: Board of Governors of the Federal Reserve System.

Anbil, Sriya, Alyssa Anderson, and Zeynep Senyuz, 2023, The importance of repo trading relationships during market stress, *SSRN* .

Anbil, Sriya, and Zeynep Senyuz, 2022, How has monetary and regulatory policy affected trading relationships in the U.S. repo market?, *International Journal of Central Banking* 18(4), 1–52.

Anderson, Alyssa, Alessandro Barbarino, Anthony M. Diercks, and Stephen Miran, 2026, A user's guide to reducing the Federal Reserve's balance sheet, *Finance and Economics Discussion Series* 2026-019.

Anderson, Alyssa, and John Kandrac, 2018, Monetary Policy Implementation and Financial Vulnerability: Evidence from the Overnight Reverse Repurchase Facility, *The Review of Financial Studies* 31, 3643–3686.

Arazi, Martin, 2025, Monetary policy normalization in the new normal: The role of quantitative tightening, *Working Paper* .

Armenter, Roc, and Benjamin Lester, 2017, Excess reserves and monetary policy implementation, *Review of Economic Dynamics* 23, 212–235.

Bauer, Michael D., and Glenn D. Rudebusch, 2014, The signaling channel for Federal Reserve bond purchases, *International Journal of Central Banking* 10, 233–289.

Begenau, Juliane, and Erik Stafford, 2025, Uniform rate setting, branch expansion, and the deposit channel, *SSRN* .

Bernanke, Ben S., and Alan S. Blinder, 1988, Credit, money, and aggregate demand, *The American Economic Review* 78, 435–439.

Bhattarai, Saroj, Gauti B Eggertsson, and Bulat Gafarov, 2022, Time consistency and duration of government debt: A model of quantitative easing, *The Review of Economic Studies* 90, 1759–1799.

Bluwstein, Kristina, Ethan Cohen, Michael McLeay, and Jacob Stevens, 2026, Monetary policy transmission with endogenous reserve supply, *Working Paper* .

Bostrom, Erik, 2025, Insights from MMF portfolio allocations amid balance sheet normalization, *FEDS Notes* Washington: Board of Governors of the Federal Reserve System.

Boyarchenko, Nina, David O. Lucca, and Laura Veldkamp, 2021, Taking orders and taking notes: Dealer information sharing in Treasury auctions, *Journal of Political Economy* 129, 607–645.

Brainard, William C., 1967, Uncertainty and the effectiveness of policy, *The American Economic Review* 57, 411–425.

Carlson, Mark, and Burcu Duygan-Bump, 2021, "Unconventional" monetary policy as conventional monetary policy: A perspective from the United States in the 1920s, *International Journal of Central Banking* 17, 207–253.

Carlson, Mark A., and David C. Wheelock, 2014, Navigating constraints: The evolution of Federal Reserve monetary policy, 1935-59, *Finance and Economics Discussion Series* 2014-44, Washington: Board of Governors of the Federal Reserve System.

Chabot, Lia, Paul Cochran, Sebastian Infante, and Benjamin Iorio, 2024, Dealer balance sheet constraints evidence from dealer-level data across repo market segments, *FEDS Notes* .

Chiu, Ching-Wai (Jeremy), and John Hill, 2018, The rate elasticity of retail deposits in the United Kingdom: A macroeconomic investigation, *International Journal of Central Banking* 14, 113–158.

Copeland, Adam, Darrell Duffie, and Yilin (David) Yang, 2025, Reserves were not so ample after all, *Quarterly Journal of Economics* 140, 239–281.

Corbae, Dean, and Pablo D'Erasmo, 2021, Capital Buffers in a Quantitative Model of Banking Industry Dynamics, *Econometrica* 89, 2975–3023.

Cordes, Lucy, and Erin Ferris, 2024, Who buys Treasuries when the Fed reduces its holdings, *FEDS Notes* Washington: Board of Governors of the Federal Reserve System.

Correa, Ricardo, Wenxin Du, and Gordon Y Liao, 2025, U.S. banks and global liquidity, *Journal of Political Economy Macroeconomics* 3, 574–620.

Crawley, Edmund, Etienne Gagnon, James Hebden, and James Trevino, 2022, Substitutability between balance sheet reductions and policy rate hikes: Some illustrations and a discussion, *FEDS Notes* Washington: Board of Governors of the Federal Reserve System, June 03, 2022.

d'Avernas, Adrien, Damon Petersen, and Quentin Vandeweyer, 2025, The central bank's balance sheet and Treasury market disruptions, *SSRN* .

d'Avernas, Adrien, and Quentin Vandeweyer, 2024, Treasury bill shortages and the pricing of short-term assets, *The Journal of Finance* 79, 4083–4141.

Diamond, William F, Zhengyang Jiang, and Yiming Ma, 2024, The reserve supply channel of unconventional monetary policy, *Journal of Financial Economics* 159.

Drechsler, Itamar, Alexi Savov, and Philipp Schnabl, 2017, The deposits channel of monetary policy, *The Quarterly Journal of Economics* 132, 1819–1876.

Du, Wenxin, Kristin Forbes, and Matthew N Luzzetti, 2024, Quantitative tightening around the globe: What have we learned?, *NBER Working Paper No. 32321* .

Duffie, Darrell, 2026, The payment system puts a floor in the Fed's balance sheet, *Brookings Papers on Economic Activity* .

D'Amico, Stefania, and Thomas B. King, 2013, Flow and stock effects of large-scale Treasury purchases: Evidence on the importance of local supply, *Journal of Financial Economics* 108, 425–448.

Egan, Mark, Ali Hortaçsu, and Gregor Matvos, 2017, Deposit competition and financial fragility: Evidence from the US banking sector, *American Economic Review* 107, 169–216.

Ennis, Huberto M, 2018, A simple general equilibrium model of large excess reserves, *Journal of Monetary Economics* 98, 50–65.

Fleming, Michael, Giang Nguyen, and Joshua Rosenberg, 2024, How do Treasury dealers manage their positions?, *Journal of Financial Economics* 158.

FOMC, 2019, Statement regarding monetary policy implementation and balance sheet normalization, Available at https://www.federalreserve.gov/newsevents/pressreleases/monetary20190130c.htm.

FOMC, 2022a, Federal Reserve issues FOMC statement, Available at https://www.federalreserve.gov/newsevents/pressreleases/monetary20220316a.htm.

FOMC, 2022b, Plans for reducing the size of the Federal Reserve's balance sheet, Available at https://www.federalreserve.gov/newsevents/pressreleases/monetary20220504b.htm.

FRBNY, 2016, Operating policy: Administration of relationships with primary dealers, Available at https://www.newyorkfed.org/markets/pridealers_policies.html.

Gorton, Gary, and Andrew Metrick, 2012, Securitized banking and the run on repo, *Journal of Financial Economics* 104, 425–451.

He, Zhiguo, Stefan Nagel, and Zhaogang Song, 2022, Treasury inconvenience yields during the COVID-19 crisis, *Journal of Financial Economics* 143, 57–79.

Holmstrom, Bengt, and Jean Tirole, 1997, Financial intermediation, loanable funds, and the real sector, *The Quarterly Journal of Economics* 112, 663–691.

Im, Jay, Yi Li, and Ashley Wang, 2023, Investor flows, monetary policy, and portfolio management of money market funds, *SSRN* .

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

Kiley, Michael T., 2014, The response of equity prices to movements in long-term interest rates associated with monetary policy statements: Before and after the zero lower bound, *Journal of Money, Credit and Banking* 46, 1057–1071.

Koijen, Ralph S.J., François Koulischer, Benoît Nguyen, and Motohiro Yogo, 2021, Inspecting the mechanism of quantitative easing in the euro area, *Journal of Financial Economics* 140, 1–20.

Krishnamurthy, Arvind, and Wenhao Li, 2023, The demand for money, near-money, and Treasury bonds, *The Review of Financial Studies* 36, 2091–2130.

Krishnamurthy, Arvind, and Annette Vissing-Jorgensen, 2011, The effects of quantitative easing on interest rates: Channels and implications for policy, *Brookings Papers on Economic Activity* Fall 2011, 215–287.

Kumhof, Michael, and Mauricio Salgado-Moreno, 2024, Quantitative easing and quantitative tightening: The money channel, *Bank of England Financial Stability Paper* 1090.

Kurlat, Pablo, 2019, Deposit spreads and the welfare cost of inflation, *Journal of Monetary Economics* 106, 78–93.

Lagos, Ricardo, and Gastón Navarro, 2023, Monetary policy operations: Theory, evidence, and tools for quantitative analysis, *NBER Working Paper No. 31370* .

Li, Wenhao, Yiming Ma, and Yang Zhao, 2026, The passthrough of Treasury supply to bank deposit funding, *Journal of Financial Economics* forthcoming.

Logan, Lorie, 2024, Normalizing the FOMC's monetary policy tools, *Speech by Lorie Logan, President of the Federal Reserve Bank of Dallas - Remarks at the Securities Industries and Financial Markets Association Annual Meeting* .

Logan, Lorie, and Sam Schulhofer-Wohl, 2026, Options for reducing the size of the Fed's balance sheet, Working paper, Federal Reserve Bank of Dallas.

Lopez-Salido, David, and Annette Vissing-Jorgensen, 2025, Reserve demand, interest rate control, and quantitative tightening, Working paper, Federal Reserve Board.

Martin, Antoine, James McAndrews, and David Skeie, 2016, Bank lending in times of large bank reserves, *International Journal of Central Banking* 12, 192–222.

Mojir, Navid, and Sriya Anbil, 2025, The value of professional ties in B2B markets, *Marketing Science* 44, 975–1215.

Perli, Roberto, 2024, Balance sheet normalization: Monitoring reserve conditions and understanding repo market pressures, *Speech by Roberto Perli, Manager of the System Open Market Account - Remarks at the 2024 U.S. Treasury Market Conference, Federal Reserve Bank of New York, New York City* .

Sims, Eric, and Jing Cynthia Wu, 2020, Are QE and conventional monetary policy substitutable?, *International Journal of Central Banking* 16, 195–230.

Smith, A. Lee, and Victor J. Valcarcel, 2023, The financial market effects of unwinding the Federal Reserve's balance sheet, *Journal of Economic Dynamics and Control* 146.

Stein, Jeremy C., 2012, Monetary Policy as Financial Stability Regulation, *The Quarterly Journal of Economics* 127, 57–95.

Wang, Yifei, Toni M. Whited, Yufeng Wu, and Kairong Xiao, 2022, Bank market power and monetary policy transmission: Evidence from a structural estimation, *The Journal of Finance* 77, 2093–2141.

Williams, John C., 2013, A defense of moderation in monetary policy, *Journal of Macroeconomics* 38, 137–150.

Xiao, Kairong, 2020, Monetary transmission through shadow banks, *The Review of Financial Studies* 33, 2379–2420.

Yang, Yilin (David), 2026, What quantity of reserves is sufficient?, *SSRN* .

Yellen, Janet, 2016, The Federal Reserve's monetary policy toolkit: Past, present, and future, Available at `https://www.federalreserve.gov/newsevents/speech/yellen20160826a.htm`.

Tables and Figures

Table 1. Independent Parameters

Parameter	Description	Source	Value
R	IORB	Data	4.57%
r	ON RRP offering rate	Data	4.47%
i_ℓ	Average interest rate on banks' outside investments	Data	7.51%
ι	Volume-weighted average rate that dealers lend to hedge funds	Data	4.55%
$p^L b^{L,cb}$	Nominal quantity of Treasury securities held by the Fed	Data	\$5.09T
$p^L B^L$	Nominal quantity bonds in the economy	Data	\$5.77T
$p^S B^S$	Nominal quantity bills in the economy	Data	\$1.62T
δ	Minimal reserve-to-deposit ratio	Data	0.13

This table displays the independent parameters used to calibrate the model. The values are averages from the data during our calibration period between June 1, 2022, and August 31, 2024. Variable descriptions and data sources are available in Table D.1.

Table 2. Model Fit and Internally Calibrated Parameters

Panel A. Moments				
Parameter	**Description**	**Data**	**Model**	**Error**
i_{d^b}	Interest rate on bank deposits	0.663%	0.663%	(0)
i_{d^m}	Interest rate on MMF deposits	4.29%	4.29%	(0)
d^m	MMF Deposits	\$3.43T	\$3.48T	+(0.05)
d^b	Bank Deposits	\$15.60T	\$15.60T	(0)
d^{RRP}	ON RRP usage	\$1.19T	\$1.18T	−(0.01)
m^r	Reserves	\$3.28T	\$3.28T	(0)
$p^L b^{L,d}$	Bonds bought and held by dealers	\$0.13T	\$0.13T	(0)
Panel B. Calibrated Parameters				
Parameter	**Description**		**Source**	**Value**
m	Money endowment to households		Internal	\$19.08T
ϕ	Price level in $t = 2$		Internal	0.0143
k^b	Balance sheet costs of banks		Internal	1.47%
k^m	Balance sheet costs of MMF		Internal	0.18%
α	Relative risk aversion		Internal	0.024
γ	Dealers balance sheet cost function		Internal	0.005
β	Loan cost function		Internal	0.0024

Panel A of this table displays the moments targeted in our calibration. The "Data" column shows the average value over our calibration period (June 1, 2022 to August 31, 2024). The "Model" column displays the model's generated moments. "Error" displays the difference between the two columns. Variable descriptions and data sources are available in Table D.1. Panel B of this table shows the values of the model parameters that are jointly calibrated.

Table 3. In-Sample Model Performance

Panel A. Model Performance – Targeted				
Parameter	**Description**	**Data**	**Model**	**Error**
i_{d^b}	Interest rate on bank deposits	0.663%	0.663%	(0)
i_{d^m}	Interest rate on MMF deposits	4.29%	4.29%	(0)
d^m	MMF Deposits	\$3.43T	\$3.48T	+(0.05)
d^b	Bank Deposits	\$15.60T	\$15.60T	(0)
d^{RRP}	ON RRP usage	\$1.19T	\$1.18T	−(0.01)
m^r	Reserves	\$3.28T	\$3.28T	(0)
Panel B. Model Performance – Untargeted				
Parameter	**Description**	**Data**	**Model**	**Error**
ρ	TGCR	4.46%	4.47%	+(0.01)
η	Price Elasticity of Bank Deposits	0.32	0.27	−(0.05)
z^m	Repo Volumes	\$0.685T	\$0.681T	−(0.004)
ℓ	Bank Loans	\$12.13T	\$12.32T	+(0.19)

Panel A of this table displays the values the model generates, under the internally calibrated parameters, for our targeted moments in comparison to the data. The average values from the data during the calibration period (June 1, 2022 to August 31, 2024) are displayed in the column "Data" and the model's generated moments are in the column "Model". "Error" displays the difference between the two columns. Panel B of this table displays predicted values generated by the model that we do not target in the calibration. The average values from the data during the calibration period are displayed in the column "Data" and the model's generated values are in the column "Model". "Error" displays the difference between the two columns. For both panels, variable descriptions and data sources are available in Table D.1.

Table 4. Out-of-Sample Model Performance: September 2024

Panel A. Parameter Changes			
Parameter	**Description**	**Source**	**Value**
R	IORB	Data	4.90%
r	ON RRP offering rate	Data	4.80%
$p^L b^{L,cb}$	Nominal quantity of bonds held by the Fed	Data	\$4.39T
$p^S B^S$	Nominal quantity of bills in the economy	Data	\$2.52T
m	Money endowment to households	Data	\$19.61T

Panel B. Model Performance – Targeted				
Parameter	**Description**	**Data**	**Model**	**Error**
i_{d^b}	Interest rate on bank deposits	1.35%	0.98%	$-(0.37)$
i_{d^m}	Interest rate on MMF deposits	4.72%	4.62%	$-(0.1)$
d^m	MMF Deposits	\$4.11T	\$4.04T	$-(0.07)$
d^b	Bank Deposits	\$15.44T	\$15.57T	$+(0.13)$
d^{RRP}	ON RRP usage	\$0.18T	\$0.15T	$-(0.03)$
m^r	Reserves	\$3.20T	\$3.25T	$+(0.05)$

Panel C. Model Performance – Untargeted				
Parameter	**Description**	**Data**	**Model**	**Error**
ρ	TGCR	4.83%	4.80%	$-(0.03)$
z^m	Repo volumes	\$1.20T	\$1.37T	$+(0.17)$
ℓ	Bank Loans	\$12.54T	\$12.32T	$-(0.22)$

Panel A of this table displays the model's independent parameters that are adjusted when running the out-of-sample test. These include changes both in policy (the IORB rate R, the ON RRP offering rate r, and the nominal quantity of bonds held by the Fed $p^L b^{L,cb}$) and in endowments in the environment (the nominal quantity of bills in the economy $p^S B^S$ and the money endowment to households m). Panel B of this table displays the values the model generates, under the internally calibrated parameters, for our targeted moments in comparison to the data. The average value from the data during the calibration period between our calibration period (June 1, 2022 to August 31, 2024) are displayed in "Data" and the model's generated moments are in the column "Model". "Error" displays the difference between the two columns. Panel C of this table displays predicted values generated by the model that we do not target in the calibration. The average value from the data during the calibration period are displayed in the column "Data" and the model's generated values are in the column "Model". "Error" displays the difference between the two columns. For both panels, variable descriptions and data sources are available in Table D.1.

Table 5. Out-of-Sample Model Performance: November 2024

Panel A. Parameter Changes			
Parameter	**Description**	**Source**	**Value**
R	IORB	Data	4.65%
r	ON RRP offering rate	Data	4.55%
$p^L b^{L,cb}$	Nominal quantity of bonds held by the Fed	Data	\$4.35T
$p^S B^S$	Nominal quantity of bills in the economy	Data	\$2.64T
m	Money endowment to households	Data	\$19.76T

Panel B. Model Performance – Targeted				
Parameter	**Description**	**Data**	**Model**	**Error**
i_{d^b}	Interest rate on bank deposits	1.32%	0.74%	−(0.58)
i_{d^m}	Interest rate on MMF deposits	4.48%	4.37%	−(0.11)
d^m	MMF Deposits	\$4.23T	\$4.18T	−(0.05)
d^b	Bank Deposits	\$15.47T	\$15.58T	+(0.11)
d^{RRP}	ON RRP usage	\$0.11T	\$0.12T	+(0.01)
m^r	Reserves	\$3.25T	\$3.27T	+(0.02)

Panel C. Model Performance – Untargeted				
Parameter	**Description**	**Data**	**Model**	**Error**
ρ	TGCR	4.58%	4.55%	−(0.03)
z^m	Repo volumes	\$1.34T	\$1.42T	+(0.08)
ℓ	Bank Loans	\$12.59T	\$12.32T	−(0.27)

Panel A of this table displays the model's independent parameters that are adjusted when running the out-of-sample test. These include changes both in policy (the IORB rate R, the ON RRP offering rate r, and the nominal quantity of bonds held by the Fed $p^L b^{L,cb}$) and in endowments in the environment (the nominal quantity of bills in the economy $p^S B^S$ and the money endowment to households m). Panel B of this table displays the values the model generates, under the internally calibrated parameters, for our targeted moments in comparison to the data. The average value from the data during the calibration period between our calibration period (June 1, 2022 to August 31, 2024) are displayed in "Data" and the model's generated moments are in the column "Model". "Error" displays the difference between the two columns. Panel C of this table displays predicted values generated by the model that we do not target in the calibration. The average value from the data during the calibration period are displayed in the column "Data" and the model's generated values are in the column "Model". "Error" displays the difference between the two columns. For both panels, variable descriptions and data sources are available in Table D.1.

Table 6. Out-of-Sample Model Performance: September 2019

Panel A. Parameter Changes			
Parameter	**Description**	**Source**	**Value**
R	IORB	Data	2.10%
r	ON RRP offering rate	Data	2.00%
$p^L b^{L,cb}$	Nominal quantity of bonds held by the Fed	Data	\$2.11T
$p^L B^L$	Nominal quantity of bonds in the economy	Data	\$2.66T
$p^S B^S$	Nominal quantity of bills in the economy	Data	\$0.89T
m	Money endowment to households	Data	\$12.70T

Panel B. Model Performance − Targeted				
Parameter	**Description**	**Data**	**Model**	**Error**
i_{d^b}	Interest rate on bank deposits	0.14%	0.00%	$-(0.14)$
i_{d^m}	Interest rate on MMF deposits	1.92%	4.42%	$+(2.50)$
d^m	MMF Deposits	\$1.57T	\$1.43T	$-(0.14)$
d^b	Bank Deposits	\$11.09T	\$11.26T	$+(0.17)$
d^{RRP}	ON RRP usage	\$0.0023T	\$0T	$-(0.0023)$
m^r	Reserves	\$1.47T	\$1.46T	$-(0.01)$

Panel C. Model Performance − Untargeted				
Parameter	**Description**	**Data**	**Model**	**Error**
ρ	TGCR	5.25%	4.61%	$-(0.64)$
z^m	Repo volumes	\$0.638T	\$0.548T	$-(0.09)$
ℓ	Bank Loans	\$9.96T	\$9.81T	$-(0.15)$

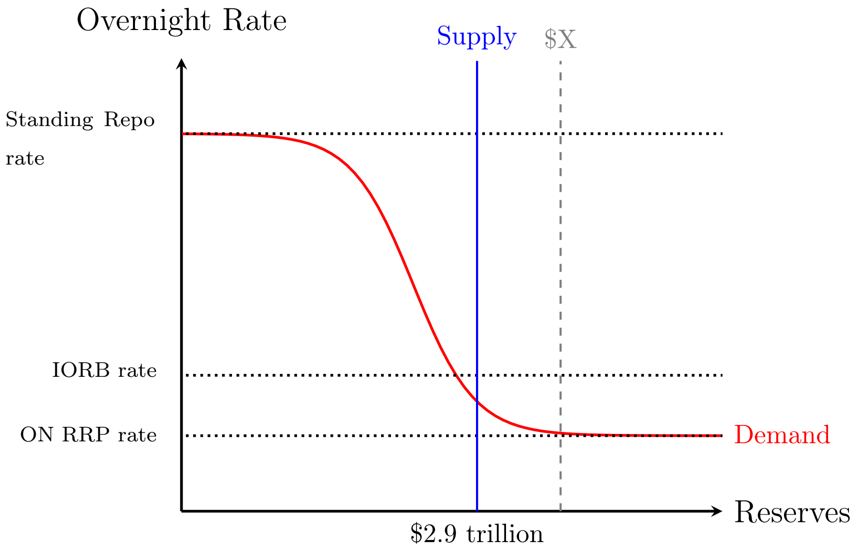
Panel A of this table displays the model's independent parameters that are adjusted when running the out-of-sample test. These include changes both in policy (the IORB rate R, the ON RRP offering rate r, and the nominal quantity of bonds held by the Fed $p^L b^{L,cb}$) and in endowments in the environment (the nominal quantity of bonds in the economy $p^S B^L$, the nominal quantity of bills in the economy $p^S B^S$ and the money endowment to households m). Panel B of this table displays the values the model generates, under the internally calibrated parameters, for our targeted moments in comparison to the data. The average value from the data during the calibration period between our calibration period (June 1, 2022 to August 31, 2024) are displayed in "Data" and the model's generated moments are in the column "Model". "Error" displays the difference between the two columns. Panel C of this table displays predicted values generated by the model that we do not target in the calibration. The average value from the data during the calibration period are displayed in the column "Data" and the model's generated values are in the column "Model". "Error" displays the difference between the two columns. For both panels, variable descriptions and data sources are available in Table D.1.

Table 7. How Different Reserve Levels Affect the Repo Rate

Reserves (m^r)	Repo spread ($\rho - R$)
\$3.31T	-10 bps
\$3.00T	-5 bps
\$2.69T	0 bps
\$2.39T	+5 bps
\$2.09T	+10 bps
\$1.79T	+15 bps
\$1.50T	+20 bps

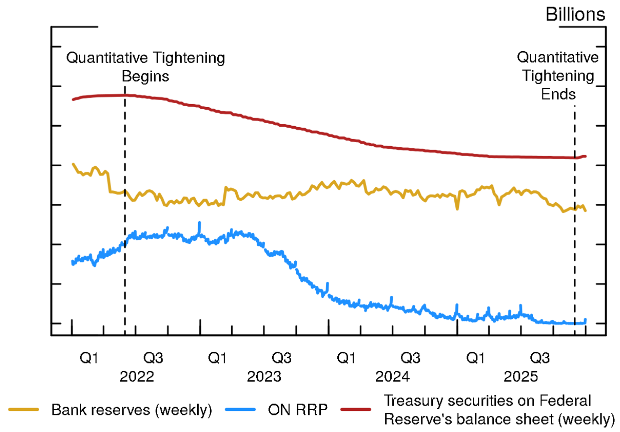
This table displays the calibrated model's predictions of the spread between the repo rate ρ and the IORB rate R, in basis points, associated with different levels of aggregate reserves m^r, in trillions. This table illustrates the continuum of equilibria demonstrated by the calibrated model, and that the level of m^r can be adjusted depending on the central bank's preferred spread of $\rho - R$. For example, if the desired the spread of $\rho - R$ is -10 basis points (i.e. ρ is to be equal to the ON RRP rate), then the model predicts $m^r = \$3.31$ trillion. If the desired spread of $\rho - R$ is zero, then the model predicts $m^r = \$2.69$ trillion.

Figure 1. The Demand Curve for Bank Reserves



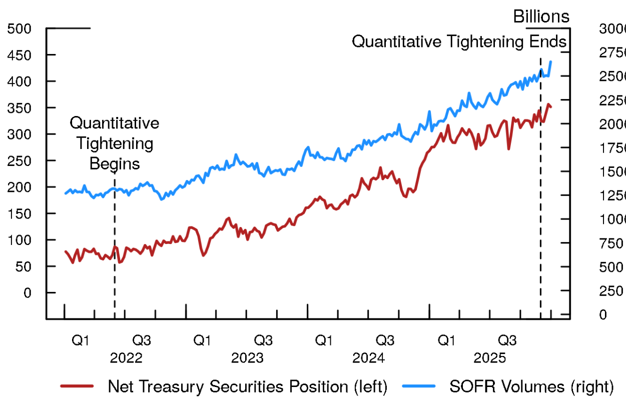
This figure shows the relationship between overnight rates and reserves held at the Federal Reserve. The red line denotes the reserve demand curve and the blue line denotes reserve supply as of April 2026 ($2.9 trillion). $X denotes the maximum level of reserves consistent with an ample reserves environment. To the right of $X is abundant reserves. The IORB rate is the interest rate on reserve balances at which banks can lend to the Fed. The ON RRP rate is the offering rate at the Overnight Reverse Repo Facility at which non-banks can lend to the Fed. The Standing Repo rate is the rate at which institutions can borrow from the Fed.

Figure 2. Evolution of the Fed's Balance Sheet



This figure displays selected items on the Federal Reserve's balance sheet: Treasury securities (red line), bank reserves (orange line), and non-bank reserves (investments at the ON RRP, blue line). Quantitative tightening began on June 1, 2022 and ended on December 1, 2025. Source: Federal Reserve Board H.4.1., Federal Reserve Bank of New York.

Figure 3. Dealer Positions of Treasury Securities & Repo Financing Volumes



This figure displays the net position of Treasury coupons and Treasury Inflation Protected Securities (red line) and overnight repurchase (repo) volumes (blue line) on primary dealers' balance sheets. Repo volumes are the sum of overnight and continuing repo positions against Treasury security collateral in the cleared bilateral, tri-party, General Collateralized Financing, and uncleared bilateral segments. Quantitative tightening began on June 1, 2022 and ended on December 1, 2025. Source: Primary Dealer Statistics from the Federal Reserve Bank of New York.

Figure 4. Auctioned Treasury Securities Financed in the Repo Market

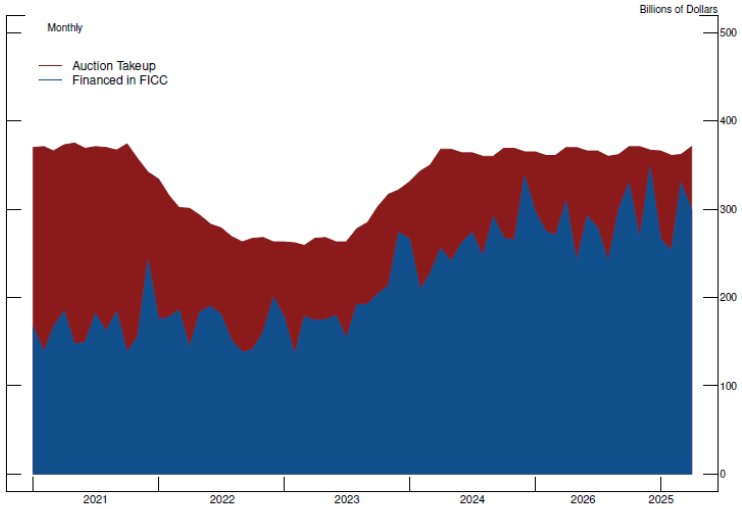


Figure 4 displays the monthly volume of auction-financed Treasury coupon securities in the Delivery-versus-Payment (DVP) repo market from 2021 to 2025, measured in billions of dollars. The blue area represents securities financed through the Fixed Income Clearing Corporation (FICC), while the red area shows direct auction takeup. The data indicates a clear shift in how these securities are being financed, with participants increasingly utilizing FICC clearing services rather than outright financing. Source: OFR Repo Collection, Repo Mapper Bostrom (2025), Treasury Direct.

Figure 5. Auctioned Treasury Bills Financed in the Repo Market

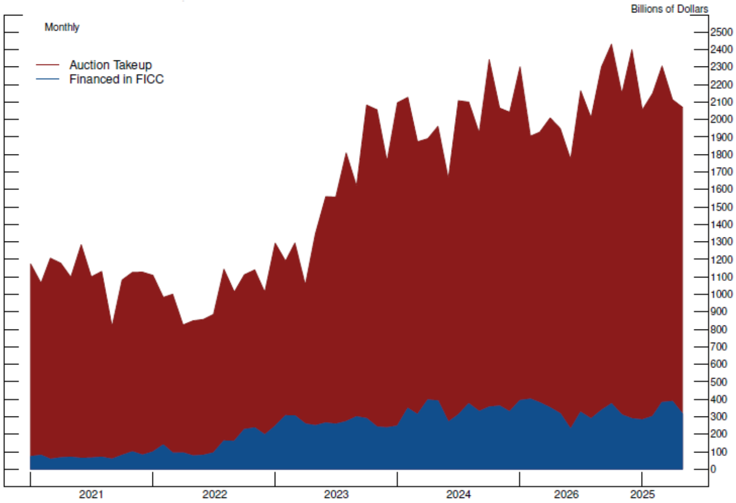


Figure 5 displays the monthly volume of auction-financed Treasury bills in the DVP repo market from 2021 to 2025, measured in billions of dollars. The blue area represents bills financed through FICC, while the red area shows direct auction takeup. The chart demonstrates that only a small fraction of Treasury bills are financed in the repo market through FICC. The dominant red area indicates that the vast majority of auctioned Treasury bills are taken up directly rather than being financed through the repo market. Source: Source: OFR Repo Collection, Repo Mapper Bostrom (2025), Treasury Direct.

Figure 6. Treasury Securities Financed in the Repo Market by Hedge Funds

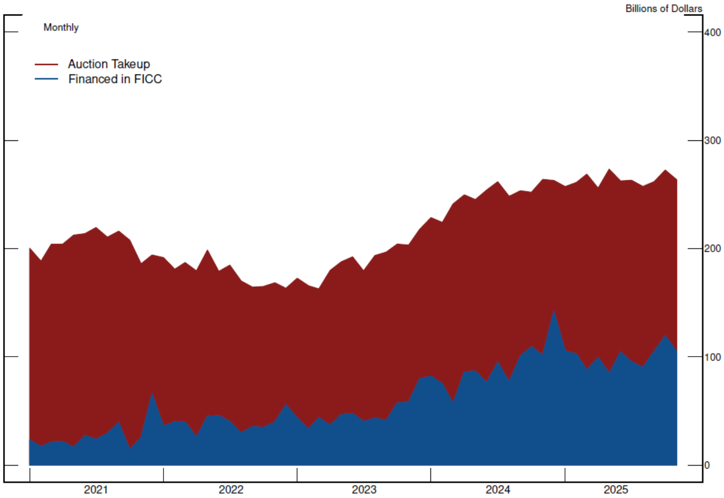
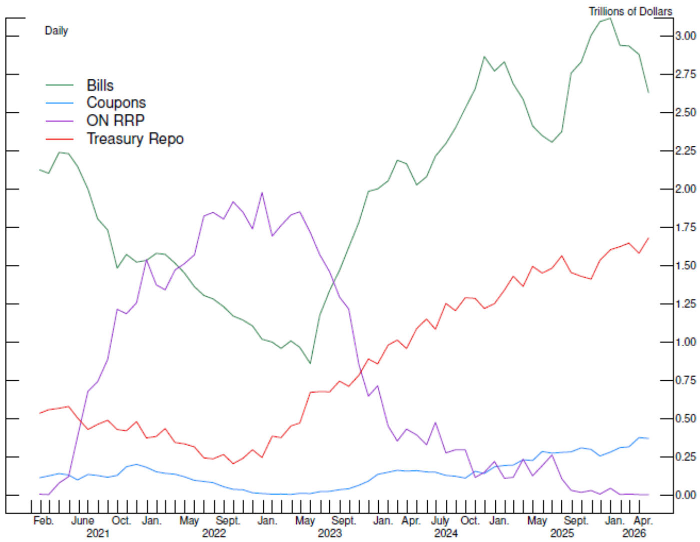


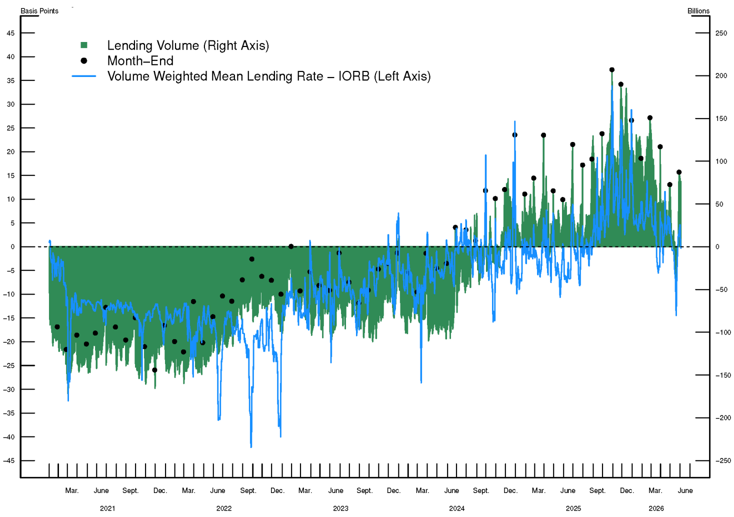
Figure 6 displays the monthly volume of hedge fund auction-financed Treasury securities in the DVP repo market from 2021 to 2025, measured in billions of dollars. The blue area represents securities financed through FICC, while the red area shows direct auction takeup by investment funds. This pattern provides empirical support for modeling levered investors as hedge fund-type participants who rely heavily on the repo market to finance their Treasury securities holdings: the data shows that hedge funds systematically finance their auction purchases through the repo market, with an increasing proportion utilizing centralized clearing through FICC, making them the natural real-world counterpart to the levered investors in our theoretical framework. Source: OFR Repo Collection, Repo Mapper Bostrom (2025), Treasury Direct.

Figure 7. Money Market Fund Holdings by Investment Type



This figure displays daily money market fund (MMF) holdings across four investment categories from February 2021 through July 2025, measured in trillions of dollars. Treasury bills (dark green) declined from over \$2.2 trillion in early 2021 to below \$1 trillion by mid-2023, before surging to nearly \$3 trillion by late 2024. The Overnight Reverse Repo (ON RRP) facility (purple) rose sharply from negligible levels in 2021 to peak around \$2 trillion in 2022-2023, then declined precipitously to near zero by 2024 as money funds shifted to higher-yielding alternatives. Treasury repo (red) exhibited steady growth from approximately \$0.5 trillion to \$1.5 trillion over the period, while Treasury coupon securities (blue) remained relatively stable at low levels around \$0.1-0.2 trillion throughout. Source: SEC N-MFP filings.

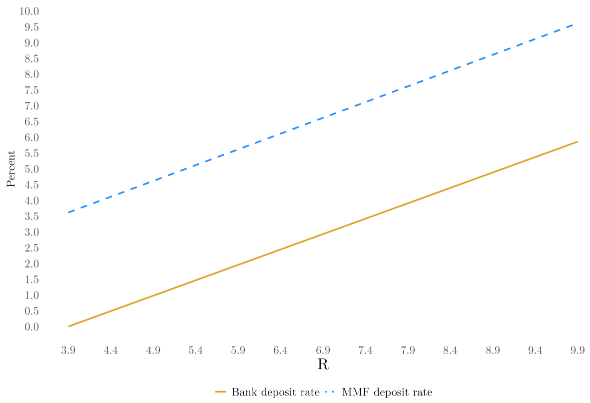
Figure 8. Bank Net Lending in Overnight Treasury Repo



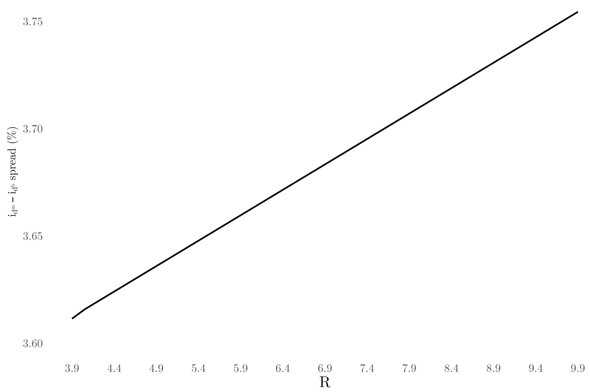
This figure shows bank net lending (lending minus borrowing) in the overnight Treasury repo market from January 2021 through March 2026. The green area (right axis) displays net lending volume in billions of dollars, with black dots marking month-end observations. The blue line (left axis) tracks the volume-weighted mean lending rate relative to the Interest on Reserve Balances (IORB) rate in basis points. Source: OFR Repo Collection, Repo Mapper Bostrom (2025).

Figure 9. Deposits and the Policy Rate

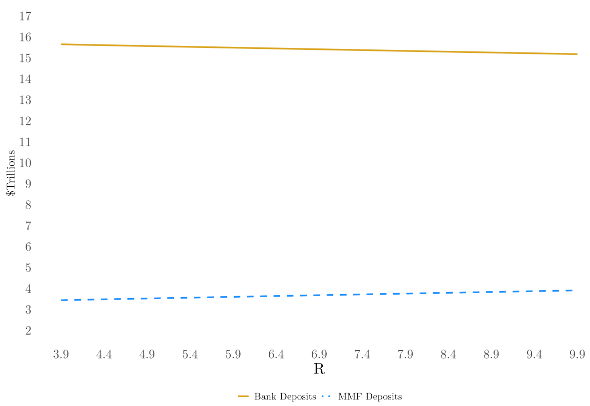
(a) Deposit Rate



(b) Deposit Spread

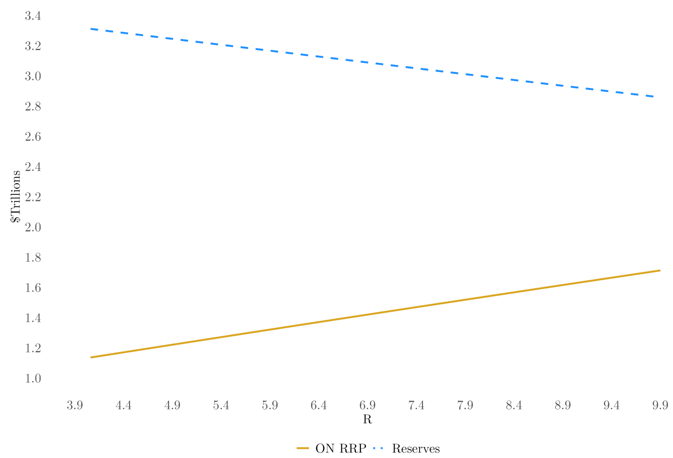


(c) Deposit Allocation



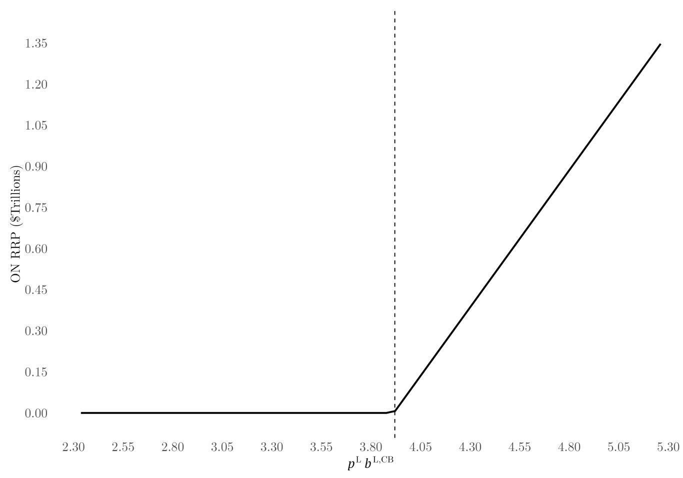
This figure shows how key deposit variables change with the policy rate R. The top panel (a) displays the deposit rate at banks i_{d^b} (the orange line) and the deposit rate at MMFs i_{d^m} (the blue dashed line). The middle panel (b) displays the spread between the MMF deposit rate and the bank deposit rate. The positive slope indicates that the pass-through of an increase in the policy rate R is larger for MMF deposit rates than it is for bank deposit rates. The bottom panel (c) displays the deposit allocation at banks d^b (the orange line) and MMFs d^m (the blue dashed line).

Figure 10. Reserves, ON RRP Investment, and the Policy Rate



This figure shows how reserve balances (the blue dashed line) and ON RRP investment (the solid orange line) change with the policy rate R, holding fixed the spread to r. As the policy rate R increases, i_{d^m} increases more than i_{d^b}. This causes households to allocate more deposits to MMFs and less deposits to the commercial bank. Holding constant the size of the central bank's balance sheet, the demand for funding in the repo market is unchanged, so MMFs place these extra deposits at the ON RRP facility. When the reserve constraint for the commercial bank is not binding, the bank continues to lend at the optimal loan amount, and the reduction in bank deposits is matched with a reduction in reserves

Figure 11. ON RRP Investment and the Federal Reserve's Balance Sheet



This figure shows how ON RRP investment changes with the size of the Federal Reserve's balance sheet $b^{L,CB}$. The dashed vertical line denotes $(p^L \times \widetilde{b^{CB}})$, the nominal critical threshold of the Federal Reserve's balance sheet where the repo market switches from "abundant liquidity" (right of the line) to "ample liquidity" (left of the line). In abundant liquidity equilibria, ON RRP investment declines with the size of the balance sheet until it reaches zero, whereas in ample liquidity equilibria, ON RRP investment is constant at zero.

Figure 12. The Repo Rate and the Federal Reserve's Balance Sheet

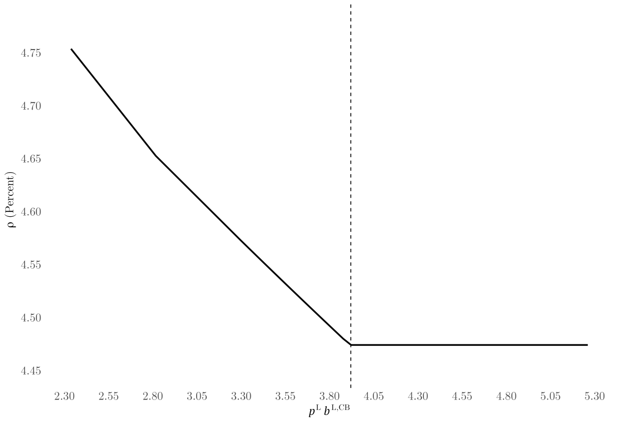


Figure 13. Deposits and the Federal Reserve's Balance Sheet

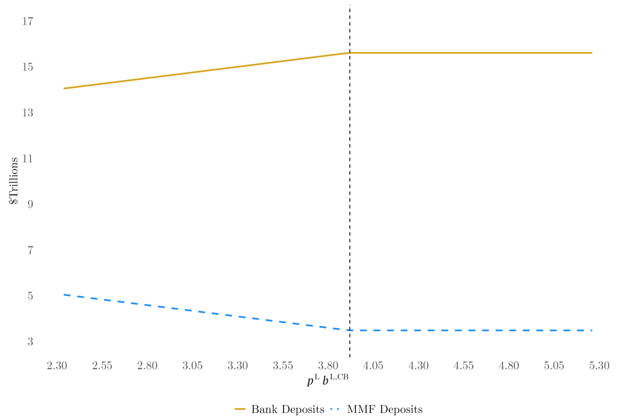


Figure 14. Reserves and the Federal Reserve's Balance Sheet

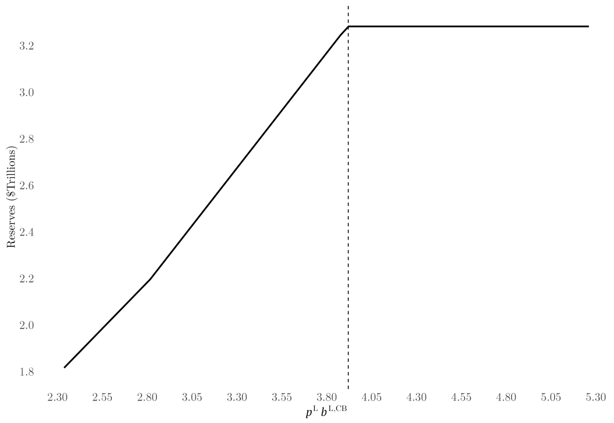
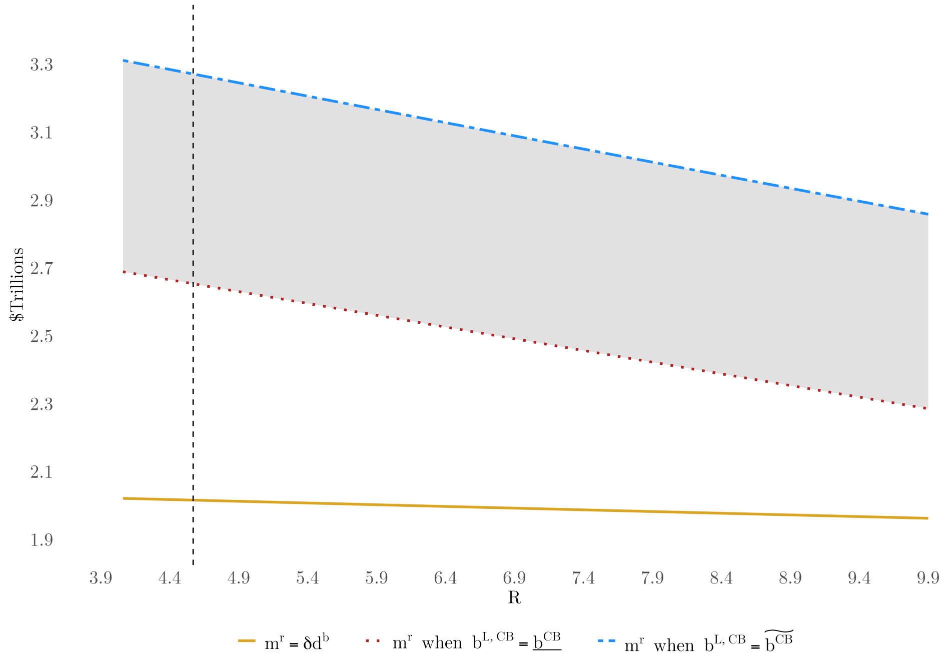


Figure 12 shows how the repo rate ρ changes with the size of the Federal Reserve's balance sheet $b^{L,CB}$. The dashed vertical line denotes $(p^L \times \widetilde{b^{CB}})$, the nominal critical threshold of the Federal Reserve's balance sheet where the repo market switches from "abundant liquidity" (right of the line) to "ample liquidity" (left of the line). In abundant liquidity equilibria, the repo rate ρ equals the ON RRP rate, whereas in ample liquidity equilibria, the repo rate increases as the size of the Federal Reserve's balance sheet declines. Figure 13 shows how deposit allocation changes with the size of the Fed's balance sheet $b^{L,CB}$. In abundant liquidity equilibria, deposits are unchanged with the size of the Fed's balance sheet, whereas in ample equilibria, bank deposits decrease and MMF deposits increase as the Federal Reserve's balance sheet decreases, given the higher rates that MMFs offer relative to bank deposit rates. Figure 14 shows how reserves change with the size of the Fed's balance sheet $b^{L,CB}$. In abundant liquidity equilibria, reserves are unchanged with the size of the Fed's balance sheet, whereas in ample equilibria, reserves decrease as the Federal Reserve's balance sheet decreases, due to the decreasing allocation of bank deposits from the household.

Figure 15. Repo Market Capacity vs. Bank Reserve Demand



This figure shows how the level of reserves consistent with the model thresholds change with the policy rate R. The dashed blue line is the threshold between abundant and ample liquidity, $\widetilde{b^{CB}}$, where $\rho = r$ and $d^{RRP} = 0$. The dotted red line is the threshold between ample and scarce liquidity, $\underline{b^{CB}}$, where $\rho = R$. The shaded region represents ample liquidity in the repo market. The solid orange line is the threshold where reserves are equal to the minimum level of reserve demand. This figure illustrates different types of equilibria when the policy rate $R \in [4.0\%, 9.5\%]$. Starting from a very large balance sheet, if the Fed's government securities holdings are larger than the critical threshold $\widetilde{b^{CB}}$, there is abundant liquidity in the repo market and excess reserves (above the dashed blue line). As the Fed reduces its balance sheet beyond $\widetilde{b^{CB}}$, the repo rate starts to increase and the repo market has ample liquidity, but banks still hold excess reserves (between the dashed blue line and the solid orange line). The dotted red line demonstrates that within this set of equilibria, reserves can be lower if the Fed is willing to tolerate higher repo rates. The repo rate equals IORB at the dotted red line; below this level, liquidity in the repo market is scarce. If the Fed reduces its balance sheet further, such that the reserve constraint for banks becomes binding (at the solid orange line), there is both scarce liquidity in the repo market and scarce reserves, as banks only hold the minimum level of reserve demand.

Internet Appendix

A Government, Central Bank, Households and Firms

In this appendix, we will provide the consolidated government budget constraint (A.1), and the maximization problems of firms (A.2) and households (A.3).

A.1 Government and Central Bank

In the model, we assume that the government has to satisfy a government budget constraint and can impose lump-sum taxes on households to balance its budget.

Households are assumed to be the owners of the monopoly bank, the dealers, MMFs, and the investors. All profits of these agents are paid out as dividends to households at the beginning of period $t = 2$. Households then pay lump-sum taxes and use the remaining funds to consume the consumption good in period $t = 2$. We assume that ownership of dealers, MMFs, investors and the monopoly bank is distributed uniformly such that each household receives the average profits as dividends.

Here, we derive the consolidated government budget constraint, starting with the balance sheet of the central bank. The central bank holds government bonds as assets, $b^{L,CB}$, and reserves, m^r_{CB}, ON RRP balances, d^{RRP}_{CB}, and equity, e_{CB}, as liabilities:

$$b^{L,CB} = m^r_{CB} + d^{RRP}_{CB} + e_{CB}. \tag{A.1}$$

The central bank's balance sheet identity must hold in equilibrium. Any changes in government bond holdings will be fully reflected in the change of the sum of reserves, ON RRP balances and central bank equity.

The central bank purchases government bonds, $b^{L,CB}$, at the market clearing price p^L and it pays interest R and r on reserves and ON RRP balances, respectively, at the end of period $t = 1$. Consequently, the surplus of the central bank satisfies

$$S^{CB} = b^{L,CB}(1 - p^L) - Rm^r_{CB} - rd^{RRP}_{CB}$$

By definition, reserve balances of banks are equal to reserve balances held at the central bank and ON RRP balances of MMFs are equal to ON RRP balances at the central bank, that is $m^r = m^r_{CB}$ and $d^{RRP} = d^{RRP}_{CB}$.

The government issues government bonds, B^L and bills B^S respectively, which yield one unit of fiat money in period $t = 2$. It uses the money from the sale of government bonds to finance a public good, denoted g, and distributes an initial endowment m to all households. In order to balance the consolidated government budget constraint, it levies lump-sum taxes or transfers lump-sum subsidies to households in period $t = 2$. Moreover, it receives any surplus or deficit from the central bank. Thus, the consolidated government budget constraint is:

$$(1 - p^L)B^L + (1 - p^S)B^S + g + m = \tau_h + S^{CB},$$

where τ_h denotes lump-sum taxes or subsidies to households.

Combining both equations yields the consolidated government budget constraint

$$(1 - p^L)(b^{L,d} + b^{L,i}) + (1 - p^S)B^S + g + m + Rm^r_{CB} + rd^{RRP}_{CB} = \tau_h \tag{A.2}$$

A.2 Firms

Firms produce a perishable, non-storable consumption good in period $t = 1$ and period $t = 2$. They can hold their earnings in bank deposits, earning a deposit rate i_{d^b}, and receive utility from money holdings in period $t = 2$. This implies that firms are always willing to produce consumption goods in return for money.[24] Their preferences are

$$\mathbb{E}\left[p_1 q_1 (1 + i_{d^b}) + p_2 q_2\right]. \tag{A.3}$$

Finally, we assume firms only accept bank deposits as payment in period $t = 1$.[25]

In the goods markets in periods $t = 1$ and $t = 2$, firms produce a perishable consumption good, q^s_t, at linear costs. Firms deposit their $t = 1$ earnings in their bank account until period $t = 2$, earning interest on these deposits. The problem of the firm is then to solve:

$$\max_{q^s_1, q^s_2} \; - q^s_1 - q^s_2 + (\phi p_1 q^s_1 (1 + i_{d^b}) + \phi p_2 q^s_2) \tag{A.4}$$

where ϕ is value of money in terms of the consumption good in period $t = 2$ and thus

[24] Without such an assumption, firms would not be willing to produce goods in period $t = 2$. This is because the economy ends after period $t = 2$, and firms would not be able to spend their money holdings. Thus by backward induction, there cannot be a monetary equilibrium where fiat money is valued. The assumption that firms receive utility from money holdings serves as a tool to ensure a monetary equilibrium.

[25] The assumption that only bank deposits are accepted as payment is used to generate demand for bank deposits even if MMF deposits dominate bank deposits in terms of their return, $i_{d^b} < i_{d^m}$, as is often observed in reality.

$1/\ p_2 = \phi$. The first-order conditions satisfy:

$$q_1^s \quad : \quad -1 + \phi p_1(1 + i_{d^b}) = 0, \tag{A.5}$$
$$q_2^s \quad : \quad -1 + \phi p_2 = 0. \tag{A.6}$$

Firms are willing to produce any quantity if the price compensates them for the marginal production cost and, in the case of period $t = 1$, for the cost of holding bank deposits across periods. Thus, if Equations (A.5) and (A.6) hold, firms produce goods to meet the demand.

A.3 Households

Households receive a lump-sum endowment from the central bank, denoted m, at the beginning of period $t = 1$. Households can allocate their endowment between bank deposits (d_h^b) and MMF deposits (d_h^m) in $t = 1$. We assume that the market for MMF deposits is perfectly competitive and MMFs pay a market clearing interest rate on their deposits. The bank deposit market is monopolistic with a monopoly bank demanding bank deposits.[26] Households want to consume a consumption good in periods $t = 1$ and $t = 2$. Their preferences satisfy

$$\mathbb{E}\left[u(q_1) + q_2\right]. \tag{A.7}$$

To generate positive demand for bank deposits even when bank deposit rates are lower than MMF deposit rates, we assume that only bank deposits can be used to purchase the consumption good in period $t = 1$. While in reality households can withdraw funds from MMFs to make payments, this simplifying assumption introduces imperfect substitutability between the two assets, in line with d'Avernas and Vandeweyer (2024) and d'Avernas et al. (2025). The key result is that households endogenously allocate funds between MMFs and bank deposits and but an incentive to hold some bank deposits, even when MMFs deposits dominate in return.

Additionally, since we assume that MMFs, households, the monopoly bank, and investors are owned by households, all profits $\sum_{j=1}^{4} \pi_j$ for $j = \{dealers, MMF, investors, bank\}$, are paid out to households at the end of period $t = 2$ and households pay lump-sum taxes or subsidies τ_h.

[26]The key ingredient in the bank deposit market is that banks have bargaining power and can therefore extract some rent from bank depositors. For simplicity, we model it as a monopolistic market.

The households' maximization problem is then to solve:

$$\max_{q_1, q_2, d_h^b, d_h^m} u(q_1) + q_2 \tag{A.8}$$

$$\text{s.t. } d_h^b - p_1 q_1 \geq 0,$$

$$(d_h^b - p_1 q_1)(1 + i_{d^b}) + d_h^m (1 + i_{d^m}) - p_2 q_2 + \left(\sum_{j=1}^{4} \pi_j - \tau_h \right) \geq 0,$$

$$m - d_h^b - d_h^m \geq 0.$$

The first constraint implies that households can only use bank deposits to purchase the consumption good in period $t = 1$ and has the Lagrange multiplier $\lambda_{q,1}$. The second constraint implies that households cannot spend more than their funds available in $t = 2$ on the consumption good and has the Lagrange multiplier $\lambda_{q,2}$. The third constraint implies that households cannot allocate more funds than their endowment to bank deposits and MMF deposits and has the Lagrange multiplier λ_h. The first order conditions satisfy:

$$q_1 : \quad u'(q_1) - p_1 \lambda_{q1} - p_1 (1 + i_{d^b}) \lambda_{q2} = 0, \tag{A.9}$$

$$q_2 : \quad 1 - p_2 \lambda_{q2} = 0, \tag{A.10}$$

$$d_h^b : \quad \lambda_{q1} + \lambda_{q2} (1 + i_{d^b}) - \lambda_h = 0, \tag{A.11}$$

$$d_h^m : \quad \lambda_{q2} (1 + i_{d^m}) - \lambda_h = 0. \tag{A.12}$$

Combining Equations (A.9) - (A.12) and using Equation (A.5) and $1/p_2 = \phi$ yields

$$u'(q_1) = \left(\frac{1 + i_{d^m}}{1 + i_{d^b}} \right) \tag{A.13}$$

Thus consumption will depend on the spread between the MMF and the bank deposit rates. We now discuss the optimal choice of allocation for households under different assumptions for i_{d^m} and i_{d^b}. First, note that $i_{d^m} < i_{d^b}$ implies $d_h^m = 0$, which cannot be an equilibrium outcome. From the dealers maximization problem, there will always be a positive demand for repo borrowing in the repo market in period $t = 1$ for $B^L > b^{L,CB}$. Since MMFs are the lenders in the repo market, the MMF deposit rate will increase in equilibrium until MMFs can attract some positive amount of deposits, which can then be lent out in the repo market, deposited at the ON RRP, or used to purchase bills.

Thus, we have two possible cases: $i_{d^m} > i_{d^b}$ and $i_{d^m} = i_{d^b}$

Case 1: $i_{d^m} > i_{d^b}$ In the data, we observe that $i_{d^m} > i_{d^b}$. Since bank deposits can be used as a means of payment unlike MMF deposits, banks can set lower deposit rates and still attract some deposits, consistent with the bank deposit channel (Bernanke and Blinder, 1988; Drechsler et al., 2017; Begenau and Stafford, 2025).

For $i_{d^m} > i_{d^b}$, households spend all of their deposit holdings on the consumption good in $t = 1$ and use MMF deposits to consume in period $t = 2$. Thus $d_h^b = p_1 q_1$ and

$$u'(d_h^b/p_1) = \left(\frac{1 + i_{d^m}}{1 + i_{d^b}}\right). \tag{A.14}$$

Consequently, MMF deposits satisfy:

$$d_h^m = m - p_1 u'^{-1}\left(\frac{1 + i_{d^m}}{1 + i_{d^b}}\right). \tag{A.15}$$

Case 2: $i_{d^m} = i_{d^b}$ First, note that in this case $u'(q_1) = 1$, which means that households consume the first-best quantity in period $t = 1$ (since the firms have linear production costs). In that case, households have at least the following quantity as deposits

$$d_h^b = p_1 u'^{-1}(1). \tag{A.16}$$

For the remaining funds, households are indifferent between allocating them between banks or MMFs. However, as mentioned above, $d^m = 0$ cannot be an equilibrium outcome. Therefore, we can assume without loss of generality, that households allocate the remaining funds to MMF deposits and therefore

$$d_h^m = m - p_1 u'^{-1}(1). \tag{A.17}$$

Finally, in period $t = 2$, demand for the consumption good satisfies

$$q_2 = \frac{d_h^m(1 + i_{d^m}) + \left(\sum_{j=1}^{4} \pi_j - \tau_h\right)}{p_2}, \tag{A.18}$$

where π_j are the profits that are paid out to households and τ_h are lump-sum taxes or subsidies.

In the main text, we focus our analysis on case 1, $i_{d^m} > i_{d^b}$, since this is consistent with the empirical observation that MMFs typically offer higher returns than banks.

B Proofs

Proof of Lemma 2. First, note that by bond market clearing $b^{CB} = B - b^p$. for any fixed $B > 0$, it suffices to show the existence and properties of a critical threshold for the private sector bond holdings $\widetilde{b^p}$, since then this immediately implies the existence of a critical threshold for the central bank bond holdings $\widetilde{b^{CB}} = B - \widetilde{b^p}$ which inherits the properties of the $\widetilde{b^p}$ threshold with reversed signs.

From the MMFs optimization problem, we know that the first-order conditions are

$$\rho \leq \lambda_m \quad \text{(equality if } z^m > 0) \tag{A.19}$$

$$r \leq \lambda_m \quad \text{(equality if } d^{RRP} > 0) \tag{A.20}$$

$$\frac{1 - p^S}{p^S} \leq \lambda_m \quad \text{(equality if } b^S > 0) \tag{A.21}$$

where λ_m is the Lagrange multiplier on the budget constraint in period 1.

The bill supply $B^S > 0$ is exogenous, and by market clearing we must have $b^S = B^S > 0$ in equilibrium. Therefore the first-order condition with respect to b^S holds with equality:

$$\frac{1 - p^S}{p^S} = \lambda_m.$$

Next, in the Repo Market in $t = 1$, market clearing requires

$$z^m = \frac{b^{L,d} + b^{L,i}}{1 + \rho}.$$

Dealers demand is exogenous since they must finance $(b^{L,d} + b^{L,i}) = B^L - b^{L,CB} > 0$ chosen by the central bank. Then immediately we have that dealer demand is $z^d > 0$ always. By market clearing we have that $z^d = z^m > 0$, and the first order condition with respect to z^m holds with equality:

$$\rho = \lambda_m.$$

Then, given the above two facts, we have that

$$\rho = \frac{1 - p^S}{p^S} \quad \Longleftrightarrow \quad p^S = \frac{1}{1 + \rho}$$

i.e. the return on repo lending is equated with the return on bills.

Suppose that $\rho = r$. Then from the above we have that

$$p^S = \frac{1}{1+r}.$$

From the MMF budget constraint, and combining with market clearing, we have that

$$\begin{aligned} d^m &= z^m + d^{RRP} + p^S b^S \\ &= \frac{b^{L,d} + b^{L,i}}{1+r} + d^{RRP} + \frac{1}{1+r} B^S \end{aligned}$$

which gives that ON RRP take-up is

$$d^{RRP} = d^m - \frac{b^{L,d} + b^{L,i} + B^S}{1+r}.$$

By assumption there is a non-negativity constraint on ON RRP take-up, $d^{RRP} \geq 0$, which from above implies that

$$\begin{aligned} d^m &\geq \frac{b^{L,d} + b^{L,i} + B^S}{1+r} \\ (1+r)d^m &\geq b^{L,d} + b^{L,i} + B^S \\ b^{L,d} + b^{L,i} &\leq (1+r)d^m - B^S \\ b^{L,d} + b^{L,i} &\leq (1+r)\left(d^m - \frac{B^S}{1+r}\right). \end{aligned}$$

Then since $p^S = \frac{1}{1+r}$ when $\rho = r$, this becomes

$$b^{L,d} + b^{L,i} \leq (1+r)\left(d^m - p^S B^S\right).$$

Next, let the total quantity of bonds held by the private sector be $b^p = (b^{L,d} + b^{L,i})$. Define the threshold value $\widetilde{b^p}$ as the $b^p \in (0, B^L)$ where

$$\widetilde{b^p} \equiv (1+r)\left(\widetilde{d^m} - p^S B^S\right)$$

where $\widetilde{d^m}$ is the d^m generated at the threshold equilibrium where $\rho = r$ and $d^{RRP} = 0$.

Note that at the equilibrium threshold $\widetilde{d^m}$ where $\rho = r$ and $d^{RRP} = 0$, from the MMFs zero-profit condition equation (22) combined with the MMFs balance sheet constraint, the zero-profit condition reduces to

$$i_{d^m} = r - k^m$$

Call this threshold MMF interest rate $\widetilde{i_{d^m}}$.

Claim 1. *If $b^p > \widetilde{b^p}$, then $d^{RRP} = 0$ and $\rho > r$.*

Proof. By contradiction.

First, suppose $b^p > \widetilde{b^p}$, but $d^{RRP} > 0$. By the MMFs first-order conditions, $d^{RRP} > 0$ and $z^m > 0$ give that $\rho = r$. By the definition, then d^m must equal $\widetilde{d^m}$. Then, by the MMFs budget constraint with market clearing imposed, we have that

$$(1+r)\widetilde{d^m} = b^p + B^s + (1+r)d^{RRP} > b^p + B^s.$$

Then, by the definition of $\widetilde{b^P}$, we have that

$$(1+r)\widetilde{d^m} = \widetilde{b^p} + B^S.$$

Combining the above two gives that

$$\begin{aligned} \widetilde{b^p} + B^S &> b^p + B^s \\ \iff \quad \widetilde{b^p} &> b^p \end{aligned}$$

which is an immediate contradiction of the initial assumption that $b^p > \widetilde{b^p}$. Therefore when $b^p > \widetilde{b^p}$ it must be that $d^{RRP} = 0$.

Second, suppose that $b^p > \widetilde{b^p}$ and $d^{RRP} = 0$, but $\rho = r$. By the definition, then d^m must equal $\widetilde{d^m}$. Then, by the MMFs budget constraint with market clearing imposed, we have that

$$(1+r)\widetilde{d^m} = b^p + B^s$$

Then, by the definition of $\widetilde{b^p}$, we have that

$$(1+r)\widetilde{d^m} = \widetilde{b^p} + B^S.$$

Combining the above two gives that

$$\widetilde{b^p} + B^S = b^p + B^s$$
$$\iff \quad \widetilde{b^p} = b^p$$

which is an immediate contradiction of the initial assumption that $b^p > \widetilde{b^p}$. Therefore when $b^p > \widetilde{b^p}$ and $d^{RRP} = 0$, it must be that $\rho > r$. ■

Claim 2. *If $b^p < \widetilde{b^p}$, then $d^{RRP} > 0$ and $\rho = r$.*

Proof. By contradiction.

Suppose that $b^p < \widetilde{b^p}$ but $d^{RRP} = 0$. By the definition of $\widetilde{b^p}$, we have that

$$\frac{b^p + B^S}{1+r} < \frac{\widetilde{b^p} + B^S}{1+r} = \widetilde{d^m}.$$

Since the repo market is active, $z^m > 0$, we know that $\rho = \lambda_m$. Since $d^{RRP} = 0$ by assumption, we know that $r - \lambda_m \leq 0$. Hence:

$$\rho = \lambda_m \geq r.$$

Since $b^S = B^S > 0$ by market clearing, from the MMFs first order conditions we know that $p^S = 1/(1+\rho)$. Next, given repo market clearing and $d^{RRP} = 0$, the MMF balance sheet constraint can be written as

$$d^m = \frac{b^p + B^S}{1+\rho}.$$

Since $\rho = \lambda_m \geq r$, $i_{d^m} = \rho - k^m$, and d^m is monotone in the rate, we have that

$$d^m \geq \widetilde{d^m}.$$

However, we now have that

$$d^m = \frac{b^p + B^S}{1+\rho} \leq \frac{b^p + B^S}{1+r} < \frac{\widetilde{b^p} + B^S}{1+r} = \widetilde{d^m}.$$

Together, this is

$$d^m \geq \widetilde{d^m} \quad \text{and} \quad d^m < \widetilde{d^m}$$

a contradiction. Therefore, when $b^p < \widetilde{b^p}$ we must have $d^{RRP} > 0$. Finally, given this, when $z^m > 0$, by the MMFs first order conditions we have that $\rho = r$. ■

Therefore, all together, we have the threshold $\widetilde{b^p}$ defined as

$$\widetilde{b^p} \equiv (1+r)\left(\widetilde{d^m} - p^S B^S\right)$$

where $\widetilde{d^m}$ is the d^m generated at the threshold equilibrium where $\rho = r$ and $d^{RRP} = 0$, and where the threshold MMF interest rate is $\widetilde{i_{d^m}} = r - k^m$. We have shown that there are three cases:

1. When $b^p < \widetilde{b^p}$, $\rho = r$ and $d^{RRP} > 0$.
2. When $b^p = \widetilde{b^p}$, $\rho = r$ and $d^{RRP} = 0$.
3. When $b^p > \widetilde{b^p}$, $\rho > r$ and $d^{RRP} = 0$.

Finally, by the above conclusion and bond market clearing, we have that there exists a critical threshold for the central bank bond holdings $\widetilde{b^{CB}} = B - \widetilde{b^p}$ which then inherits the properties of $\widetilde{b^p}$ with reversed signs, i.e. where

1. When $b^{CB} > \widetilde{b^{CB}}$, $\rho = r$ and $d^{RRP} > 0$.
2. When $b^{CB} = \widetilde{b^{CB}}$, $\rho = r$ and $d^{RRP} = 0$.
3. When $b^{CB} < \widetilde{b^{CB}}$, $\rho > r$ and $d^{RRP} = 0$.

and thus concludes the proof.

■

Proof of Proposition 3. To show that the thresholds $\widetilde{b^{CB}}$ is decreasing in policy rates R, r for a constant spread $s = R - r$, first note that it suffices to show that d^m is increasing in policy rates, since

$$\mathrm{d}\widetilde{b^{CB}} = -\mathrm{d}d^m(1+r) - d^m \mathrm{d}r.$$

Thus, for

$$\frac{\mathrm{d}d^m}{\mathrm{d}r} > 0$$

$\widetilde{b^{CB}}$ is decreasing in policy rates.

Using the micro-founded solutions for d^m in Appendix (A.3), we can furthermore show that $m_h = d^b + d^m$ must hold. Thus, in order to show that d^m is increasing in policy rates, it suffices to show that bank deposits are decreasing in policy rates.

From A.2 and A.3, $d^b = p_1 q$. Using Equations (A.5) and (A.13) and taking the total differential of $p_1 q$ yields

$$\frac{\mathrm{d}d^b}{\mathrm{d}r} = \frac{\mathrm{d}i_{d^m}}{\mathrm{d}r}\left(\frac{p_1}{u''(q)}\frac{1}{(1+i_{d^b})}\right) - \frac{\mathrm{d}i_{d^b}}{\mathrm{d}r}\left(\frac{p_1}{u''(q)}\frac{(1+i_{d^m})}{(1+i_{d^b})^2} + \frac{p_1 q}{(1+i_{d^b})}\right).$$

Rearranging yields

$$\frac{\mathrm{d}d^b}{\mathrm{d}r} = \frac{p}{1+i_{d^b}}\left[\frac{1}{u''(q)}\frac{\mathrm{d}i_{d^m}}{\mathrm{d}r} - \frac{\mathrm{d}i_{d^b}}{\mathrm{d}r}\left(\frac{u'(q)}{u''(q)} + q\right)\right] < 0.$$

Since, $p/(1+i_{d^b}) > 0$, the second term on the RHS must be negative for the inequality to hold. Rearranging this term yields

$$\frac{\mathrm{d}i_{d^m}/\mathrm{d}r}{\mathrm{d}i_{d^b}/\mathrm{d}r} > u'(q) + u''(q)\cdot q. \tag{25}$$

Thus, for $\frac{\mathrm{d}d^b}{\mathrm{d}r} < 0$ to hold, Equation (25) must hold.

Similarly, to show that $\underline{b^{CB}}$ is decreasing in policy rates R, r, it suffices to show that

$$\mathrm{d}\underline{b^{CB}} = -\mathrm{d}d^m(1 + r + \bar{\varepsilon}) - d^m \mathrm{d}\bar{r} < 0,$$

which follows directly under Equation (25).

■

Proof of Proposition 4. In order to show how the thresholds for banks reserve demand and the repo market liquidity compare to each other, first note that for any monetary policy (b^{CB}, R, r), we can determine the bank's and the MMFs' liquidity holdings. Thus, to compare the two thresholds, we can simply compare the size of the bank's balance sheet at the bank's reserve demand threshold for scarce reserves to the size of the bank's balance sheet at the repo market threshold for scarce liquidity, $b^{CB} = \underline{b^{CB}}$.

Consider first the bank's reserve demand. Reserves become scarce, once the bank can no longer invest $\ell = \ell^* = x'^{-1}(i_\ell - R)$ and still hold $m^r = \delta d^b$ reserves. Reserves at the threshold between ample and scarce reserves satisfy $m^r = \delta d^b$. Using the balance sheet constraint, we can write bank deposits $\tilde{d^b}$ at this threshold

$$\tilde{d^b} = \frac{\chi'^{-1}(i_\ell - R)}{1-\delta}.$$

Using the micro-foundations in Appendices A.3 and A.2, bank deposits satisfy

$$d^b = p_1 u'^{-1}\left(\frac{1+i_{d^m}}{1+i_{d^b}}\right).$$

Define p_1', and i'_{d^b} as the price of the consumption good in $t=1$ and bank deposit rates when $R = R'$, $r = r' = R' - s$ and $\rho = r + \bar{\varepsilon}$. Note further than when $b^{CB} = \underline{b^{CB}}$, $i_{d^m} = r + \bar{\varepsilon} - k^m$. Thus, there exists a critical threshold R' for which the bank's demand for reserves and the repo market capacity correspond to equal amounts of the central bank's balance sheet size:

$$\frac{\chi'^{-1}(i_\ell - R')}{1-\delta} = p_1' u'^{-1}\left(\frac{1+r+\bar{\varepsilon}-k^m}{1+i'_{d^b}}\right),$$

which is Equation (26) and where the RHS is bank deposits at $R = R'$, $r = R' - s$ and $\rho = r + \bar{\varepsilon}$, using the micro-foundations derived in Appendices A.3 and A.2.

For a constant spread $i_\ell - R$, the LHS is constant in R, whereas the RHS is decreasing in R under Proposition 3. Thus, there exists a unique R' for which the thresholds for scarce reserves and scarce liquidity are identical in terms of the size of the central bank's balance sheet. Thus, for $R < R'$, the repo market capacity is the binding constraint for the central bank's balance sheet.

Note further that if we were to relax our assumption that $i_\ell - R$ is constant, we could still get this result under the following conditions:

$i_\ell - R$ is increasing in R. Assume for now that the spread $i_\ell - R$ is increasing in R, which would imply that the LHS of Equation is increasing in R. Since the RHS is decreasing in R under Proposition 3, that would still imply that there is a R' for which the two thresholds are identical and therefore would also imply that for $R < R'$, the repo market liquidity constraint is the binding constraint for the central bank.

$i_\ell - R$ is decreasing in R. First, we rearrange Equation (26):

$$\frac{1}{1-\delta} = \frac{1}{\chi'^{-1}(i_\ell - R')} p_1' u'^{-1}\left(\frac{1+i'_{d^m}}{1+i'_{d^b}}\right),$$

The LHS is again constant in R, whereas the RHS is not. The first term of RHS is now decreasing in R and the second term is increasing in R. Thus, we would still have a level of R for which the two thresholds are the same as long as the decline in the first term of the RHS is not equal to the change in the second term of the RHS. Intuitively, we would still get a threshold R' as long as the change in loans does not exactly offset the change in bank deposits when R changes.

■

C Calibration and Simulations

C.1 Functional Form Assumptions

We make the following functional form assumptions. We assume that: 1) the household's utility over consumption in period one is CRRA with parameter α,

$$u(q_1) = \frac{q_1^{1-\alpha}}{1-\alpha},$$

2) the banks loan monitoring cost is quadratic with parameter β,

$$\chi(\ell) = \frac{1}{2}\beta\ell^2,$$

and 3) the dealers balance sheet cost is quadratic with parameter γ,

$$c(b^{L,d}) = \frac{1}{2}\gamma(b^{L,d})^2.$$

We also assume that the bank's loan rate i_ℓ is a function of the IORB rate R, and in particular is a fixed spread over R, i.e.

$$i_\ell(R) = R + \mathcal{S}$$

for some $\mathcal{S} > 0$. Similarly, we assume that the investors' outside option ι is a function of the repo rate ρ, and in particular is a fixed spread over ρ, i.e.

$$\iota(\rho) = \rho + \widetilde{\mathcal{S}}$$

for some $\widetilde{\mathcal{S}} > 0$.

C.2 Calibration System

We have seven parameters that we must jointly calibrate. Let this vector of parameters be

$$\Xi \equiv \left\{\alpha, k^m, k^b, m, \gamma, \beta, \phi\right\}.$$

Let $S^i(\Xi)$ be a vector of moments generated under the parameter vector Ξ where $i \in \{model, data\}$. The objective is then to solve

$$\begin{aligned} \min_{\Xi} \quad & \left(S^{data}(\Xi) - S^{model}(\Xi)\right)^2 \\ \text{s.t.} \quad & EC(\Xi) = 0 \end{aligned}$$

where $EC(\Xi)$ is a system of equilibrium conditions evaluated with Ξ. For seven parameters, to be just identified we need at least seven moments in the objective – we use the following seven:

$$\begin{aligned} i\_db\_eq &= \left(\frac{i_{d^b}}{i\_db\_data} - 1\right)^2 \\ i\_dm\_eq &= \left(\frac{i_{d^m}}{i\_dm\_data} - 1\right)^2 \\ m\_r\_eq &= \left(\frac{m^r}{m\_r\_data} - 1\right)^2 \\ d\_b\_eq &= \left(\frac{d^b}{d\_b\_data} - 1\right)^2 \\ d\_onrrp\_eq &= \left(\frac{d^{RRP}}{d\_onrrp\_data} - 1\right)^2 \\ d\_m\_eq &= \left(\frac{d^m}{d\_m\_data} - 1\right)^2 \\ b\_L\_d\_eq &= \left(\frac{b^{L,d}}{b\_L\_d\_data} - 1\right)^2 \end{aligned}$$

Next, suppose we have the seven parameters in the vector Ξ, and the following independent values from the data

$$(R, r, \mathcal{S}, \widetilde{\mathcal{S}}, B^L, B^S, b^{L,cb})$$

then we can solve the following 11-equation system for the following 11-unknowns:

$$(i_{d^m}, i_{d^b}, \rho, d^b, d^m, d^{RRP}, b^{L,d}, b^{L,i}, \ell, m^r, \eta)$$

$$
\begin{aligned}
EC1 &= \left(\frac{1}{\phi(1+i_{d^b})} \left(\frac{1+i_{d^b}}{1+i_{d^m}} \right)^{1/\alpha} \right) - d^b \\
EC2 &= \left(m - \frac{1}{\phi(1+i_{d^b})} \left(\frac{1+i_{d^b}}{1+i_{d^m}} \right)^{1/\alpha} \right) - d^m \\
EC3 &= \left(d^m - \frac{b^{L,d} + b^{L,i} + B^S}{1+\rho} \right) - d^{RRP} \\
EC4 &= \left[\frac{\eta(i_{d^b})}{1+\eta(i_{d^b})} \right] \left(R - k^b \right) - i_{d^b} \\
EC5 &= \left(\frac{1-\alpha}{\alpha} \frac{i_{d^b}}{1+i_{d^b}} \right) - \eta(i_{d^b}) \\
EC6 &= \frac{1}{\gamma} \left(\frac{\widetilde{S}+\rho}{1+\widetilde{S}+\rho} - \frac{\rho}{1+\rho} \right) - b^{L,d} \\
EC7 &= \frac{\rho}{1+\rho}(b^{L,d} + b^{L,i} + B^S) + r d^{RRP} - d^m(i_{d^m} + k^m) \\
EC8 &= B^L - b^{L,cb} - b^{L,d} - b^{L,i} \\
EC9 &= r - \rho \\
EC10 &= \frac{1}{\beta}\mathcal{S} - \ell \\
EC11 &= d^b - \ell - m^r
\end{aligned}
$$

where we have imposed that we are in the abundant liquidity regime during the calibration period.

C.3 Simulation System

Suppose we have the seven parameters in the vector Ξ, and the following independent values from the data

$$(R, r, \mathcal{S}, \widetilde{\mathcal{S}}, B^L, B^S, b^{L,cb}, \delta)$$

then we can solve the following 11-equation system for the following 11-unknowns:

$$(i_{d^m}, i_{d^b}, \rho, d^b, d^m, d^{RRP}, b^{L,d}, b^{L,i}, \ell, m^r, \eta)$$

Letting $\mathcal{L}^\star = \frac{1}{\beta}\mathcal{S}$,

$$EC1 = \left(\frac{1}{\phi(1+i_{d^b})}\left(\frac{1+i_{d^b}}{1+i_{d^m}}\right)^{1/\alpha}\right) - d^b$$

$$EC2 = \left(m - \frac{1}{\phi(1+i_{d^b})}\left(\frac{1+i_{d^b}}{1+i_{d^m}}\right)^{1/\alpha}\right) - d^m$$

$$EC3 = \left(d^m - \frac{b^{L,d}+b^{L,i}+B^S}{1+\rho}\right) - d^{RRP}$$

$$EC4 = \left[\frac{\eta(i_{d^b})}{1+\eta(i_{d^b})}\right]\big(R - k^b + (1-\delta)[\mathcal{S} - \beta\ell]\big) - i_{d^b}$$

$$EC5 = \left(\frac{1-\alpha}{\alpha}\frac{i_{d^b}}{1+i_{d^b}}\right) - \eta(i_{d^b})$$

$$EC6 = \frac{1}{\gamma}\left(\frac{\widetilde{S}+\rho}{1+\widetilde{S}+\rho} - \frac{\rho}{1+\rho}\right) - b^{L,d}$$

$$EC7 = \frac{\rho}{1+\rho}(b^{L,d}+b^{L,i}+B^S) + rd^{RRP} - d^m(i_{d^m}+k^m)$$

$$EC8 = B^L - b^{L,cb} - b^{L,d} - b^{L,i}$$

if $\quad (b^{L,d}+b^{L,i}) \leq (1+r)\left(d^m - \frac{B^s}{1+r}\right)$

then $\quad EC9 = r - \rho$

else

then $\quad EC9 = \left(\frac{b^{L,d}+b^{L,i}}{1+\rho}\right) - \left(d^m - \frac{B^s}{1+\rho}\right)$

if $\quad (1-\delta)d^b \geq \mathcal{L}^\star$

then $\quad EC10 = (\ell - \mathcal{L}^\star) \quad$ and $\quad EC11 = (d^b - \ell - m^r)$

else

then $\quad EC10 = (\ell - (1-\delta)d^b) \quad$ and $\quad EC11 = (\delta d^b - m^r)$

D Data

Table D.1. Description of Model Moments & Data Sources

Variable	**Mapping to the Data**	**Data Source**	**Freq.**
R	Interest Rate on Reserve Balances (IORB)	FRB	Daily
r	Overnight Reverse Repo Facility rate (ON RRP)	FRBNY	Daily
Repo Rate, ρ	Tri-Party General Collateral Rate (TGCR)	FRBNY	Daily
Central bank's gov't bonds, $b^{L,cb}$	Federal Reserve's Treasury security holdings	FRB H.4.1 Data Release	Weekly
Dealer purchases of gov't bonds, $b^{L,d}$	Primary Dealer positions of Treasury securities	Primary Dealer Statistics, FRBNY	Weekly
Total gov't bonds in the economy, B^L	Treasury securities held by the Federal Reserve + Treasury securities held by primary dealers	FRB H.4.1, Primary Dealer Statistics, FRBNY	Weekly
Total gov't bills in the economy B^S	Treasury securities held by gov't MMFs	SEC N-MFP filing detail from FRB	Monthly
Reserves, m_r	Aggregate reserves provided by the Federal Reserve	FRB H.4.1 Data Release	Weekly
MMF repo lending, z^m	Gov't MMF lending in overnight Treasury repo	SEC N-MFP filing detail from FRB	Monthly
MMF investment at ON RRP, d^{ONRRP}	Gov't MMF investment at ON RRP	Confidential FRB data set	Daily
MMF deposits, d^m	Gov't MMF Overnight Treasury repo + Gov't MMF investment at the ON RRP + Gov't MMF Treasury securities	SEC N-MFP filings	Monthly
MMF deposit rate, i_{d^m}	Gov't MMF average annualized seven day yield, assuming no reinvestment, net of expenses	iMoneyNet, Inc., iMoneyNet Bulk Data - Offshore Analyzer & Gold Analyzer	Weekly
Bank deposits, d^b	Demand deposits + small-time deposits at commercial banks (i.e. "Other Deposits")	FRB H.8 Data Release	Weekly
Bank deposit rate, i_{d^b}	Average rate on demand deposits and small-time deposits at commercial banks	RateWatch	Weekly
Bank lending rate, i_ℓ	Average rates offered on new loan products by commercial banks (specifically consumer facing loan products, e.g. loan rates for new or used cars for 36/60/72 months, 15/20/30 year fixed rate mortgages, adjustable rate mortgages, 5/10/15 year home equity loans and home equity lines of credit)	RateWatch	Weekly
Dealer lending rate to investors ι	Volume-weighted average rate dealers lend to hedge funds (classified as funds that file an SEC Form PF)	sponsored service FICC	Daily
δ	Average reserves / deposits ratio between Thurs. Sept. 5 to Wed. Sept. 11, 2019	FRB H.4.1, FRB H.8 Data Release	Weekly

Notes: The public data sources are all hyperlinked. We remove the last three business days of the year from daily data series because of volatility in repo markets from year-end (Anbil and Senyuz, 2022).

E Robustness Tests

E.1 Alternative Lending Rate Assumptions

E.1.1 $\mathcal{S}$ increases in R.

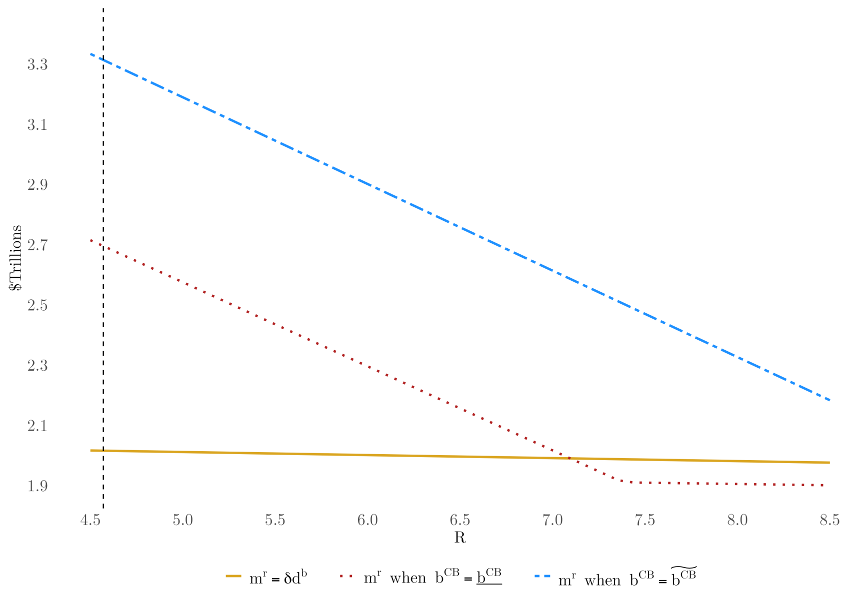
In our calibration, we assume that the spread between the loan rate i_ℓ and R is positive and constant, i.e. $\mathcal{S} \equiv i_\ell - R > 0$ and $d\mathcal{S}/dR = 0$. Here, we relax that assumption and consider an alternative specification where i_ℓ increases faster than R, i.e. $d\mathcal{S}/dR > 0$. In particular, we consider a parameterization of i_ℓ such that the implied spread $\mathcal{S}$ coincides with the value of the spread used in our calibration when $R = 4.57\%$ (our calibration average), and increases monotonically thereafter.

We investigate whether changing this assumption alters qualitatively our headline results in Figure 15. Computing the equilibria for increasing values of R, for the different central bank balance sheet critical values, under this alternative spread assumption gives the results displayed in Figure E.1. As one can see, the main result still holds – for the majority of the relevant range of R values, it is still the case that the capacity of the repo market is the relevant margin that binds first before the minimum level of reserve demand. This is expected, given Proposition 4.

Letting $\mathcal{S}$ be increasing in R changes reserve demand behavior relative to the baseline scenario where $\mathcal{S}$ is constant, and this induces the observed differences between Figure 15 and Figure E.1. Recall that in the baseline, reserve demand is decreasing in R since, even though the optimal loan amount $\mathcal{L}$ is constant, the endowment allocated to bank deposits is decreasing in R. When $\mathcal{S}$ is increasing in R, this original mechanism is occurring, but there is an additional force – the optimal loan amount $\mathcal{L}$ is increasing in R, inducing a further asset reallocation from reserves and to loans. Therefore, insofar as the bank is not constrained by the δ reserve constraint, fewer reserves are demanded at every R when $\mathcal{S}$ is increasing in R vis-a-vis when $\mathcal{S}$ is constant in R. This is what delivers the steeper decline of reserve levels seen in Figure E.1.

In our model, as shown by Proposition 4, for any assumption on the lending technology, there exists a level of the policy rate R' such that for all $R \leq R'$ the repo market capacity is the binding constraint, and conversely for all $R > R'$, lending is suboptimal and the bank is constrained such that $m^r = \delta d^b$ and $\ell = (1-\delta)d^b$. In Figure 15, this critical value R' is sufficiently large that it is off the range of the chart. In Figure E.1, under the alternative loan rate assumption, this critical value R' is less than the critical value with the constant spread assumption, and these critical value is visible on the chart for the case where $b^{L,cb} = \underline{b^{L,cb}}$ and $\rho = R$.

Figure E.1. Reserve Levels under Balance Sheet Critical Values when $d\mathcal{S}/dR > 0$



This figure shows the level of reserves when reserves are equal to the minimum level of reserve demand (solid orange line), when the Fed's government bond holdings are equal to $\widetilde{b^{L,cb}}$ (dashed blue line), and when the Fed's government bond holdings are equal to $\underline{b^{L,cb}}$ (dotted red line) – all under the alternative assumption that the loan rate increases faster than IORB, i.e. $d\mathcal{S}/dR > 0$. The interpretation of the thresholds is identical to that in Figure 15. Note that the flat portion of the dotted red line lies below the solid orange line. This is expected since the orange line represents $m^r = \delta d^b$ where d^b is from the abundant liquidity/reserve regime where $\rho = r$. The flat portion of the dotted red line coincides with an alternative $\widehat{m^r} = \delta\widehat{d^b}$ where $\widehat{d^b}$ is from the marginal scarce liquidity/reserve regime where $\rho = R$. Since the pass-through to i_{d^b} is less than that of i_{d^m}, we know that $\widehat{d^b} < d^b$ and therefore immediately that $\widehat{m^r} = \delta\widehat{d^b} < \delta d^b = m^r$.

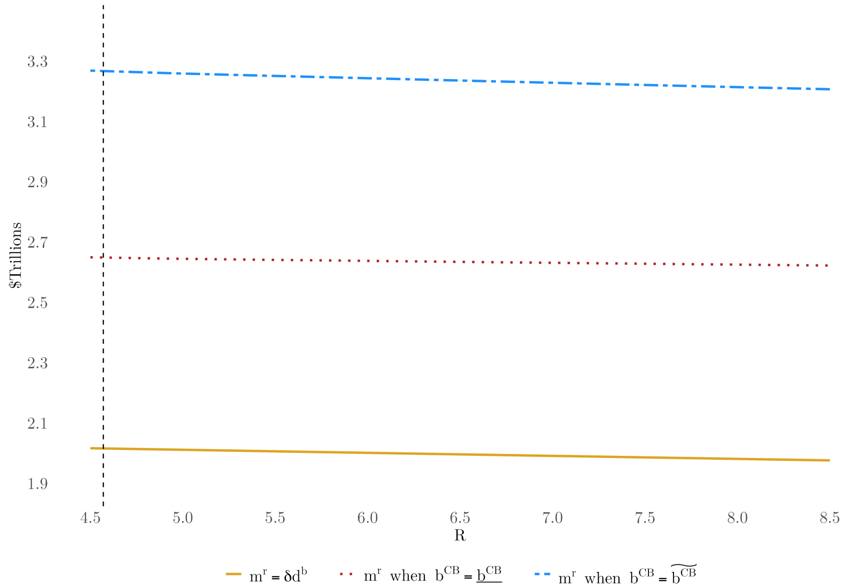
E.1.2 $\mathcal{S}$ decreases in R.

Next, we consider an alternative specification where i_ℓ increases slower than R, i.e. $d\mathcal{S}/dR < 0$. In particular, we consider a parameterization of i_ℓ such that the implied spread $\mathcal{S}$ coincides with the value of the spread used in our calibration when $R = 4.57\%$ (our calibration average), and decreases monotonically thereafter.

We investigate whether changing this assumption alters qualitatively our headline results in Figure 15. Computing the equilibria for increasing values of R, for the different central bank balance sheet critical values, under this alternative spread assumption gives the results displayed in Figure E.2. As above, the main result still holds.

Letting $\mathcal{S}$ be decreasing in R changes reserve demand behavior relative to the baseline scenario where $\mathcal{S}$ is constant, and as expected this induces the observed differences between Figure 15 and Figure E.2. In the baseline, reserve demand is decreasing in R since bank deposits are decreasing in R and loans are constant. Here, this original mechanism is attenuated – the optimal loan amount $\mathcal{L}$ is decreasing in R, so as d^b decreases, m^r does not decrease as much as in the baseline assumption. Therefore, in the region where the banks minimum reserve constraint is not binding, more reserves are demanded at every R, and this is what delivers the shallower decline in reserve levels seen in Figure E.2 vis-a-vis Figure 15. Similar to the baseline assumption, as implied by Proposition 4, the critical value of the policy rate R', where for all $R > R'$ lending is suboptimal and the bank is constrained, is sufficiently large that it is off the range of the chart.

Figure E.2. Reserve Levels under Balance Sheet Critical Values when $d\mathcal{S}/dR < 0$



This figure shows the level of reserves when reserves are equal to the minimum level of reserve demand (solid orange line), when the Fed's government bond holdings are equal to $\widetilde{b^{L,cb}}$ (dashed blue line), and when the Fed's government bond holdings are equal to $\underline{b^{L,cb}}$ (dotted red line) – all under the alternative assumption that the loan rate increases slower than IORB, i.e. $d\mathcal{S}/dR < 0$. The interpretation of the thresholds is identical to that in Figure 15.