Federal Reserve Tools for Managing Rates and Reserves[1]

Antoine Martin[2]
Federal Reserve Bank of New York

James McAndrews[3]
Wharton Financial Institutions Center

Ali Palida[4]
MIT

David Skeie[5]
Texas A&M University

December 2018

Abstract

The Federal Reserve began the reduction of its balance sheet in October 2017, but the eventual quantity of its outstanding reserves has not yet been decided. We show that the optimal supply of reserves equates banks' deposit rates to the interest rate paid on reserves (IOER). Bank deposit rates are determined by two frictions, banks' liquidity costs and balance sheet costs. Reserves should be reduced until deposit rates rise to the IOER. Raising the Fed's overnight reverse repo rate up to IOER would implement a higher optimal supply of reserves, expediently reduce the overabundance of reserves, and stabilize the volatility of overnight market rates.

Keywords: banks, balance sheet costs, liquidity, Federal Reserve, reserves, overnight reverse repurchases
JEL classification: G21, E5

---

[1]We thank Alex Bloedel and Sean Myers for research assistance and seminar participants at several seminars and conferences for helpful comments. The views expressed in this paper are those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of New York or the Federal Reserve System. Any errors are the responsibility of the authors.

[2]Email: antoine.martin@ny.frb.org.

[3]Email: jamesjmcandrews@gmail.com.

[4]Email: apalida@mit.edu.

[5]Corresponding author. Department of Finance, Mays Business School, Texas A&M University, College Station, TX 77843-4218. Email: dskeie@mays.tamu.edu.

1 Introduction

The Federal Reserve's balance sheet grew dramatically in the past decade to over $4 trillion in size, with reserves increasing from approximately $30 billion to over $2 trillion. In October 2017, the Fed began the gradual reduction of its balance sheet, but the eventual size of the balance sheet and quantity of reserves is still an open question. There is a currently a debate between advocates for abundant reserves, as has been the case post crisis, or very scarce reserves, as pre crisis. Indeed, market expectations on the longer-term level of reserves range from a few hundred billion dollars to over one trillion dollars.[6] In addition, the Fed has announced that it would phase out the overnight reverse repurchase (RRP) facility with non-banks when it is no longer needed to help control the federal funds rate.[7]

We analyze the optimal supply of reserves and use of overnight RRPs from the critical perspective of their impact through the banking system onto the economy. The unprecedented amount of reserves and the new overnight RRP operate through banks in ways not previously seen in practice and not studied until very recently.

Our general equilibrium approach delivers a number of novel results that provide strong insight and important policy guidance. We derive an optimal policy rule showing that the optimal supply of reserves is determined when bank wholesale deposit rates equal the interest rate on excess reserves (IOER). The corresponding supply of reserves equates banks' marginal liquidity and balance sheet costs. We also show that it is optimal to set the overnight RRP rate equal to the IOER rate. This efficiently reduces the overabundance of reserves to their optimal level and absorb bank liquidity shocks to stabilize the volatility of overnight market rates.

Our results derive from three main ingredients. First, reserves are banks' most immediate source of liquidity, and the cost of borrowing reserves on the interbank market is higher when reserves are more scarce. Second, bank moral hazard necessitates capital requirement regulation. Third, capital requirements create balance sheet costs for banks,

[6] The New York Fed reports that according to its survey of market participants, the expected level of reserves in 2025 ranges from $406 billion to $1 trillion at the 25th and 75th percentile levels, respectively. Source: https://www.newyorkfed.org/medialibrary/media/markets/omo/SOMAPortfolioandIncomeProjections_July2017Update.p

[7] https://www.federalreserve.gov/monetarypolicy/policy-normalization.htm

because equity is costly relative to deposits. Indeed, deposits provide households with liquidity whereas equity does not.

Bank liquidity costs and balance sheet costs have received particular attention in the last few years. The impact of larger bank balance sheet costs has drawn focus from regulators internationally.[8] Importantly, we show both costs are intimately tied to the central bank's balance sheet. Historically, the extreme scarcity of reserves caused large short-term liquidity costs because banks had to depend on very large amounts of borrowing in the interbank market to maintain liquidity. Before the adoption of IOER, the Fed had to rely on such scarcity of reserves to maintain positive short-term interest rates. The massive increase in reserves has more than satiated banks' short-term liquidity needs, as reflected by the near disappearance of the fed funds market. However, bank balance sheet costs arise from regulatory and economic costs that are proportional primarily to the size, but not the composition, of commercial banks' balance sheets, such as the Basel III leverage ratio and the FDIC assessment fee on banks' non-equity liabilities.

Our model delivers a particularly sharp policy rule for the optimal supply of reserves. The Fed should set the quantity of reserves such that the equilibrium bank deposit rate equals IOER. This result follows from the optimal balance between the benefit of reserves to provide liquidity to banks with the cost of reserves that follows from their requirement to be held only by banks. Given that deposit rates are below the rate on reserves, our model indicates that the Fed's supply of reserves is above the optimal quantity at present. Our model also highlights that it would be unwise to try to go back to the pre-crisis regime, with a bank deposit rate exceeding the implicit interest rate on reserves of zero. Overall, our results fall within the broad guidance of the Federal Open Market Committee's (FOMC) most recent normalization principles and plans, which state that the Fed's balance sheet will decrease to a level "appreciably below that seen in recent years but larger than before the financial crisis."[9] Importantly, we provide the first guide as to how an optimal moderate supply of reserves may be determined.

By focusing on the economic frictions that affect short-term rates broadly, our model also sheds new light on the important role that the Fed's overnight RRP plays in U.S.

[8] See, for example, the GCFS report on repo market functioning: http://www.bis.org/publ/cgfs59.htm.

[9] Source: https://www.federalreserve.gov/newsevents/pressreleases/monetary20170614c.htm.

money markets. Overnight RRPs are available to non-bank as well as bank counterparties at a rate set at a spread below IOER. We show that, in equilibrium, overnight RRPs with non-banks increase welfare by absorbing bank liquidity shocks, which reduces balance sheet costs and increases bank liquidity by enabling a higher optimal supply of reserves. The overnight RRP rate can be increased to equal IOER to ensure that the RRP provides the maximum welfare value. Increasing the RRP rate to IOER can also efficiently reduce the current overabundance of reserves to their optimal level more rapidly than relying on the Fed's current strategy of gradual asset run-offs alone. The overnight RRP also stabilizes the volatility of interest rates better than IOER does alone, something that is clearly observed empirically.

These insights are important as the Federal Reserve will have to make choices regarding whether to maintain the overnight RRP facility. In its most recent normalization principles and plans, the FOMC states: "During normalization, the Federal Reserve intends to use an overnight reverse repurchase agreement facility and other supplementary tools as needed to help control the federal funds rate. The Committee will use an overnight reverse repurchase agreement facility only to the extent necessary and will phase it out when it is no longer needed to help control the federal funds rate."[10] Our results point out that the overnight RRP should not be discontinued since it decreases equilibrium bank costs and rate volatility. We also argue against the use of the 'other supplementary tools.' These include longer-term RRPs, which we show are not as effective as the overnight RRP, and the term deposit facility (TDF) offered to banks, which we show is inefficient for reducing the overabundance of reserves.

The optimal quantities of reserves and the overnight RRP, which determine the size of the Fed's balance sheet, is a topic of growing policy and market attention. Several authors have recently analyzed this topic with contrasting views. Greenwood, Hansen, and Stein (2015, 2016) argue for a large Fed balance sheet to supply large amounts of overnight RRPs for financial stability reasons. They reason that overnight RRPs can act similarly to their documented findings of Treasury bills crowding out banks' production of money-like assets, complementing the work of Nagel (2016).[11] Gagnon and Sack (2014) advocate

[10]Source: https://www.federalreserve.gov/monetarypolicy/policy-normalization.htm.

[11]Additional advocates for a large Fed balance sheet include Cochrane (2014), who argues for sizable

for setting the overnight RRP rate at IOER and maintaining an abundance of reserves as optimal. We show that the RRP rate set to IOER will prevent an abundant quantity of reserves in equilibrium, and that a moderate equilibrium quantity of reserves is optimal. Sims (2016) argues for a small Fed balance sheet based on the maturity mismatch between the Fed's assets and reserves liabilities, which creates risks to the Fed's net worth, political support, and policy independence. Williamson (2016) shows that excessive reserves cause bank balance sheet costs because of binding capital requirements. He argues for reducing the overabundance of reserves either directly or else indirectly through the overnight RRP to improve welfare.[12]

In contrast to this recent studies, which generally advocate either for a very "large" Fed balance sheet, as current, or a very "small" Fed balance sheet, as before 2008, we find that the Fed balance sheet size should be neither too small nor too large. Rather, the optimal size is determined by a moderate quantity of reserves and overnight RRPs. A larger quantity of reserves provides diminishing liquidity benefits outweighed by the costs of bank balance sheet expansion, whereas a smaller quantity creates increasing bank illiquidity costs that outweigh the benefits of bank balance sheet reduction. The overnight RRP also has a moderate size determined in equilibrium by market demand when the RRP rate is optimally set equal to IOER. The range of costs and benefits of a larger or smaller Fed balance sheet in the recent academic literature could also likely be incorporated into our marginal cost and benefit approach to result in a related moderate-size solution as optimal.

The paper proceeds with an institutional background on the Fed's balance sheet policies, followed by the analysis of optimal reserves and then the overnight RRP. Proofs and extensions are located in the appendices.

---

overnight RRPs to crowd out private short-term debt creation and large reserves to satiate monetary liquidity needs. Duffie and Krishnamurthy (2016) argue for sizable overnight RRPs to produce a more direct and effective transmission of interest rate policy to money markets and financial markets. Bernanke (2016) supports extensive overnight RRPs and large reserves for better establishing the Fed as lender-of-last-resort. Additional arguments for large reserves are for banks' provision of payment services (Curdia and Woodford, 2011), broad liquidity services (Goodfriend, 2002), and the effective transmission of monetary policy (Reis, 2016). Gertler and Karadi (2013) argue for a large amount of Fed assets to help circumvent bank leverage constraints.

[12] Stein (2012) and Kashyap and Stein (2012) advocate for small reserves to maintain their scarcity value and limit bank creation of money-like deposits for financial stability reasons. Bindseil (2016) argues that central banks should have small balance sheets in order to focus on their core mandate.

2 Institutional background

Historically, the Federal Reserve supplied a scarce amount of reserves in the banking system to maintain positive interest rates. Depository institutions (DIs) in the U.S. hold 'required reserves' for reserve requirements and 'excess reserves' for precautionary reasons to meet liquidity shocks. DIs in need borrow reserves from each other at the federal funds rate.[13] This rate represented the marginal cost for banks' most immediate liquidity shocks and influences other bank funding rates and money market rates through arbitrage and substitution. The Fed targeted the fed funds rate by adjusting the supply of reserves using open market operations (OMOs) as its primary policy. In an OMO, the open market trading desk at the Federal Reserve Bank of New York would buy or sell government securities with 'primary dealer' counterparties either on a temporary basis (using repurchase agreements, i.e., RPs) or on a permanent basis (using outright transactions).[14] For example, purchasing Treasuries would increase the supply of reserves and decrease the fed funds rate.[15]

This scarcity method for managing short-term rates was no longer available once the level of reserves dramatically increased starting in late 2008. Reserves grew as a by-product of the Fed's crisis liquidity operations and subsequent large-scale asset purchases (i.e., quantitative easing) aimed at lowering longer-term rates.[16] With a supply of reserves far beyond banks' demand for reserve requirements and precautionary liquidity, banks' have a zero-interest rate marginal value for reserves that do not pay interest.

IOER To manage rates with abundant reserves, the Fed prepared a variety of new policy tools.[17] The Fed began to pay IOER to DIs in October 2008, following Congressional

[13] A few other type of institutions, such as government-sponsored enterprises (GSEs), also participate in the federal funds market.

[14] The list of primary dealers is at http://www.newyorkfed.org/markets/primarydealers.html Assets eligible for OMOs are Treasuries, agency debt, and agency mortgage-backed securities (MBS).

[15] See Ennis and Keister (2008) and Keister et al. (2008) for a more detailed introduction to traditional Federal Reserve monetary policy and OMOs.

[16] Gagnon et al. (2011) provide an analysis of the first LSAP. Details on the Fed's asset purchase program are at: https://www.federalreserve.gov/faqs/what-were-the-federal-reserves-large-scale-asset-purchases.htm.

[17] The list of Fed policy tools is at: https://www.federalreserve.gov/monetarypolicy/policytools.htm.

authorization included in the Emergency Economic Stabilization Act of 2008.[18,19] The Fed distinguishes between IOER as its primary policy rate and the fed funds rate as its policy target. In December 2008, the FOMC lowered IOER to 25 basis points and set the fed funds target rate to a range of 0 to 25 basis points, its effective zero bound.[20] Since December 2015, the FOMC has raised IOER and the target range for the fed funds rate several times in increments of 25 basis points.

RRPs and TDF The development of RRPs and the TDF began in 2008 to act as the "suspenders" supporting the IOER "belt" for the Fed to "retain control of monetary policy," according to NY Fed President Dudley in a 2009 speech.[21] RRPs are economically analogous to collateralized loans made to the Fed by its expanded set of 164 counterparties, which include DIs and non-DIs such as money market mutual funds (MMFs), GSEs, and securities dealers.[22] The TDF allows DIs to make term deposits of reserves at a rate above IOER.[23] RRPs and the TDF do not change the size of the Fed's balance sheet, but modify the composition of its liabilities by reducing reserves.

Each of these tools has been developed and tested with both fixed-quantity auctions at market determined rates and at fixed-rates with market determined quantities. Small-value testing of the term RRP from 2009 until 2015 and the TDF since 2010 have been conducted for a range of maturities up to 28 days. Small-value testing of the overnight RRP began in 2009 with a fixed-quantity before switching in 2013 to a fixed-rate.[24] The overnight RRP with a fixed-rate set at the bottom of the fed funds target range was

[18] Source: http://www.newyorkfed.org/newsevents/speeches/2009/dud090729.html.

[19] In principle, the Federal Reserve could set a different rate for the interest on required reserves and the interest on excess reserves. In practice, the two have been the same since 2009.

[20] Source: http://www.federalreserve.gov/monetarypolicy/files/fomcminutes20081216.pdf.

[21] Source: http://www.newyorkfed.org/newsevents/speeches/2009/dud090729.html.

[22] Pre-crisis, RRPs with primary dealers were used on a very infrequent basis by the Fed as a part of standard OMOs. Details on RRPs and the list of RRP counterparties are at the following two links, respectively: http://data.newyorkfed.org/aboutthefed/fedpoint/fed04.html https://www.newyorkfed.org/markets/rrp_counterpartiesarties.html.

[23] Details on the TDF are at: https://www.federalreserve.gov/monetarypolicy/tdf.htm. The TDF was approved April 2010, following the approval of amendments to Regulation D (Reserve Requirements of Depository Institutions), allowing Federal Reserve Banks to offer term deposits to institutions eligible to earn interest on reserves. Source: http://www.federalreserve.gov/newsevents/press/monetary/20100430a.htm.

[24] Source: http://www.federalreserve.gov/monetarypolicy/files/fomcminutes20130731.pdf.

formally adopted by the FOMC in March 2015.[25]

Normalization In September 2017, the Fed announced its highly anticipated normalization plan to gradually reduce its balance sheet. Directly selling bonds had been previously ruled out as too risky after the 2013 "taper-tantrum," when bond market volatility erupted based on then-Chairman Bernanke's comments about the end of Fed asset purchases. Instead, the Fed has stopped reinvesting capped amounts of proceeds from its maturing Treasuries and agency mortgage-backed securities. This strategy allows for a gradual run-off of the Fed's assets and hence gradual decrease in the supply of reserves. The end-point for this run-off and the eventual new-normal quantity of reserves has not been yet decided.[26]

3 Model

In this section, we develop the model with interest on reserves to determine the optimal supply of reserves in section 4. In Section 5, we add the overnight RRP to the model to analyze its optimal policy use.

3.1 Set up

The economy lasts for three dates 0, 1, 2, and is populated by competitive risk-neutral households who maximize expected utility, banks and firms that maximize expected profits, and a central bank and government. At date 0, households have an endowment of g goods. Banks, firms, and the central bank do not receive an endowment.

At date 0, firms borrow from banks to buy goods as inputs into their production technology. The government buys goods for an exogenous amount of fiscal consumption financed by issuing government bonds. Households use proceeds from selling goods to acquire government bonds, bank deposits, and preferred bank equity, which we refer to as *equity*. The central bank buys government bonds with newly created reserves, which

[25] Source: https://www.federalreserve.gov/monetarypolicy/policy-normalization.htm.

[26] Details are at: https://www.federalreserve.gov/monetarypolicy/fomcminutes20170920.htm.

act as the nominal unit of account in the economy.[27] Reserves can only be held by banks and earn a rate of interest set by the central bank.

To create a motive for interbank trading we assume that bank deposit markets are segmented and banks are subject to a liquidity shock at date 1. Households can only deposit at the banks in their sector and, for simplicity, we also assume that households only hold equity in banks in their sector, and firms only borrow from banks in their sector. These latter two assumptions are not required for the results.

The liquidity shock takes the form of a deposit withdrawal from households in one of the sectors, which is used to buy government bonds. The shocked-sector banks may meet the withdrawal with their reserve holdings. If they do not have enough reserves, the can borrow in the interbank market. Households in the other sector sell bonds and deposit the proceeds with their banks.

At date 2, firms sell the goods they produced. Households receive the proceeds from their deposit, equity, and government bond holdings, as well as the (common equity) residual profits of their sector's banks and firms. Households also pay a lump-sum tax out of their nominal revenues, and buy goods from firms to consume.

Ingredients Our results derive from three standard ingredients from the banking literature, which are detailed in the next subsection. These frictions are added in a reduced form manner to focus the analysis on the logical implications of these ingredients for central bank policy. First, interbank market frictions lead to bank liquidity costs when there is a scarcity of reserves. Second, bank moral hazard arises from inefficient risky projects that banks can take to shift risk onto depositors. The government, as regulator, imposes capital requirements in the form of a minimum amount of equity. Third, equity is relatively costly because it does not provide households the liquidity value of deposits and bonds.

[27] We consider only the central bank and households as holders of bonds, as banks and firms would not hold bonds in equilibrium.

3.2 Optimizations

Households Following the recent literature establishing the liquidity value of money-like assets, households have a liquidity factor on deposits and bonds, denoted by $\theta > 1$.[28] We assume that the size of the liquidity shock is proportional to the bank assets held by the household in the shocked sector. Those assets are deposits, denoted by D, and preferred equity, denoted by E. The amount of early withdrawals can thus be written as $D^W \equiv \lambda(D + E)$, where λ is the size of the liquidity shock.

Each household receives the return on the equity it holds ($R^E E$) as well as the residual profits from the bank (Π^{Bj}) and firm (Π^{Fj}) in its sector, and pays a lump sum tax Υ. In addition, the households' nominal income depends on whether it is in the shocked sector, $j = s$, or not, $j = n$. We use $\mathbf{1}$ as the indicator function to show the additional returns received by the shocked household and nonshocked households, respectively, where $\mathbb{E}[\mathbf{1}_{j=s}] = \mathbb{E}[\mathbf{1}_{j=n}] = \frac{1}{2}$.

Households in the shocked sector decrease their deposits by D^W and increase their holdings of government bonds by a corresponding amount, denoted B_1^s. Households in the nonshocked sector sell B_1^n government bonds and deposit a corresponding amount with the bank in their sector, denoted D_1. In addition, these households have to hold the additional preferred equity, E_1, that the bank in their sector issues. Formally, we can write the household's expected nominal income as:

$$\begin{aligned}\Pi^{Hj} \quad = \quad & R^E E + \Pi^{Bj} + \Pi^{Fj} - \Upsilon \\ & +\mathbf{1}_{j=s}\theta[R^D(D - D^W) + R^B(B^H + B_1^s)] \\ & +\mathbf{1}_{j=n}\{R^{E1}E_1 + \theta[R^D D + R^{D1}D_1 + R^B(B^H - B_1^n)]\}.\end{aligned} \tag{1}$$

Note that equity doesn't benefit from the the liquidity factor $\theta > 1$.

A household's utility is the real value of its nominal profit at date 2, $u^j \equiv \frac{\Pi^{Hj}}{P_2}$. The

[28] The liquidity value of money-like short-term bank debt is the centerpiece of Greenwood, Hanson and Stein (2015, 2016) and is based on monetary benefit services for households in Stein (2012) and a transactions value from lower information sensitivity in Dang et al. (2009, 2015, 2017) and originally in Gorton and Pennacchi (1990), which are motivated by the liquidity value of deposits originating in Diamond and Dybvig (1983) and Bryant (1980). More broadly, the safety and liquidity premia of Treasuries is demonstrated by Krishnamurthy and Vissing-Jorgensen (2011, 2012), Caballero and Krishnamurthy (2009), and Krishnamurthy (2002).

household optimization is:

$$\left.\begin{aligned} &\max_{\mathcal{Q}^{Hj}} \mathbb{E}[u^j] \\ &\text{s.t.} \quad D + E + B^H \leq G \\ &\qquad P_1^B B_1^s \leq R^W D^W \\ &\qquad D_1 + E_1 \leq P_1^B B_1^n \\ &\qquad B_1^n \leq B^H. \end{aligned}\right\} \tag{2}$$

where $\mathcal{Q}^{Hj} \equiv \{D, E, B^H, D_1, E_1, B_1^j\}$. G denotes the household's nominal quantity of goods at date 0. Without loss of generality we normalize the price level at date 0 to equal 1, so $G = g$. The first three inequalities are the budget constraints at date 0 for each household and at date 1 for the shocked and nonshocked households, respectively. The term structure of deposit returns implies that the return on early withdrawals at date 1 at the shocked bank is $R^W = \frac{R^D}{R^{D1}}$. The last inequality is a feasibility constraint.

Firms The firm in sector $j \in \{n, s\}$ demands loans to buy the quantity L of date-0 goods, which they use to produce goods sold at date 2. The price, P_2, at which these goods are sold is endogenously determined. The firm chooses L to maximize profits:

$$\max_L \Pi^{Fj} = P_2 \int_0^L r(\hat{L}) d\hat{L} - R^L L. \tag{3}$$

We assume that the marginal real return on production by firms is greater than one and follows standard Inada conditions: $r(L) > 0$, $r'(L) < 0$, $r(0) = \infty$, $r(\infty) = 1$.

Banks We start this section by deriving the banks' profits. Banks earn interest income on their assets: loans, reserves, and potentially interbank loans. They pay interest on their liabilities: deposits, equity, and potentially interbank borrowing.

All banks hold the same amount of loans, L. The amount of reserves held by a bank at the end of date 1 depends on whether or not the bank is in the shocked sector. A bank in the shocked sector has to pay $R^W D^W$ to withdrawing depositors and can borrow reserves in the interbank market. We denote interbank borrowing by I^s. A bank in the nonshocked sector accrues reserves corresponding to additional deposits, D_1, and equity, E_1, and can lend in the interbank market. Interbank market loans are denoted by I^n. In addition, banks in both sectors receive interest from the central bank on the reserves they hold at the end of date 0. Formally, we have

$$M_1^s = R^M M + I^s - R^W D^W \tag{4}$$

$$M_1^n = R^M M - I^n + D_1 + E_1 \tag{5}$$

for the shocked and nonshocked banks, respectively.

With this we can write the date 2 profit for a bank in sector $j \in \{n, s\}$ as:

$$\begin{aligned} \Pi^{Bj} = & \; R^L L + R^M M_1^j \\ & -\mathbf{1}_{j=s}[R^E E + R^D(D - D^W) + R^I I^s] \\ & -\mathbf{1}_{j=n}\{[R^E E + R^D D + R^{E1} E_1 + R^{D1} D_1 + \textstyle\int_0^{I^n} Y(\hat{I}^n) d\hat{I}^n] - R^I I^s\}. \end{aligned} \tag{6}$$

The first line represents the interest income on reserves and loans. The second and third lines represent the cost of the bank's liabilities, except for the case where a bank lends in the interbank market, in which case I^s represents an asset. The second line corresponds to the bank in the shocked sector and the third line to the bank in the sector that did not experience a shock. The cost of the bank's liabilities is affected by two frictions, a liquidity cost and capital requirements, to which we turn now.

Liquidity costs In the third line, $Y(I^n)$ represents the marginal cost of monitoring interbank loans. We assume that $Y(0) = 0$ and $Y'(\cdot) > 0$ to capture the fact that as a bank's interbank lending exposure grows, more monitoring is necessary. Hence, the cost of providing liquidity to the interbank market increases with the amount of liquidity provided.

In practice banks assign maximum allowable counterparty exposure limits that imply an increasing marginal shadow cost of interbank lending that is ultimately prohibitive. The interbank cost can also be interpreted as representing other costs in the interbank market, such as search costs, that increase when reserves are more scarce in equilibrium.[29]

Below, we show that, in equilibrium, the shocked bank bears the entire interbank cost through the interbank rate it pays because that bank has an inelastic demand for

[29] A more general model of various interbank market costs paid directly by both the shocked and nonshocked banks would not change the paper's results.

interbank borrowing to meet its early withdrawals.[30]

Capital requirements and balance sheet costs Banks face moral hazard in the form of risk-shifting onto depositors. A bank can take an unobservable, negative expected NPV project that has a positive or negative return with equal probability. Specifically, each bank can take the project at date 0 for a marginal return, if there is a positive realization, of $R^{\alpha}(A)$ on the bank's date 0 assets A, where

$$A = L + M.$$

The nonshocked bank can also take the project at date 1 for a marginal return, if there is a positive realization, of $R^{\alpha}(A + A_1)$ on the bank's new date 1 assets A_1. A_1 is equal to the amount of the bank's new liabilities at date 1:

$$A_1 = D_1 + E_1.$$

We assume that $R^{\alpha\prime}(\cdot) > 0$, which implies that the bank's ability for risk-taking is increasing in the size of the bank's balance sheet at dates 0 and 1.

In practice, as banks grow larger, they can undertake greater amounts of hidden risk-taking, such as through derivatives and off-balance sheet exposures that do not require more verifiable initial investments as with loans. For example, banks are the predominant participant in the trillion-dollar market for interest rate swaps. Determining whether derivatives are used for hedging or speculating depends on sophisticated quantitative models of detailed bank-specific information on its assets and liabilities.

The bank chooses whether to take the risky project at dates 0 and 1 to maximize profit subject to the expected return required by equity. If there is a negative realization on the project, the bank and equityholders bear a complete loss. The depositors bear a partial loss and lose their liquidity value on deposits, such that $\theta = 1$. Hence, the risky project is a form of risk-shifting that is socially inefficient.

[30] The large literature on interbank monitoring costs originates with Rochet and Tirole (1996) and is detailed and broadly developed by Freixas and Rochet (2008) and Rochet (2008). Interbank market costs are also extensively studied in the more recent literature on the search and matching frictions in the bilateral, OTC fed funds market. See, e.g., Afonso and Lagos (2015), Armenter and Lester (2017), Ashcraft and Duffie (2007), Atkeson et al. (2015), and Bech and Monnet (2016). Limits to interbank borrowing capacity are studied as arising from liquidity and credit constraints by Acharya and Skeie (2011) and Ashcraft et al. (2011), and from moral hazard in monitoring by Acharya et al. (2012).

The government as regulator imposes a capital requirement that is sufficient to incentive banks not to take the risky project. The full analysis analyzed in appendix B shows that the capital requirement is

$$\begin{aligned}\underline{E}(A) &\equiv \tfrac{R^{\alpha}(A)A}{R^{E}} \\ \underline{E}_1(A, A_1) &\equiv \tfrac{R^{\alpha}(A+A_1)A_1}{R^{E1}}.\end{aligned} \qquad (7)$$

The capital requirement acts as a "balance sheet cost" because it increases the cost of growing a bank's balance sheet.[31] Frictions related to bank balance sheet size are motivated in part by the analysis of market observers. For example, interbank broker Wrightson ICAP (2008) voiced concerns that large reserves could "clog up bank balance sheets." Under Basel III, U.S. banks are subject to a leverage ratio that is higher, and assessed on a broader base, than was the case pre-crisis. Thus, regulatory-based bank balance sheet costs are likely to be relevant.[32] An additional source of balance sheets costs is the FDIC deposit insurance assessment fees that are applied to all non-equity bank liabilities and increase with banks' balance sheet size. The balance sheet cost in our model can be interpreted as also capturing the effect of the FDIC assessment.[33] Furthermore, banks have tended to reduce the size of their balance sheets as the quantity of reserves increased, which Martin et al. (2016) demonstrate may be explainable by the presence of balance sheet costs and a large level of reserves partially crowding out bank lending.[34]

[31] The impact of balance sheet costs based on the leverage and capital constraints for financial intermediaries has been studied for asset pricing, financial crises, and monetary policy in the recent literature, most notably by Adrian and Shin (2008, 2009, 2014), Adrian et al. (2014), He and Krishnamurthy (2012, 2013), He et al. (2010), and He et al. (2017).

[32] Portfolio manager of the Fed's balance sheet Simon Potter (2014) states that "An increase in bank reserves that increases bank assets makes regulatory leverage ratios more binding, raising the shadow marginal cost of bank balance sheets...as the level of reserves declines during normalization, marginal balance sheet costs should fall."

[33] Armenter and Lester (2017) show in a calibrated model that bank balance sheet costs are increasing in reserves and deposits and are driven by direct expenses, such as FDIC fees, and indirect expenses, such as capital requirements and leverage ratios.

[34] Additionally, NY Fed President William Dudley states that "to the extent that the banks worry about their overall leverage ratios, it is possible that a large increase in excess reserves could conceivably diminish the willingness of banks to lend." Source: http://www.newyorkfed.org/newsevents/speeches/2009/dud090729.html.

Bank optimization We can now write the bank's problem, which is given by

$$\left.\begin{aligned} &\max_{\mathcal{Q}^{Bj}} \mathbb{E}[\Pi^{Bj}] \\ &\text{s.t.} \quad L + M \leq D + E \\ &\qquad M_1^j \geq 0 \qquad \text{for } j \in \{n, s\} \\ &\qquad E \geq \underline{E} \\ &\qquad E_1 \geq \underline{E}_1, \end{aligned}\right\} \tag{8}$$

where $\mathcal{Q}^{Bj} \equiv \{L, M, D, E, I^j, D_1, E_1\}$. The first two inequalities are the bank's budget constraints for dates 0 and 1, respectively.[35] The last two inequalities are the bank's capital requirements to issue a minimum amount of equity, $\underline{E}$ and $\underline{E}_1$ at dates 0 and 1, respectively.

Central bank and government budget constraints The budget constraint for the central bank at date 0 is

$$B^{CB} \leq M, \tag{9}$$

which shows that the quantity of bonds that the central bank buys as assets is limited by the quantity of reserves the central bank supplies at date 0.[36] In subsection 3.3 below, an equilibrium is defined for any choice of the supply of reserves, M. The central bank's optimization problem for choosing the optimal supply of reserves is presented in section 4.2 on welfare and optimal policy. We consider the return on reserves (R^M) as exogenous since it is a policy tool of the central bank. The level of R^M doesn't affect welfare in our model as we focus on the optimal supply of reserves for any given rate on reserves.

At date 2, the government pays the return $R^B B$ on bonds and receives the lump sum tax Υ. The central bank receives the return $R^B B^{CB}$ on its bond holdings and pays banks the return $R^{M^2} M$ on reserves. The date 2 consolidated budget constraint for the government and the central bank is

$$R^B B + R^{M^2} M \leq \Upsilon + R^B B^{CB}.$$

[35] For simplicity, we abstract from required reserves for banks, as they do not play a meaningful role and would not alter our results. Bennett and Peristiani (2002) show that reserve requirements have been largely avoided in the U.S. since the 1980s by sweep accounts. The small amounts of remaining reserve requirements are largely met by vault cash that banks hold for retail purposes and so pose little cost for banks.

[36] Abstracting from currency as a liability of the central bank does not effect our analysis.

Since M and hence B^{CB} are chosen by the central bank, B and R^{M^2} are taken as exogenous, and R^B is an endogenous equilibrium variable, the lump sum tax Υ required[37] to meet net government liabilities at date 2 is

$$\Upsilon \geq R^B(B - B^{CB}) + R^{M^2}M. \tag{10}$$

Goods market at dates 0 and 2 At date 0, firms buy L goods, the government buys B goods, and households sell their endowment of goods g at a price normalized to 1. For simplicity, we do not consider households storing goods, which they would not do in equilibrium except for instances of extreme balance sheet costs.

At date 2, firms sell their production of goods, $\int_0^L r(\hat{L})d\hat{L}$. The household in sector j buys consumption goods c^j with its net nominal revenue, which is equal to the household's profit given by equation (1) with the household liquidity factor set to one. Thus, $c^j P_2 = \Pi^{Hj}(\theta = 1)$. The marginal real cost for interbank lending across sectors by the nonshocked bank is $y(I^n) = \frac{Y(I^n)}{P_2}$, which is considered to be a cost in terms of real resources. These resources are acquired by the nonshocked buying a total of $\int_0^{I^n} y(\hat{I}^n)d\hat{I}^n$ goods at date 2.

3.3 Equilibrium

Assumptions To focus on a size of liquidity shocks that is consistent with common money market flows, we assume that λ that is less than $\frac{R^M}{R^W}$, $\frac{D}{D+E}$, and $\frac{B-M}{G-(B-M)}$. This ensures that, at date 1, the amount of interbank borrowing weakly decreases in the quantity of reserves, shocked household withdrawals of deposits are feasible, and nonshocked household sales of bonds are feasible, respectively.

Definition 1 *An equilibrium in the economy consists of the two-period returns for date 0 assets* $(R^L, R^D, R^E, R^B) > 0$*, the one-period returns for date 1 assets* $(R^I, R^{D1}, R^{E1}) > 0$*, the price of bonds at date 1* $P_1^B > 0$*, and the price of goods at date 2* $P_2 > 0$*; such that, at the optimizing quantities* Q^{Hj} *for the household in each sector* $j \in \{n, s\}$ *given by (2),* L *for each firm given by (3), and* Q^{Bj} *for the bank in each sector* $j \in \{n, s\}$

[37] Since the government tax is determined based on the central bank's choice of reserves, the determination of P_2 in equilibrium represents a simple form of the fiscal theory of the price level with monetary dominance and passive fiscal policy.

given by (8); and at the central bank and government exogenous quantities M and B, and endogenous quantities B^{CB} and Υ given by binding budget constraints (9) and (10), respectively; markets clear for:

(a) deposits (D), equity (E) and loans (L) within each sector $j \in \{n, s\}$ at date 0;

(b) deposits (D_1) and equity (E_1) within sector $j = n$ at date 1;

(c) reserves at date 0 (M) and date 1, $M_1^n + M_1^s = 2R^M M$;

(d) bonds at date 0, $B^{CB} + B^H = B$;

(e) bonds at date 1, $B_1^n = B_1^s$;

(f) interbank loans at date 1, $I^n = I^s$;

(g) goods at date 0, $L + B = g$; and

(h) goods at date 2, $\sum_{j\in\{n,s\}} c^j + \int_0^{I^n} y(\hat{I}^n)d\hat{I}^n = 2\int_0^L r(\hat{L})d\hat{L}$.

Proposition 1 *There exists a unique equilibrium.*

We proceed by analyzing the equilibrium effects of reserves on bank liquidity and balance sheet costs in order to then determine the optimal supply of reserves.

4 Reserves

Banks face both benefits and costs of increasing their reserves holdings. On the one hand, holding more reserves creates a buffer against demand shocks, which protects the bank from having to fund withdrawals with costly interbank loans. On the other hand, banks face capital requirements and must hold more equity when they hold more reserves. Equity is costly because it is not liquid from the perspective of households. In this section we study this tradeoff in more detail.

4.1 Balance sheet costs and liquidity costs

First, we look at how the bank deposit rate is determined. We show that, in equilibrium, the deposit rate is priced at a spread to IOER, which measures banks' liquidity costs relative to balance sheet costs.

It will be useful to have an expression for the expected return on loans. Because banks behave competitively, the return on loans is equal to the (expected) cost of funding a loan. This cost includes the deposit rate (R^D), as well as the costs corresponding to the capital requirements and liquidity shocks. The *balance sheet cost* is denoted $K(A)$ and the (expected) *liquidity cost of deposits* is ($\frac{1}{2}\lambda R^W Y(I)$). We derive these expressions below. The expected return on loans can thus be written as

$$R^L = R^D + K(A) + \tfrac{1}{2}\lambda R^W Y(I). \tag{11}$$

Bank balance sheet costs The balance sheet cost is the (expected) marginal capital requirement cost for a marginal increase in balance sheet size at date 0:

$$K(A) \equiv (R^E - R^D)\tfrac{dE}{dA} + \tfrac{1}{2}(R^{E1} - R^{D1})\tfrac{dE_1}{dA}. \tag{12}$$

The spreads $(R^E - R^D) = (\theta - 1)R^D$ and $(R^{E1} - R^{D1}) = (\theta - 1)R^{D1}$ are the premiums on equity returns relative to deposit returns at dates 0 and 1, respectively. These spreads are required by households to hold equity, which does not provide households the liquidity value ($\theta > 1$) on deposits. For a marginal increase in the bank's balance sheet size at date 0, the expected additional amounts of required equity are $\frac{dE}{dA}$ at date 0 and $\frac{dE_1}{dA_1}$ at date 1, respectively. This highlights the impact of equity requirements on the balance sheet cost at date 0 as arising from two sources in equation (12). The first term is the immediate marginal equity required at date 0. The second term reflects that the bank prices in the effect of the current balance sheet size on increases in its expected date 1 equity requirement.

Liquidity cost of deposits If a bank has a withdrawal shock at date 1, it has to borrow on the interbank market if it is illiquid. This is the case if the early withdrawal amount $R^W D^W$ is greater than the bank's date 1 entering reserves $R^M M$. Hence, the shocked bank's borrowing on the interbank market is

$$I = (R^W D^W - R^M M)^+. \tag{13}$$

If the shocked bank is liquid, then $I = 0$ and $Y(I = 0) = 0$, and the liquidity cost of deposits in the second term of equation (11) is zero. If the shocked bank is illiquid, then the bank has an inelastic demand for borrowing $I > 0$. This implies that the shocked bank bears the interbank cost $Y(I)$ in equilibrium through the interbank rate spread above the rate on reserves, which is $R^I - R^M = Y(I > 0) > 0$. The amount that the bank must borrow for date 1 withdrawals per unit of the bank's date 0 deposits is $\frac{dD^W(D)}{dD} = \lambda R^W$. Since each bank has an ex-ante one-half probability of being shocked and paying the marginal cost $Y(I)$ for borrowing λR^W, at date 0 each bank has an expected marginal cost of $\frac{1}{2}\lambda R^W Y(I)$ for interbank borrowing to cover depositor liquidity shock withdrawals.

Liquidity value of reserves The return on loans can also be expressed in terms of the asset liquidity premium required by banks to lend to firms rather than hold reserves:

$$R^L = R^M[R^M + \tfrac{1}{2}(R^I - R^M)]. \tag{14}$$

On the RHS of equation (14), the first factor R^M is the bank's opportunity cost of holding reserves from date 0 to 1. In the brackets, the first term R^M is the opportunity cost of holding reserves from date 1 to 2. The second term in the brackets gives the one-half probability that the bank has the liquidity shock and has to pay the interbank borrowing spread, which is positive if the shocked bank is illiquid: $R^I - R^M = Y(I > 0) > 0$. In contrast to reserves, loans are illiquid assets that are not available for paying out to meet liquidity shock withdrawals. Equation (14) can be rewritten to show the asset liquidity premium spread between the returns on loans and reserves:

$$R^L - R^{M^2} = \tfrac{1}{2}R^M Y(I). \tag{15}$$

We define $\frac{1}{2}R^M Y(I)$ as the bank's (expected, marginal) *liquidity value of reserves.*

Net liquidity value of reserves We can now derive the bank's implicit funding cost for holding reserves. Substituting for the return on loans in the asset liquidity premium equation (15) with the funding cost of loans in equation (11), and solving for R^{M^2}, gives

$$R^{M^2} = R^D + K(A) - \tfrac{1}{2}(R^M - \lambda R^W)Y(I). \tag{16}$$

In equilibrium, the exogenous return on reserves must equal the deposit rate cost for funding reserves, plus the balance sheet cost, minus the term $\frac{1}{2}(R^M - \lambda R^W)Y(I)$. This term expresses the bank's *net marginal liquidity value of reserves*, which is the difference between the bank's expected liquidity value of an additional reserve, $\frac{1}{2}R^M Y(I)$, and the bank's expected liquidity cost of funding the additional reserve with an additional deposit, $\frac{1}{2}\lambda R^W Y(I)$. The net marginal liquidity value of reserves is positive because a bank's expected liquidity value of an additional reserve, $\frac{1}{2}R^M Y(I)$, is greater than the bank's expected liquidity cost of funding the additional reserve with an additional deposit, $\frac{1}{2}\lambda R^W Y(I)$.

Bank liquidity cost The net marginal liquidity value of reserves represents the value that a bank saves in expected liquidity costs from interbank borrowing. With a one-half probability, a bank has the liquidity shock and pays the marginal interbank cost $Y(I)$ on its interbank borrowing. Since $D^W = \lambda A$, and bank assets increase linearly with reserves in equilibrium, $\frac{dA(M)}{dM} = 1$, the quantity of the shocked bank's interbank borrowing given in equation (13) decreases with a marginal increase in reserves: $\frac{dI(M)}{dM} = -(R^M - \lambda R^W)\mathbf{1}_{I>0}$. Equivalently, the bank's interbank borrowing is reduced by $\frac{d(-I)}{dM}$.

Thus, a marginal increase in reserves reduces a bank's expected marginal interbank borrowing cost by $\frac{1}{2}Y(I)\frac{d(-I)}{dM}$, which we define as the *bank liquidity cost*. Since

$$\frac{1}{2}Y(I)\frac{d(-I)}{dM} = \frac{1}{2}(R^M - \lambda R^W)Y(I), \tag{17}$$

the bank liquidity cost is equal to the bank's net marginal liquidity value of reserves. Next, we define the net cost for banks to hold reserves.

Definition 2 *The net cost of reserves, $C(M)$, is defined as the IOER-deposit rate spread:*

$$C(M) \equiv R^{M^2} - R^D. \tag{18}$$

Proposition 2 *The net cost of reserves is equal to the bank's balance sheet cost minus the bank's liquidity cost:*

$$C(M) = K(A) - \frac{1}{2}Y(I)\frac{d(-I)}{dM}. \tag{19}$$

This result follows directly from substituting from equations (16) and (17) into equation (18). The net cost of reserves represents the net marginal effect of an increase in reserves on bank profits. Since an increase in reserves is funded by an increase in deposits, the competitively determined IOER-deposit rate spread naturally captures this marginal effect on profits.[38] The following lemma establishes the comparative statics result that $C(M)$ is increasing in the central bank's supply of reserves.

Corollary 1 *A bank's net cost of reserves increases with the supply of reserves in the banking system:* $\frac{dC(M)}{dM} > 0$.

This result shows that with an increase in reserves, the increase in balance sheet costs outweighs the decrease in bank liquidity costs. This result follows from the impact of reserves on the balance sheet cost and liquidity cost terms in equation (19).

Since the Fed is the monopoly supplier of reserves, the Fed's choice of reserves determines the equilibrium $C(M)$ for each bank. An increase in the Fed's supply of reserves in the banking system increases the equilibrium size of each bank's balance sheets and hence balance sheet costs.

More plentiful reserves also lower bank liquidity costs. More reserves imply a decrease in the amount of interbank borrowing required to meet liquidity shocks. This results in a lower marginal interbank market cost, $Y(I)$, and lower interbank rate spread to IOER, $R^I - R^M = Y(I)$. The interbank rate spread is positive if there is a positive amount of interbank borrowing in equilibrium: $I > 0$. This occurs if there is at least a partial scarcity of reserves, $M < \bar{M}$, where

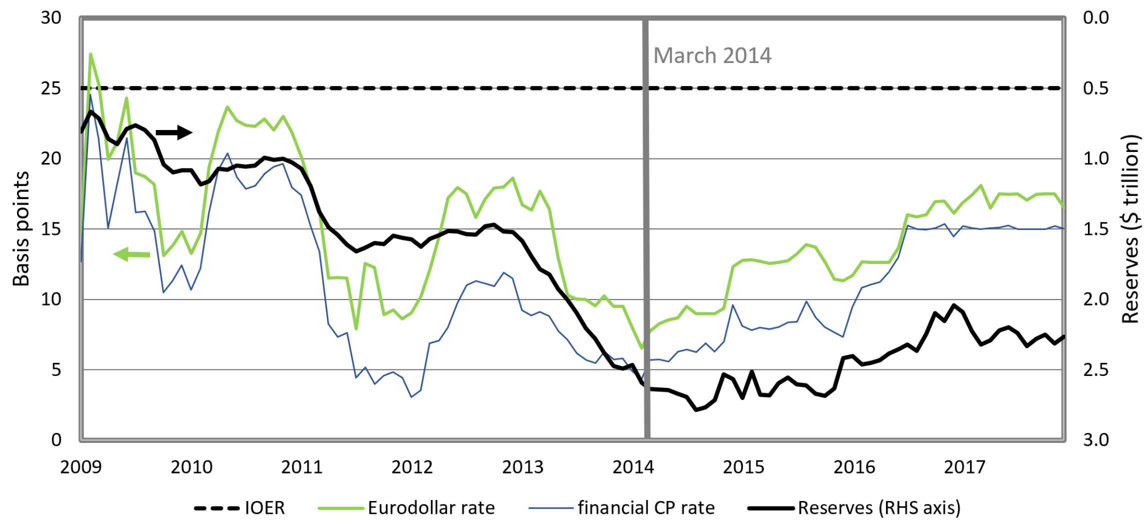
$$\bar{M} \equiv \frac{\lambda R^W (G-B)}{R^M - \lambda R^W}$$

is the threshold amount of reserves in the banking system below which the interbank market is active.

With an overabundance of reserves, $M \geq \bar{M}$, bank liquidity shocks can be met without interbank borrowing. This demonstrates how the large increase in reserves beginning in late 2008 led to the actual interbank component of the fed funds market effectively

[38] An increase in reserves increases one-for-one bank assets and hence bank liabilities, i.e., deposits plus equity. A partial increase in bank liabilities could occur instead if an increase in reserves only leads to a partial increase in bank assets. Such is the case when there is a partial decrease in bank loans caused by a crowding-out effect from reserves. This occurs at the zero lower bound on real deposit rates, when households would store some of their goods at date 0. For simplicity, we do not consider this situation in the paper, as it does not meaningfully alter our results.

Figure 1: Bank borrowing rates and reserves



Rates are measured on the LHS y-axis. Reserves are measured with a reverse scale on the RHS y-axis. Rates and reserves are monthly averages. Bank borrowing rates and IOER are normalized to present the spreads in basis points of borrowing rates to IOER, where IOER is normalized to a constant rate of 25 basis points according to:
$\text{Plotted\_Rate}_{i,t} = \text{Actual\_Rate}_{i,t} - (\text{Actual\_IOER}_t - 25 \text{ basis points})$, for each month (t), and each bank borrowing rate (i).
Hence, the adjustments are effective beginning with the first increase in IOER above 25 basis points in December 2015.
The vertical line at March 2014 represents when the overnight RRP became well established at a regular fixed rate.

disappearing.[39] Without a need for interbank borrowing, the deposit rate decreases below IOER with an increase in reserves driven by the increasing balance sheet cost, $R^D = R^{M^2} - K(A)$.

Figure 1 shows how bank borrowing rates have broadly decreased with the increase in reserves since 2009, and how these borrowing rates have generally fluctuated with changes in reserves.[40] The figure plots for 2009-2017 the monthly average for reserves with a reverse scale on the RHS axis. The LHS axis measures the spreads of the Eurodollar deposit rate

[39] The remaining activity in the fed funds market has consisted primarily of GSE lending to banks. GSEs lend in the fed funds market because they hold reserves but are not eligible for IOER (Bech and Klee, 2011). Banks borrow from GSEs in a similar manner and at similar rates as from depositors. Klee et al. (2016) show that since late 2008, the fed funds rate has provided a weak anchor and transmission mechanism to other short-term funding rates. Gagnon and Sack (2014) highlight that the fed funds rate does not perform a policy role when there is an abundance of reserves. Hence, IOER has effectively replaced the fed funds rate as the marginal rate for bank liquidity.

[40] Figure 1 data sources are, for reserves, "Reserve balances with Federal Reserve Banks" (H41/H41/RESPPLLDD_N.WW), https://www.federalreserve.gov/releases/h41/; for IOER, https://www.federalreserve.gov/datadownload/Choose.aspx?rel=PRates; for the overnight Eurodollar rate, ICAP Capital Markets Eurodollar Rates O/N (EDDR01D Index) obtained from Bloomberg; and for the overnight AA financial commercial paper (CP) rate, https://www.federalreserve.gov/releases/cp/.

and AA financial commercial paper (CP) rate to IOER, which is normalized to a constant 25 basis points when IOER increased beginning in 2015. The vertical line at March 2014 marks a structural break when the fixed-rate overnight RRP became well established after its initial testing at a variety of frequently changing rates ranging from one to five basis points.[41] For a given level of reserves, there is a modest but marked increase in bank borrowing rates after March 2014. However, the correlations of borrowing rates and reserves after March 2014 are consistent with the correlations before March 2014.

Table 1 lists these correlations of the bank borrowing rate spreads to IOER and the quantity of reserves, which range from -0.76 to -0.91. The negative correlations may suggest a role for higher quantities of reserves leading to increased bank balance sheet costs that are pushed onto depositors and commercial paper lenders through lower bank borrowing rates.

Table 1: Correlations between bank borrowing rates and reserves

Bank borrowing rate	Correlation	
	2009 - Feb 2014	Mar 2014 - 2017
Eurodollar rate	-0.76	-0.83
Financial CP rate	-0.80	-0.91

Correlations are between monthly averages for bank borrowing rates and reserves, with rates adjusted as spreads to IOER normalized at 25 basis points as specified in Figure 1.

4.2 Economic welfare and optimal policy

The trade-off banks face in deciding whether or not to expand their balance sheets is also the main driver of overall welfare in our model. More reserves increase bank liquidity and reduce interbank borrowing costs. However, the increase in banks' balance sheets raises equity requirements, which decreases households' liquid assets. Both of these margins affect not only bank profits and equilibrium returns, but also welfare.

The optimal supply of reserves is the quantity that maximizes welfare, which is household expected utility and is equal to the total surplus in the economy. At equilibrium

[41] Source: https://www.newyorkfed.org/markets/opolicy/operating_policy_140225.html.

returns and quantities, this is equal to

$$\mathbb{E}[u^j] = \frac{1}{P_2}\left[P_2\int_0^L r(\hat{L})d\hat{L} + (\theta-1)R^D G - (R^E - R^D)E - \frac{1}{2}(R^{E1} - R^{D1})E_1 - \frac{1}{2}\int_0^I Y(\hat{I})d\hat{I}\right]. \tag{20}$$

The first term is the total production of goods in a sector. The second and third terms represent the household liquidity value $(\theta - 1)$ received on the amount of household endowment G that is invested at date 0 in liquid deposits and bonds but not equity. To see this, substitute for G with $D+B^H+E$ from the household budget constraint in the second term. Substitute for $(R^E - R^D)$ with $(\theta - 1)R^D$, the equity premium, in the third term. The second and third terms of equation (20) simplify to equal the net liquidity value on deposits and bonds, $(\theta - 1)(D + B^H)R^D$, which also reflects that the return on bonds must equal the return on deposits in equilibrium, $R^B = R^D$, to ensure that households are indifferent between holding bonds and deposits. The fourth term of equation (20) is the expected cost of the equity premium on additional equity at date 1.[42] Finally, the last term of equation (20) is the expected interbank market cost.

To maximize welfare, the central bank chooses the optimal supply of reserves, M^*, that maximizes expected household utility $\mathbb{E}[u^j]$ given by equation (20).[43]

A marginal increase in reserves increases the quantities of equity in the third and fourth terms of equation (20) by $\frac{dE(A)}{dM} = \frac{dE(A)}{dA}$ and $\frac{dE_1(A)}{dM} = \frac{dE_1(A)}{dA}$, respectively. This reflects that bank assets directly increase with reserves in equilibrium. A marginal increase in reserves decreases welfare by the amounts $(R^E - R^D)\frac{dE}{dA}$ and $\frac{1}{2}(R^{E1} - R^{D1})\frac{dE_1}{dA}$, which

[42] The new deposit or bond liquid assets that a household acquires at date 1 replace the household's bonds that are sold or deposits that are withdrawn at date 1, respectively, and so do not receive additional liquidity benefits. However, the amount of date 1 bond sales that go toward new date 1 equity decreases the household's liquidity benefit.

[43] To focus on central bank policy for balance sheet size, we do not examine optimal price inflation, P_2, arising from interest rate policy, R^M. The price level is determined as a simple application of the fiscal theory of the price level as in Stein (2012), Cochrane (2005, 2014), Sims (2013), and as distinguished by Freixas et al. (2009, 2011). Examining price inflation would show neo-Fisherian inflationary impacts (Cochrane 2014) from higher nominal interest rates, R^M, since the equilibrium price level $P_2 = \frac{{R^M}^2}{r(L)-\frac{1}{2}R^M y(I)}$ represents a modified version of the Fisher equation with inflation equal to the ratio of the nominal policy rate to the net real marginal return of production in the economy. Instead, we consider the direct welfare effects of the central bank optimizing over reserves, M, where interest rate policy, R^M, holds second-order price and rate effects fixed. Equivalent would be a log-linearization approach using a first-order approximation to a Taylor series expansion of price and rates. Also equivalent would be to assume that rates of interest are additive rather than compounded over the two periods in the model, which reflects the practice for interest on reserves paid by the Fed at biweekly periods rather than daily.

correspond to the two equity terms in equation (20). Thus, increasing reserves decreases welfare by an amount equal to the bank balance sheet cost, which is $K(A) = (R^E - R^D)\frac{dE}{dA} + \frac{1}{2}(R^{E1} - R^{D1})\frac{dE_1}{dA}$.

The expected interbank market cost in the last term of equation (20), $\frac{1}{2}\int_0^I Y(\hat{I})d\hat{I}$, decreases with an increase in reserves. A marginal increase in reserves increases welfare by the amount

$$\frac{d}{dM}\left[-\frac{1}{2}\int_0^I Y(\hat{I})d\hat{I}\right] = -\frac{1}{2}Y(I)\frac{dI}{dM} = \frac{1}{2}Y(I)\frac{d(-I)}{dM} > 0,$$

which is equal to the bank liquidity cost.

Since increasing reserves increases welfare by the amount of the bank liquidity cost and decreases welfare by the amount of the bank balance sheet cost, the optimal quantity of reserves equates these two marginal bank costs. This result is stated in the following proposition.

Proposition 3 *At the central bank's optimal supply of reserves (M^*) for maximizing welfare, $\mathbb{E}[u^j]$, bank liquidity costs equal bank balance sheet costs:*

$$\frac{1}{2}Y(I)\frac{d(-I)}{dM} = K(A). \tag{21}$$

The central bank's optimization problem gives a first order condition that is equivalent to setting the bank's net cost of reserves equal to zero, $C(M) = 0$, which also directly leads to equation (21). By corollary (1), the second order condition is $\frac{dC(M)}{dM} > 0$, which shows that there is an interior optimum for the central bank's choice of M. These results lead to the following proposition.

Proposition 4 *The optimal supply of reserves is a moderate quantity $M^* \in (0, \bar{M})$. At this optimal quantity (M^*), the date 0 net cost of reserve holdings equals zero:*

$$C(M^*) = R^{M^2} - R^D = K(A) - \frac{1}{2}Y(I)\frac{d(-I)}{dM} = 0. \tag{22}$$

This proposition establishes three important implications. First, a positive amount of reserves is always desirable to provide banks with liquid assets to mitigate the cost of interbank trading. Second, in contrast, is that some amount of bank illiquidity and interbank lending is optimal. If the amount of reserves is so large that banks have no liquidity costs, then there is a benefit to decreasing the amount of reserves to mitigate the equity cost from banks' balance sheets. Thus, $M^* \in (0, \bar{M})$ provides an optimal interior solution.

Third, and perhaps most novel, is that the net cost of reserve holdings equals zero. A particularly interesting consequence of this result is that the spread between IOER and the deposit rate, $R^{M^2} - R^D$, is zero under the optimal choice of reserves. This provides a sharp characterization of the optimal supply of reserves in terms of economic variables that are easily observable.

The intuition for this result again lies in weighing the two key frictions in our model. On the one hand, banks must take into consideration the cost of capital caused by equity requirements. If this were the only effect present in our model, then we would obtain the result that the deposit rate would be below IOER regardless of the amount of reserves in the banking system. Once a bank borrows enough deposits to fund the bank's loans to firms, any further amount of deposits lead to a bank accumulating reserves. Banks would demand additional deposits only if the deposit rate is low enough below IOER to cover the marginal capital cost from balance sheet expansion.

Our second friction implies that having more reserves reduces banks' cost of having to borrow in the interbank market. This in turn gives banks an incentive to compete for deposits, consequently bidding up their rate. When the supply of reserves is below the optimum, reserves provide banks a greater net liquidity value than balance sheet cost, which is reflected by a negative net cost of reserves: $C(M < M^*) < 0$. Banks competition for reserves, and hence deposits, leads to the deposit rate above IOER: $R^D > R^M$. This is consistent with the deposit rate at a positive spread above the return on reserves, equal to one, before IOER was introduced in 2008, and when there was a scarcity of reserves: $R^D \geq R^{M^2} = 1$. This also shows that a partial scarcity of reserves, $M < M^*$, is required to maintain positive net deposit rates when IOER is at the zero lower bound.

Since the cost of capital increases with reserves, and the cost of interbank trading decreases with reserves, the supply of reserves should be chosen to equalize these costs at the margin. In our model, this corresponds to the equilibrium relation $R^D = R^{M^2}$, which serves as an optimum policy rule for the supply of reserves. At this point, there is a partial scarcity of reserves that creates a positive spread of the interbank rate above IOER, $R^I - R^M = Y(I)$.

In the context of current Federal Reserve policy, our model gives a sharp prediction regarding the optimal supply of reserves and also how to achieve it. In particular, reserves should be decreased until bank deposit rates increase to the level of IOER. This situation characterizes the equilibrium in which marginal costs arising from bank illiquidity are equated with marginal capital costs. From the perspective of a policymaker, our model shows that observing interest rate spreads can serve as a benchmark for measuring the welfare effects of current policy.

4.3 Reducing overabundant reserves

We can interpret the Fed lowering the current overabundance of reserves within the context of the model by considering at the beginning of date 0 a *starting* quantity of reserves and central bank bond holdings notated by M' and $B^{CB\prime} = M'$, respectively. We define the optimal quantity of bonds held by the Fed as B^{CB*}, which implies that $B^{CB*} = M^*$. Within date 0, the optimal choice of reserves, $M = M^*$, can be implemented by the Fed selling a quantity $\Delta = B^{CB\prime} - B^{CB*}$ of its bonds. Households buy these bonds by withdrawing Δ of deposits. For this deposit withdrawal, the households' banks pay Δ reserves to the Fed. Since banks hold reserves at accounts with the Fed, the banks' payment to the Fed reduces reserves by Δ in the banking system, and these reserves are extinguished. The Fed's outstanding reserves liabilities are decreased by Δ. Hence, as $B^{CB\prime}$ is reduced to B^{CB*}, reserves in the banking system are reduced from M' to M^*.

An equivalent interpretation is for the Fed to reduce reserves from the starting point of M' with the gradual rolling off of bonds from the Fed's balance sheet as bonds mature without the reinvestment of the proceeds. Assume that instead of the Fed selling bonds, $\Delta = B^{CB\prime} - B^{CB*}$ is the amount of bonds the Fed holds that mature and roll off the Fed's balance sheet within date 0. The government issues and sells a new amount of bonds equal to Δ to keep the aggregate supply of bonds constant at B within date 0. Households buy the new bonds with deposits by their banks paying Δ reserves to the government. The government (i.e. the Treasury Department) also holds a reserves account at the Fed. The government pays the Δ reserves received to the Fed for the Fed's Δ maturing bonds. Hence, $B^{CB\prime}$ is reduced to B^{CB*} and M' decreases to M^*, which is an equivalent outcome to the Fed selling bonds directly.

This interpretation method can also be applied to illustrate how the supply of reserves increases from Fed bond purchases, such as during the Fed's LSAP program. From a starting quantity of reserves, the Fed purchases bonds at date 0 from households by creating and paying reserves to the households' banks for credit to the households' deposit accounts. In particular, from a starting quantity of zero reserves and Fed bond holdings, the initial introduction of reserves into the banking system can be interpreted as follows. At the start of date 0, banks can lend to firms, which buy goods from households, who in turn deposit at the banks. These transactions can occur with inside money created by banks and do not require banks to hold reserves. Likewise, households can start by buying all of the bonds sold by the government, with the government using the proceeds to buy goods from households. The Fed then buys bonds from households by paying reserves to the households' banks. Alternatively, the Fed can buy bonds directly from the government by paying reserves to the government's (i.e. Treasury Department's) reserve account at the Fed. The government pays for some of the goods it buys from households by paying its reserves to the households' banks for credit to the household's deposit accounts.

5 Overnight RRP

In this section, we examine the effect of the overnight RRP on rates and reserves. We show that the optimal RRP rate is equal to IOER, which allows the RRP to reduce an overabundance of reserves, enable a higher optimal amount of reserves, and stabilize bank and bond rates. We conclude with a brief extension to analyze the term RRP and TDF.

5.1 Equilibrium effects

Expanding on the model, the central bank can offer the one-period overnight RRP with a return of R^{Qt} and quantity of Q_t at date $t \in \{0, 1\}$ to households.[44] The central bank

[44] Banks would be indifferent or prefer not to invest in the RRP because the equilibrium RRP return is less than or equal to IOER. We show in appendix C that the results of households investing directly in

offers the RRP across sectors either with a fixed-rate and market determined quantity or with an aggregate fixed-quantity supply at a market determined rate. Total central bank liabilities, which are reserves and the RRP, determine the size of central bank assets,

$$B^{CB} = M + Q_t \text{ at date } t \in \{0, 1\}, \tag{23}$$

where quantities are normalized to a per-sector basis. To simplify the analysis, we apply two conditions that hold under the optimal central bank policy: $Q_1 \leq R^W \lambda A$ and $2R^M M > R^W \lambda A$.[45] Formal details of the model and equilibrium with the addition of the overnight RRP are presented in the proof for the following lemma, which shows that the RRP acts as a rate floor for deposits by competing for household investments against bank deposits at dates 0 and 1.

Lemma 1 *The overnight RRP sets a floor on deposit rates at dates 0 and 1:* $R^D \geq R^{Q0} R^{Q1}$ *and* $R^{D1} \geq R^{Q1}$*, which are binding for* $Q_0 Q_1 > 0$ *and* $Q_1 > 0$*, respectively.*

If the central bank uses the overnight RRP with a fixed-quantity supply that varies across dates 0 and 1, the RRP can have an equivalent equilibrium outcome as the overnight RRP with a fixed-rate. We proceed by analyzing the overnight RRP with a fixed-rate that is constant at dates 0 and 1, $R^{Q0} = R^{Q1}$, in which the market determined quantities can vary across dates 0 and 1.

5.2 Reducing overabundant reserves

The overnight RRP rate can be set equal to IOER to efficiently reduce a starting overabundance of reserves to their optimal quantity at date 0. This is an alternative to the central bank directly reducing the supply of reserves through a lower quantity of bonds it holds at date 0.

the RRP are identical in the more institutionally realistic setting of allowing households to hold shares in MMFs, which in turn invest in the RRP and bonds.

[45] The first condition states that the quantity of the overnight RRP is weakly less than the size of the liquidity shock, which precludes the overnight RRP from triggering withdrawals in excess of those caused by the liquidity shock. The second condition states that the supply of aggregate reserves is greater than the size of the liquidity shock, which allows the potential for the overnight RRP to absorb the full liquidity shock.

Proposition 5 *For a starting amount of reserves $M' > M^*$, the optimal supply of reserves can be implemented with the overnight RRP rate set equal to IOER, $R^{Q0} = R^{Q1} = R^M$.*

Consider the economy starting with no RRP and $M' > M^*$, which implies the deposit rate starts at the beginning of date 0 below IOER: $R^{D\prime} < R^{M^2}$. If the RRP is then introduced within date 0 at a rate $R^{Q0}R^{Q1} > R^{D\prime}$, there is a decrease in reserves. Households decrease deposits to invest in the RRP until the deposit rate rises to $R^D = R^{Q0}R^{Q1}$. If the RRP rate is set equal to IOER, the deposit rate rises to IOER, $R^D = R^{M^2}$, and the RRP quantity is $Q_0 = M' - M^*$. The date 0 amount of reserves fall to the optimum, $M = M^*$.

This result highlights that the Fed could more rapidly achieve the optimal level of reserves by using the overnight RRP with a rate set at IOER than by its current normalization strategy of waiting for its assets to mature and roll off its balance sheet. This strategy does not risk disruption in the bond market, since the Fed's quantity of bonds within date 0 is unchanged from the starting quantity $B^{CB\prime}$. This shows that the RRP has no effect on the asset side of the Fed's balance sheet.

Complementing this strategy, the Fed could use a combination of setting the RRP rate to IOER while allowing bonds to run off as they mature. The proposition above implies that there is an indeterminacy in the optimal quantities of the overnight RRP at date 0, Q_0, and of the Fed's bonds, B^{CB}. Applying equation (23), the requirement is that $B^{CB} - Q_0 = M^*$ at the end of date 0. Setting the RRP rate at IOER immediately leads to $Q_0 = M' - M^*$ and a reduction of reserves from M' to M^*. Within date 0, as the Fed's bonds gradually decrease through roll-offs from $B^{CB\prime}$ to B^{CB*}, the RRP quantity gradually decreases from $Q_0 = M' - M^*$ to zero. When $Q_0 = 0$ is reached would signal the appropriate time to end normalization of the Fed's balance sheet and discontinue the run-off of the Fed's assets.

5.3 Increasing the optimal quantity of reserves

Apart or in conjunction with using the overnight RRP to reduce the overabundance of reserves, the overnight RRP can be used with the rate equal to IOER to absorb the full size of bank liquidity shocks at date 1. This capacity increases the optimal supply of reserves and welfare. The liquidity value of reserves is maximized at date 0 when they are equally distributed among banks as buffers against liquidity shocks. However, without the RRP, liquidity shocks lead to an imbalance of reserves beyond what is used for interbank lending at the nonshocked bank, leading to new date 1 assets A_1 and equity requirement E_1.

The overnight RRP can eliminate this date 1 equity by absorbing the liquidity shock. With the optimal supply of reserves, and the RRP rate set at $R^{Q0} = R^{Q1} \in [\bar{R}^{D1}, R^M]$, where $\bar{R}^{D1} \equiv R^M - (\frac{\theta-1}{\theta})R^{\alpha}(A)$, the take-up quantity at the RRP is $Q_0 = 0$ at date 0 and $Q_1 = R^W \lambda A$ at date 1. The date 1 RRP quantity equals the full amount of the liquidity shock. This represents that the RRP can be set at a fixed rate to flexibly absorb liquidity shocks when they occur. The nonshocked bank's new date 1 assets (A_1) and equity (E_1) are zero.

Furthermore, this use of the RRP enables a higher optimal supply of reserves. The optimal central bank policy is determined by the joint quantities of reserves and the overnight RRP, $\mathcal{Q}^{CB} \equiv \{M, Q_0, Q_1\}$, that maximize household expected utility, $\mathbb{E}[u^j]$, from equation (20). M^{**} is designated as the optimal quantity of reserves conditional on the optimal RRP quantity $Q_1 = R^W \lambda A$, which results in welfare equal to

$$\mathbb{E}[u^j] = \tfrac{1}{P_2}\left[P_2 \textstyle\int_0^L r(\hat{L}) d\hat{L} + (\theta - 1)(R^D D + R^B B^H) - \tfrac{1}{2}\int_0^I Y(\hat{I}) d\hat{I}\right]. \tag{24}$$

Equation (24) is modified from welfare without the overnight RRP by setting $E_1 = 0$ in equation (20). Thus, M^{**} is equivalently given by the maximization of equation (24) with respect to reserves.

Proposition 6 *The optimal central bank policy is implemented by the overnight RRP rate equal to IOER, $R^{Q0} = R^{Q1} = R^M$, and an optimal supply of reserves that is greater with the overnight RRP than without it, $M^{**} \in (M^*, \bar{M})$. Bank liquidity costs equal balance sheet costs, $\frac{1}{2}Y(I)\frac{d(-I)}{dM} = K(A)$, and the net cost of reserve holdings equals zero: $C(M^{**}) = R^{M^2} - R^D = 0$.*

Equating marginal bank costs is maintained as the determinant for the optimal supply of reserves with the RRP, as in equation (21) without the RRP. The balance sheet cost $K(A)$ is lower with the RRP than without the RRP through the elimination of new date

1 assets A_1 and hence the expected date 1 required equity. This requires a higher amount of reserves $M^{**} > M^*$ to lower the expected liquidity cost $\frac{1}{2}Y(I)\frac{d(-I)}{dA}$ by an equal amount that the balance sheet cost is reduced. The result $M^{**} > M^*$ reflects that reserves at date 0 are more valuable on net because their impact on the balance sheet cost is lower through the RRP elimination of date 1 equity. Thus, the RRP has the indirect effect of also increasing bank liquidity through a higher optimal supply of reserves. The policy rule for the optimal supply of reserves, $R^D = R^{M^2}$, is maintained. Starting from an overabundance, reserves do not need to be decreased by as much for R^D to rise up to IOER since $M^{**} > M^*$.

The optimal quantity of reserves and the overnight RRP imply a moderate optimal size for the central bank balance sheet of $B^{CB} = M^{**}$. The composition of the central bank's liabilities fluctuate between a quantity M^{**} of reserves at date 0 when there is no liquidity shock, and a quantity $M^{**} - Q_1$ of reserves and $Q_1 = R^W \lambda A$ of the overnight RRP at date 1 when there is a liquidity shock. Our results contrast with the current literature, which primarily advocates either for a very large or a very small size of the Fed's balance sheet. Our model instead points to a moderate Fed balance sheet size reflected by the optimal moderate quantity of reserves and the overnight RRP.

5.4 Stabilizing rates

The overnight RRP with a rate set at IOER also stabilizes volatile rates that arise when a volatile size of the bank liquidity shock is introduced. The model is extended by redefining λ as a random variable $\tilde{\lambda}$ that takes a realization, λ^i, where $i \in \{h, l\}$ corresponds to a relatively high or low shock state, with $\lambda^h > \lambda^l$ and $\mathbb{E}[\tilde{\lambda}] \in (\lambda^l, \lambda^h)$. At date 0, the central bank chooses its policy for the RRP, along with reserves, before the realization of $\tilde{\lambda}$ at date 1. We also show that the RRP can similarly stabilize the price and hence one-period return on bonds. The one-period holding return on bonds between dates 0 and 1 is P_1^B and between dates 1 and 2 is $\frac{R^B}{P_1^B}$, where $R^{B1} \equiv \frac{R^B}{P_1^B}$ is defined as the date 1 bond return.

The RRP with a rate $R^{Q0} = R^{Q1} \in [\bar{R}^{D1}, R^M]$ has an equilibrium take-up of the full amount of the liquidity shock $R^W \lambda^i A$, regardless of the shock size state $i \in \{h, l\}$. As a consequence, the equilibrium rates on deposits and bonds acquired at date 1 equal R^{Q1}.

In contrast, without the RRP, these equilibrium date 1 deposit and bond rates equal $R^{D1} = R^{B1} = R^M - (\frac{\theta-1}{\theta})R^{\alpha}(A + 2R^W \lambda^i A)$, which is lower in the high shock state with λ^h than in the low shock state with λ^l.

Proposition 7 *The overnight RRP with rate equal to IOER implements a constant one-period rate at dates 0 and 1 for deposits and bonds equal to IOER:* $R^W = P_1^B = R^M$ *and* $R^{D1} = R^{B1} = R^M$ *for both shock size states* $i \in \{h, l\}$*. In contrast, with no RRP, the date 1 deposit and bond rate is volatile with a higher rate in the low shock state than in the high shock state.*

In addition, with random shock sizes, the optimal welfare with M^{**} reserves is maintained with the RRP rate set at IOER since the RRP eliminates required date 1 equity in both the high and low shock states.

The Fed originally tested the overnight RRP using a fixed-quantity supply before testing and then adopting the fixed-rate implementation. The following corollary reflects that with a random shock size, the RRP can only stabilize rates using a fixed-rate rather than fixed-quantity implementation.

Corollary 2 *The overnight RRP with a fixed-quantity supply cannot implement a constant date 1 deposit and bond rate.*

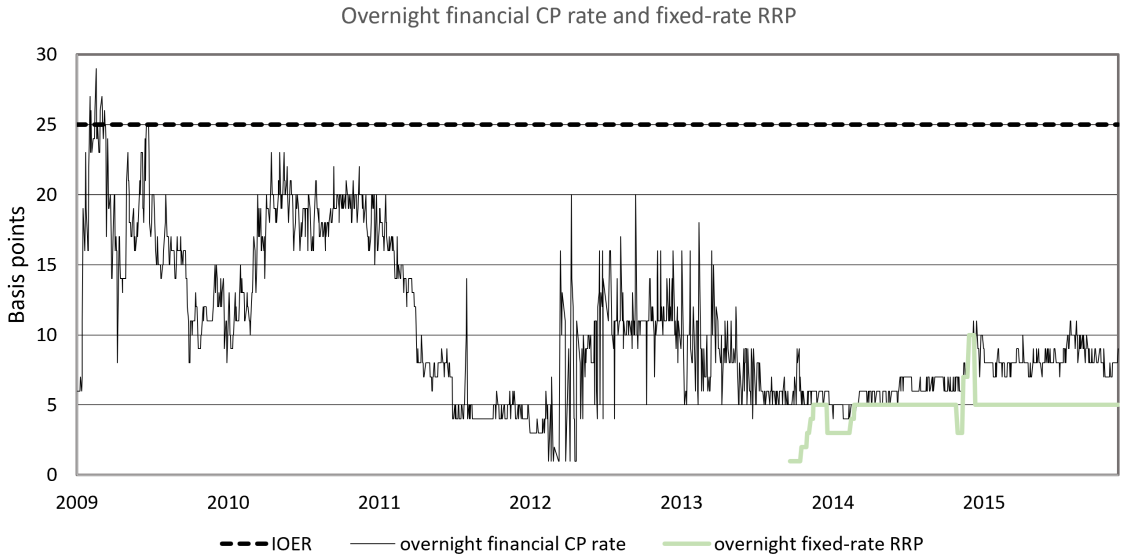
For any fixed-quantity, Q_1, the date 1 deposit and bond rates equal the equilibrium RRP rate and are greater in the low state than high state: $R^{D1}(\lambda^l) > R^{D1}(\lambda^h)$.

Figure 2 illustrates that the fixed-rate overnight RRP greatly reduced volatility and acted as a strong floor for overnight bank funding rates as exemplified by the overnight AA financial CP rate, which is seen as the lowest of the bank borrowing rates in Figure 1 and is the most volatile of these rates when measured at higher frequencies.[46]

Figure 2 plots the daily overnight financial CP rate starting in 2009 and the daily overnight RRP starting in September 2013, when the RRP was switched from a fixed-quantity to a fixed-rate implementation. The two series are plotted through December

[46] Figure 2 data sources are, for IOER, https://www.federalreserve.gov/datadownload/Choose.aspx?rel=PRates; for the overnight AA financial commercial paper (CP) rate, https://www.federalreserve.gov/releases/cp/; and for the overnight fixed-rate RRP, https://apps.newyorkfed.org/markets/autorates/tomo-search-page.

Figure 2: Overnight financial CP rate and overnight fixed-rate RRP



The overnight financial CP rates on the last day of each month are excluded because of window-dressing effects.

2015, when IOER was first increased as part of the FOMC's series of interest rate hikes and the overnight RRP rate spread below IOER was increased to 25 basis points. The figure shows that the fixed rate on the overnight RRP was set at five basis points for most of the time period, although at brief times it was set at rates ranging from one to ten basis points for testing purposes.

We observe that the volatility of the overnight CP rate is strikingly reduced during the period when the overnight fixed-rate RRP is in place. We also observe that the overnight CP rate rarely falls below the RRP rate. Whereas, prior to the overnight RRP facility, the overnight CP rate repeatedly falls to one basis point, an effective zero lower bound, from February through April 2012.

While the Fed's exercises with the overnight RRP led to confidence of it acting as a floor on funding rates, the Fed's current normalization plan is to eventually phase out its use. We argue for the continued use of the overnight RRP. Beyond acting as a floor on overnight rates, the overnight RRP stabilizes the volatility of overnight rates, reduces bank balance sheet costs, and enables a higher optimal quantity of reserves.

Additional advantages of the overnight RRP likely exist that are outside the formal benefits we consider in this model. Lower volatility and uncertainty of overnight funding rates and quantities can support a more elastic demand for short-term assets more broadly, which may further help transmit the stabilization effect of the RRP to other money market assets. Klee et al. (2016) shows empirically that the fixed-rate overnight RRP has strongly reduced the volatility and increased the co-movement of overnight funding rates. These results complement the Duffie and Krishnamurthy (2016) findings that the overnight RRP provides a better transmission of monetary policy than IOER alone, which is the basis for their advocation for the continued use of the overnight RRP.

5.5 TDF and term RRP

The Fed has not adopted the term RRP or TDF for policy but has retained the option of using these supplementary tools as needed. We analyze these tools by the central bank offering them in an analogous manner as the overnight RRP with the following differences. The term RRP and TDF are offered as two-period assets at date 0, and the TDF can only be held by banks, as in practice. The term RRP has a return of R^{TM} and quantity of Q^{TM}, while the TDF has a return of R^{TD} and quantity of Q^{TD}. The central bank's budget constraint is $B^{CB} = M + Q_t + Q^{TM} + Q^{TD}$ at date $t \in \{0, 1\}$.

The term RRP acts equivalently to the overnight RRP with a restriction that date 0 and date 1 rates and quantities are equal. Thus, the term RRP with a rate set to IOER can efficiently reduce a starting overabundance of reserves to their optimal level. But the term RRP cannot stabilize rates or implement optimal welfare with the optimal supply of reserves M^{**}, which requires a greater RRP quantity at date 1, to absorb bank liquidity shocks, than at date 0.

The TDF is inefficient for reducing a starting overabundance reserves. The overnight or term RRP reduces banks' date 0 balance sheet size in equal measure to the reduction of reserves, whereas the TDF does not reduce bank size since the TDF replaces reserves as an asset on banks' balance sheets. Thus, the TDF lowers welfare because it inefficiently increases bank liquidity costs without reducing balance sheet costs. The TDF also does not absorb liquidity shocks to support a higher optimal supply of reserves or stabilize rates.

6 Concluding remarks

In October 2017, the Fed began the normalization of its balance sheet guided by a cap on the run-off of its bonds that gives a rough pace of reduction but no determined end point. The eventual new-normal quantity of reserves and the use of the overnight RRP are open questions that have been publicly debated primarily between advocates for either a large or small Fed balance sheet size.

We analyze the optimal Fed balance sheet based on the impact of reserves and the overnight RRP on the banking system. Our model provides a sharp result that the optimal supply of reserves is determined by equating their impact at the margin on bank liquidity costs and balance sheet costs. We derive an optimal policy rule, which states that reserves should be reduced until the bank deposit rate rises to equal IOER, at which point the Fed should end its asset run-offs. Since this end-point will not be reached for a few years, we show that the Fed could more expediently reduce reserves to their optimum by using the overnight RRP with its rate raised to equal IOER. We also demonstrate that the Fed should establish the overnight RRP as a permanent policy tool rather than end its use as currently planned. With the overnight RRP rate set at IOER, it increases the optimal quantity of reserves by absorbing bank liquidity shocks, which reduces bank balance sheet costs, and it stabilizes overnight interest rates.

A moderate quantity of reserves and the overnight RRP implies that a moderate size of the Fed balance sheet is optimal. A likely reason that a moderate size has not yet been given more attention is that such a size has not been previously used by central banks in practice or studied in the academic literature. Historically, central banks operated using a *corridor* or *channel* system without paying interest on reserves. Maintaining positive (net) interest rates required central banks to create extreme scarcity of reserves. The advent of a *floor* system with relatively large central bank balance sheets did not occur until 2006, starting with the Reserve Bank of New Zealand. Positive rates can be supported when there is a large supply of reserves by paying interest on reserves. The Fed and other central banks only expanded their balance sheets with large quantities of reserves starting in 2008 as a by-product of large liquidity operations during the financial crisis and large asset purchases in response to weak economies.

The academic study of central bank balance sheet policy was traditionally based on a corridor system following Poole (1968). The central bank inelastic supply of reserves intersects with banks' downward-sloping aggregate demand for reserves in the region of a small quantity of reserves and a very active interbank market. Goodfriend (2002) originates the study of a floor system, which corresponds to banks' elastic demand at the interest rate on reserves in the region of a large quantity of reserves and an inactive interbank market.

Our solution of a moderate quantity of reserves is unique in the literature. This quantity corresponds to the small region where banks' demand for reserves is downward-sloping, just before it kinks and is inelastic. This region is characterized by a moderately active interbank market and has not been previously considered in the literature or in central bank practise. This result introduces a novel paradigm for future research on the analysis and optimality of central bank policy more broadly. Our findings also provide a new system for the practice of central bank policy, with predictions and interpretations of the relevant short-term interest rates and relative spreads. Such research will continue to shape the balance sheet policy decisions faced by the Federal Reserve over the next few years, as well as soon to be faced by other central banks with recent massively expanded balance sheets, including the ECB, Bank of England, and Bank of Japan.

Appendix A: Proofs

Proof of Proposition 1. Necessary first order conditions and sufficient second order conditions hold for L in the firm optimization (3); $\mathcal{Q}^{Bj}$ in the bank optimization (8); and $\mathcal{Q}^{Hj}$ in the household optimization (2). With market clearing, and equal date 0 asset holdings across sectors by ex-ante symmetry, all constraints bind for the agents with the exception of $B_1^n \leq B^H$, $M_1^n \geq 0$, and potentially $M_1^s \geq 0$. Binding household date 1 budget constraints give $D_1 + E_1 = R^W D^W$. Market clearing for the interbank market and reserves at date 1 imply $A_1 = R^W \lambda A$; hence, since $D^W \equiv \lambda A$, we have $A_1 = D_1 + E_1$.

In the bank optimization (8), first order conditions with respect to L, M, I^j, D, E, D_1 and E_1, with binding constraints and market clearing, give $I = (R^W \lambda A - R^M M)^+$ and the following returns:

$$
\begin{aligned}
R^{D1} &= R^M - (\tfrac{\theta-1}{\theta}) R^\alpha (A + 2A_1) && (25)\\
R^I &= R^M + Y(I) && (26)\\
R^L &= R^{M^2} + \tfrac{1}{2} R^M Y(I) && (27)\\
R^L &= (1 - \tfrac{1}{2}\lambda) R^D + \tfrac{1}{2}\lambda R^W R^I + (\tfrac{\theta-1}{\theta}) R^\alpha (2A + \tfrac{1}{2} A_1). && (28)
\end{aligned}
$$

Substituting for R^L from equation (27) into the first order condition with respect to L for the firm optimization (3) gives $P_2 = \frac{R^{M^2}}{r(L) - \frac{1}{2} R^M y(I)}$.

In the household optimization (2), first order conditions with respect to D, E, D_1, E_1 and B_1^n, with binding constraints and market clearing, give $B_1 = R^W D^W$, $P_1^B = R^W$, and the returns $R^B = R^D$, $R^{D1} = \frac{R^B}{P_1^B}$, $R^E = \theta R^D$, and $R^{E1} = \theta R^{D1}$. Market clearing and binding bank and household constraints give $A = G - B + M$.

Proof of Proposition 2. The results follows directly from substituting for R^{M^2} from equation (16) into equation (18).

Proof of Corollary 1. Since $A = L + M$, $\frac{dA(M)}{dM} = 1$. This allows for solving

$$
\begin{aligned}
\frac{dK(A)}{dM} &= \frac{dK(A)}{dA} = (\tfrac{\theta-1}{\theta}) R^\alpha [2 + (1 + \lambda R^W)\lambda R^W] > 0\\
\left.\frac{dI(M)}{dM}\right|_{R^M M \geq R^W \lambda A} &= 0\\
\left.\frac{dI(M)}{dM}\right|_{R^M M < R^W \lambda A} &= -(R^M - \lambda R^W).
\end{aligned}
$$

To evaluate $\frac{dC(M)}{dM} = \frac{d}{dM}[K(A) - \frac{1}{2}(R^M - \lambda R^W)Y(I)]$ for $R^M M \geq R^W \lambda A$,

$$\begin{aligned} \frac{dY(I)}{dM} &= \frac{dY(I)}{dI}\frac{dI(M)}{dM} = 0 \\ \frac{dC(M)}{dM} &= (\tfrac{\theta-1}{\theta})R^\alpha[2 + (1+\lambda R^W)\lambda R^W] > 0. \end{aligned}$$

For $R^M M < R^W \lambda A$,

$$\begin{aligned} \frac{dY(I)}{dM} &= \frac{dY(I)}{dI}\frac{dI(M)}{dM} = -(R^M - \lambda R^W)\frac{dY(I)}{dI} \\ \frac{dC(M)}{dM} &= (\tfrac{\theta-1}{\theta})R^\alpha[2 + (1+\lambda R^W)\lambda R^W] + \tfrac{1}{2}(R^M - \lambda R^W)^2\frac{dY(I)}{dI}, \end{aligned}$$

which implies that $\frac{dC(M)}{dM} > 0$ since $Y'(I) > 0$.

Proof of Proposition 3. The central bank's optimization is $\max_M \mathbb{E}[u^j]$. Since $\frac{dA(M)}{dM} = 1$, the first-order condition with respect to M is equivalent to the first-order condition with respect to A, which is

$$-(\tfrac{\theta-1}{\theta})\frac{d\left(R^E E\right)}{dA} - \tfrac{1}{2}(\tfrac{\theta-1}{\theta})\frac{d\left(R^{E1}E_1\right)}{dA} + \tfrac{1}{2}Y(I)\frac{d(-I)}{dA} = 0.$$

Since $\frac{\theta-1}{\theta} = \frac{R^E - R^D}{R^E} = \frac{R^{E1}-R^{D1}}{R^{E1}}$, the first two terms of the above equation equal $-K(A)$. Hence, the first-order condition gives the result in equation (21).

Proof of Proposition 4. Since $\frac{d(-I)}{dM} = \frac{d(-I)}{dA} = (R^M - \lambda R^W)$ for $M < \bar{M}$, and $Y(I) = 0$ for $M \geq \bar{M}$ following from $I(M \geq \bar{M}) = 0$, we can write

$$\tfrac{1}{2}Y(I)\frac{d(-I)}{dA} = \tfrac{1}{2}(R^M - \lambda R^W)Y(I).$$

Substituting for $\frac{1}{2}(R^M - \lambda R^W)Y(I)$ with $\frac{1}{2}Y(I)\frac{d(-I)}{dA}$ into equation (22) gives $C(M^*) = K(A) - \frac{1}{2}Y(I)\frac{d(-I)}{dA}$, from which $C(M^*) = 0$ follows directly from lemma 1.

To establish that $M^* < \bar{M}$, note that since $Y(I) = 0$ for $M \geq \bar{M}$, $C(M \geq \bar{M}) = K(A) > 0$. Since $\frac{dC(M)}{dM} > 0$, $M^* < \bar{M}$ is required for $C(M^*) = 0$. To establish that $M^* > 0$, note that for $R^M = 1$ and $R^D \geq 1$ for an arbitrarily small amount of reserves $\check{M} > 0$, as pre-IOER, implies that $C(\check{M}) = 1 - R^D \leq 0$. Since $\frac{dC(M)}{dM} > 0$, as reserves decrease from $\check{M}$ to zero, $C(M)$ decreases and hence $C(0) < 0$. Thus, $M^* > 0$ is required for $C(M^*) = 0$.

As discussed in footnote (43), an equivalent approach to holding second-order price and rate effects fixed is to instead use a log-linearization with a first-order approximation to a Taylor series expansion of price and rate effects, which results in a central bank objective function

$$\mathbb{E}[u^j] = \int_0^L r(\hat{L})d\hat{L} + (\theta - 1)W - (\tfrac{\theta-1}{\theta})\left(R^E E + \tfrac{1}{2}R^{E1}E_1\right) - \tfrac{1}{2}\int_0^I y(\hat{I})d\hat{I} \tag{29}$$

to replace the expression in equation (20). The alternative equivalent approach of assuming that one-period rates of interest are additive rather than compounded over the two periods in the model also results in the central bank objective function above. The first order condition for reserves of this objective function gives the same result, $\frac{1}{2}Y(I)\frac{d(-I)}{dA} = K(A)$, for the optimal level of reserves, M^*, which leads to the same result that $C(M^*) = 0$.

Proof of Lemma 1. The overnight RRP is incorporated into the model by redefining the household dates 0 and 1 budget constraints as $D + E + B^H + Q_0 \leq G$, $P_1^B B_1^s + Q_1^s \leq R^W D^W + R^{Q0}Q_0$, and $D_1 + E_1 + Q_1^n \leq P_1^B B_1^n + R^{Q0}Q_0$; adding $\mathbf{1}_{[j=s]}\theta R^{Q1}Q_1^s$ and $\mathbf{1}_{[j=n]}\theta R^{Q1}Q_1^n$ to the household profit Π^{Hj} in equation (1). The budget constraints are redefined for the central bank equation (9) as $B^{CB} = M + Q_0$ for date 0 and $B^{CB} = M + \frac{1}{2}Q_1^s + \frac{1}{2}Q_1^n$ for date 1, and for the government equation (10) as $\Upsilon = R^B(B - B^{CB}) + R^{M^2}M + R^{Q1}Q_1$, and the central bank's choice variables are M, Q_0, and Q_1.Added to the definition of an equilibrium for a fixed-quantity RRP are the returns $R^{Q0} > 0$ and $R^{Q1} > 0$; and market clearing for Q_0 at date 0 for household demand and central bank supply, and $\frac{1}{2}(Q_1^s + Q_1^n) = Q_1$ at date 1.

Market clearing, binding budget constraints, and equal initial asset holdings across sectors give $A = W - B + M + Q_0$, $A_1 = R^W\lambda A - Q_1$. First order conditions give $R^W \geq R^{Q0}$ and $R^{D1} \geq R^{Q1}$, which bind for $Q_0 > 0$ and $Q_1 > 0$, respectively.

The equilibrium definition is revised for a fixed-rate RRP at date $t = 0$, date $t = 1$, or both dates $t \in \{0, 1\}$, by setting R^{Qt} as an exogenous choice by the central bank rather than an equilibrium variable. The equilibrium outcome with a fixed-rate RRP at dates 0 and/or 1 is equivalent to the equilibrium outcome with a fixed-quantity RRP since there exists a one-to-one mapping between the equilibrium R^{Qt} and Q_t for all cases.

Proof of Proposition 5. For starting reserves $M' > M^*$ without the RRP, $R^D > R^{M^2}$ and $B^{CB\prime} = M'$. Following from the proof of lemma 1, with the RRP at rate $R^{Q0} = R^{Q1} = R^M$, $R^D = R^{M^2}$, $Q_0 = M' - M^*$, and hence $M = B^{CB'} - Q_0 = M^*$.

Proof of Proposition 6. The maximization of $\mathbb{E}[u^j]$ over $\mathcal{Q}^{CB}$ subject to the central bank budget constraints given in the proof of lemma 1 give the first order conditions $Q_1 = R^W \lambda A$ and $M = M^{**}$. The derivative of $\mathbb{E}[u^j]$ with respect to M with $Q_1 = R^W \lambda A$ is greater than the corresponding derivative of $\mathbb{E}[u^j]$ with $Q_1 = 0$ because of the omission of the $-\frac{1}{2}R^{E1}E_1$ term, which gives $M^{**} > M^*$. The derivative of $\mathbb{E}[u^j]$ with $Q_1 = R^W \lambda A$ is negative for $M \geq \bar{M}$, which gives $M^{**} < \bar{M}$.

Equating the solutions for R^L from the first order conditions with respect to A and M in the bank optimization and applying the equilibrium conditions in lemma 1; solving for R^D; and substituting for $K(A) = \frac{1}{2}Y(I)\frac{d(-I)}{dA}$ given by the first order condition with respect to M for the optimization of $\mathbb{E}[u^j]$, gives $C(M^{**}) = R^{M^2} - R^D = 0$.

Proof of Proposition 7. First, note that $R^{B1} \equiv \frac{R^B}{P_1^B} = R^{D1}$ is shown in the proof for proposition 1. Next, replace λ by $\psi\lambda^h + (1-\psi)\lambda^l$ for a generic probability $\psi \in (0,1)$ in the agent optimizations and market clearing conditions of the model. For all $\lambda^i \in (0,1)$ for $i \in \{h,l\}$ such that $\lambda^h > \lambda^l$, and for any choice of $R^{Q1} \in [\underline{R}^{D1}, \bar{R}^{D1}]$, we have $R^{D1} = R^{B1} = R^{Q1}$, with $Q_1 = R^W \lambda^i A$, for $i \in \{h,l\}$; whereas, with no RRP ($Q_1 = 0$), we have $R^{D1}(\lambda^h) = R^{B1}(\lambda^h) < R^{D1}(\lambda^l) = R^{B1}(\lambda^l)$.

Proof of Corollary 2. Following from the proof of proposition 6, for any $\lambda^i \in (0,1)$ for $i \in \{h,l\}$ such that $\lambda^h > \lambda^l$, and for any Q_1 that is not conditional on the state $i \in \{h,l\}$, we have $R^{D1}(\lambda^h) = R^{B1}(\lambda^h) < R^{D1}(\lambda^l) = R^{B1}(\lambda^l)$.

Appendix B: Bank moral hazard and capital requirements

At dates 0 and 1, a bank can unobservably take a risky project with a negative expected NPV and a realized marginal return of either $R^{\alpha}(\cdot) > 0$ or $\beta(\cdot) < 0$ with equal probability. $|\beta(\cdot)|$ is sufficiently large that depositors bear a partial loss and lose their liquidity value on deposits such that $\theta = 1$, which implies that the risky project is a form of bank risk-shifting that is socially inefficient. The government has the option of providing a bail-out to depositors by using a lump sum tax on all households to pay for the loss on deposits at a bank that has a negative realization of the risky project.

The bank chooses whether to take the risky project at dates 0 and 1 in order to maximize profit subject to the expected return required by equityholders. Specifically, if the bank takes the risky profit, it pays equity a sufficient return to equity when there is a positive realization to compensate for the loss to equity when there is a negative realization. Since equity has a zero return if there is a negative realization, equity receives an additional return of $R^E E$ ($R^{E1} E_1$) if there is a positive realization for a project taken at date 0 (date 1). Hence, the profit for bank $j \in \{n, s\}$ conditional on having a positive realized return from risk-shifting on its new date 0 assets A is

$$\Pi^{Bj,RS_0} = \Pi^{Bj} + R^{\alpha}(A)A - R^E E, \tag{30}$$

or on its new date 1 assets A_1 is

$$\Pi^{Bj,RS_1} = \Pi^{Bj} + R^{\alpha}(A + A_1)A_1 - R^{E1} E_1. \tag{31}$$

The first term of equations (30) and (31) is the bank's profit Π^{Bj} from equation (6) if the bank does not take the risk-shifting project at either dates 0 or 1. The second terms of equations (30) and (31) are the returns $R^{\alpha}(A)A$ and $R^{\alpha}(A + A_1)A_1$ from a positive realization of risk-shifting at dates 0 and 1, respectively. The third terms of equations (30) and (31) subtract the additional return paid to equity $R^E E$ and $R^E E_1$ for positive realizations.

Since risk-shifting is socially inefficient, the government as regulator imposes an equity capital requirement that incentives banks not to take the risky project. The constraint for a bank not to take the risk-shifting project at date $t \in \{0, 1\}$ is

$$\mathbb{E}_t[\Pi^{Bj,RS_t}] \leq \mathbb{E}_t[\Pi^{Bj}] \quad \text{for } t \in \{0, 1\}. \tag{32}$$

Note that we could add an additional constraint for the bank not to take the risk-shifting project at both dates 0 and 1, but that would be redundant as fulfillment of the no risk-shifting constraints for dates 0 and 1 individually guarantees fulfillment of such an additional constraint.

Substituting for Π^{Bj,RS_0}, Π^{Bj,RS_1} and Π^{Bj} from equations (30), (31), and (6), respectively, equation (32) with a binding inequality gives the minimum capital requirements to prevent risk-shifting as

$$\begin{aligned} \underline{E}(A) &\equiv \tfrac{R^{\alpha}(A)A}{R^E} \\ \underline{E}_1(A, A_1) &\equiv \tfrac{R^{\alpha}(A+A_1)A_1}{R^{E1}}, \end{aligned}$$

which is the result given in equation (7). Setting the capital requirement at the minimum to prevent risk-shifting at each date 0 and 1 minimizes welfare costs. Equity issued at date 0 decreases date 0 deposits and is more costly in expectation for welfare as well as for bank profits than equity issued at date 1, where the measure of welfare is given by equation (20). Date 1 equity is only necessary for the nonshocked bank with date 1 inflows. Thus, reducing expected date 1 required equity with additional date 0 equity is inefficient.

The capital requirement is a source of economic inefficiency, since equity does not provide the liquidity value of deposits, but it acts as a constrained-efficient mechanism to solve the time-inconsistency problems for the bank and government. If the project was verifiable, complete contracts for deposits would provide costless commitment for the bank not to take the project without requiring costly equity.

If the government could ex-ante commit against bailouts to depositors, banks would issue equity equal to the capital requirement as a market discipline commitment device to avoid otherwise facing higher deposit rates from the loss of liquidity value. Without commitment, the government imposes the capital requirement in place of market-discipline based equity. Alternatively, the government could ensure the depositor liquidity benefit by providing deposit insurance. However, this is less efficient than capital requirements because it would not provide the bank a commitment device against taking the risky project. This is true even if the bank were charged the expected government cost of the deposit insurance, unless such charges could be contingent on verifiability of the bank's actual risk-taking.

Appendix C: Money market funds

The MMF is incorporated into the model including the overnight RRP, which is presented in Section 5 and detailed in the proof for lemma 1. There are competitive MMFs that operate across both sectors represented by a single price-taking MMF. At date 0, the MMF buys B^{MF} bonds, invests Q_0^{MF} in the date 0 overnight RRP, and issues S shares to households. At date 1, the MMF invests Q_1^{MF} in the date 1 overnight RRP, redeems S^W shares and issues S_1 new shares. MMF bond and share quantities are normalized to a per-sector basis. Instead of transacting in bonds, households acquire and redeem MMF shares. (Allowing households to transact both in bonds and MMF shares is equivalent).

Specifically, households buy S shares instead of buying B^H bonds and investing Q_0 in the overnight RRP at date 0. The shocked household uses its early withdrawal of D^W to buy S_1 additional shares instead of buying B_1^s bonds and investing Q_1^s in the overnight RRP at date 1. The nonshocked household redeems S^W shares instead of selling B_1^n bonds. Returns at date 2 are R^S for shares issued at date 0 and R^{S1} for shares issued at date 1. The return on shares issued at date 0 and redeemed early at date 1 is R^{SW}. Households receive the liquidity benefit factor θ on MMF shares, as with deposits and bonds. In particular, the date 0 MMF shares provide a promised redemption return R^{SW}, which parallels the return R^W on early deposit withdrawals.

The MMF profit is

$$\Pi^{MF} = R^B B^{MF} + R^{Q1} Q_1^{MF} - R^S(S - S^W) - R^{S1} S_1. \tag{33}$$

At date 2, the MMF receives the R^B return on its B^{MF} bonds bought at date 0. The MMF pays the return R^S on the $(S - S^W)$ outstanding shares issued at date 0 and the return R^{S1} on shares issued at date 1.

The MMF maximizes its expected profit, Π^{MF}, as follows:

$$\left.\begin{aligned} &\max_{\mathcal{Q}^{MF}} \mathbb{E}[\Pi^{MF}] \\ &\text{s.t.} \;\; B^{MF} + Q_0^{MF} \leq S \\ &\qquad R^{SW} S^W + Q_1^{MF} \leq S_1 + R^{Q0} Q_0^{MF}, \end{aligned}\right\} \tag{34}$$

where $\mathcal{Q}^{MF} \equiv \{B^{MF}, S, S^W, S_1, Q_0^{MF}, Q_1^{MF}\}$. The two inequalities are the MMF budget constraints for dates 0 and 1, respectively. The first inequality states that the amount B^{MF} the MMF pays to buy bonds is limited to the amount S received from issuing shares at date 0. The second inequality states that the amount $R^{SW}S^W$ paid for early redemptions is limited to the amount S_1 received from issuing shares at date 1.[47]

The model is updated with the substitution of S for $(B^H + Q_0)$, B^{MF} for B^H, Q_0^{MF} for Q_0, Q_1^{MF} for $\frac{1}{2}(Q_1^s + Q_1^n)$, S^W for B_1^n, $\frac{S_1}{R^{SW}}$ for B_1^s, and R^{SW} for P_1^B. The household profit Π^{Hj} in equation (1) and optimization in equation in (2) are updated, respectively, as

$$\begin{aligned} \Pi^{Hj} \;\; = \;\; & R^E E + \Pi^{Bj} + \Pi^{Fj} - \Upsilon \\ & + \mathbf{1}_{[j=s]} \theta [R^D (D - D^W) + R^S B^S + R^{S1} S_1)] \\ & + \mathbf{1}_{[j=n]} \{\theta [R^D D + R^{D1} D_1 + R^S (B^S - S^W)] + R^{E1} E_1\}. \end{aligned}$$

$$\left. \begin{aligned} & \max_{\mathcal{Q}^{Hj}} \mathbb{E}[u^j] \\ & \text{s.t.} \;\; D + E + S \leq G \\ & \qquad S_1 \leq R^W D^W \\ & \qquad D_1 + E_1 \leq R^{SW} S^W \\ & \qquad S^W \leq S. \end{aligned} \right\} \tag{35}$$

The definition of an equilibrium is updated as follows. The returns $(R^S, R^{SW}, R^{S1}) > 0$ are added. The MMF optimizing quantities $\mathcal{Q}^{MF}$ given by (34) are added to the agents' optimizing quantities. Market clearing is added for MMF shares at date 0, S. Market clearing for bonds at date 1 is replaced by market clearing for MMF shares at date 1, $S_1^s = S_1$.

Proposition 8 *The returns on MMF shares are equal to the corresponding returns on deposits and bonds for equivalent one- or two-period holding periods:* $R^{SW} = R^W = P_1^B$, $R^{S1} = R^{D1} = \frac{R^B}{P_1^B}$, *and* $R^S = R^D = R^B$. *The MMF quantities are equal to the corresponding quantities without the MMF:* $S = B^H + Q_0$, $B^{MF} = B^H$, $Q_0^{MF} = Q_0$, $Q_1^{MF} = \frac{1}{2}(Q_1^s + Q_1^n)$, $S^W = B_1^n$, *and* $S_1 = P_1^B B_1^s$. *The MMF makes zero profits, and the results of the paper are unchanged.*

[47] Note that MMFs are not required to issue equity to create money-like assets with liquidity benefits, as MMFs invest in government bonds and the overnight RRP without the ability for risk-shifting. Stein (2012) makes this essential point by referring to the "more benign forms of money creation, for example, money market fund accounts backed exclusively by Treasury bills."

Proof. Necessary first order conditions and sufficient second order conditions hold for $\mathcal{Q}^{MF}$ in the MMF optimization (34) and the revised household optimization (35). With market clearing and and symmetric date 0 asset holdings; the results for corresponding equilibrium returns and quantities follow directly from the binding budget constraints and first order conditions with respect to S, D, S_1 and D_1 in the revised household optimization (35); the binding budget constraints and first order conditions with respect to $\mathcal{Q}^{MF}$ in the MMF optimization (34); the term structure on deposit returns $R^D = R^W R^{D1}$; and the equilibrium results $R^B = R^D$, $R^{D1} = \frac{R^B}{P_1^B}$, and $P_1^B = R^W$ following the proof of proposition 1. Substituting for the corresponding equilibrium returns and quantities into equation (33), the MMF profit is zero, and all of the results and proofs of the paper hold directly and unchanged. ■

The MMF is not a required component within the model since it does not affect the results of the paper. However, a value-added role for the MMFs can easily be incorporated into the model to make the MMF a necessary feature for achieving the optimality results of the paper. The household liquidity benefit for money-like assets can be more narrowly defined as to apply only to household assets that have a contracted one-period return. This definition implies that the liquidity benefit would apply to deposits and shares but not to bonds, which makes the role of the MMF critical. The sale at date 1 of bonds bought at date 0 occurs in the bond market. Hence, the one-period return for date 0 bonds is based on the bond market transaction rather than a contracted one-period return as is provided by the MMF along with banks.

References

Acharya, Viral V., Denis Gromb and Tanju Yorulmazer (2012). "Imperfect Competition in the Interbank Market for Liquidity as a Rationale for Central Banking," *American Economic Journal: Macroeconomics* 4(2), pp. 184-217.

Acharya, Viral V. and David Skeie (2011). "A Model of Liquidity Hoarding and Term Premia in Inter-Bank Markets" *Journal of Monetary Economics* 58(5), pp. 436-447.

Adrian, Tobias, Erkko Etula and Tyler Muir (2014). "Financial Intermediaries and the Cross-Section of Asset Returns," *Journal of Finance* 69(6), pp. 2557–2596.

Adrian, Tobias and Hyun Song Shin (2008). "Financial Intermediaries, Financial Stability, and Monetary Policy," *2008 Economic Policy Symposium Proceedings.* Jackson Hole: Federal Reserve Bank of Kansas City.

Adrian, Tobias and Hyun Song Shin (2009). "Money, Liquidity, and Monetary Policy," *American Economic Review Papers and Proceedings* 99(2), pp. 600-605.

Adrian, Tobias and Hyun Song Shin (2014). "Procyclical Leverage and Value-at-Risk," *Review of Financial Studies* 27(2), pp. 373–403.

Afonso, Gara and Ricardo Lagos (2015). "Trade Dynamics in the Market for Federal Funds," *Econometrica* 83(1), pp. 263-313.

Armenter, Roc and Benjamin Lester (2017). "Excess Reserves and Monetary Policy Implementation," *Review of Economic Dynamics* 23, pp. 212–235.

Ashcraft, Adam B. and Darrell Duffie (2007). "Systemic Illiquidity in the Federal Funds Market," *American Economic Review* 97(2), pp. 221-225.

Ashcraft, Adam B., James McAndrews and David Skeie (2011). "Precautionary Reserves and the Interbank Market," *Journal of Money, Credit and Banking* 43(S2), pp. 311 - 348.

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

Bech, Morten L. and Elizabeth Klee (2011). "The Mechanics of a Graceful Exit: Interest on Reserves and Segmentation in the Federal Funds Market," *Journal of Monetary Economics* 58(5), pp. 415-431.

Bech, Morten L. and Cyril Monnet (2016). "A Search-Based Model of the Interbank Money Market and Monetary Policy Implementation," *Journal of Economic Theory* 164(C), pp. 32-67.

Bennett, Paul and Stavros Peristiani (2002). "Are U.S. Reserve Requirements Still Binding?" Federal Reserve Bank of New York *Economic Policy Review* 8(1).

Bernanke, Ben S. (2016). "Should the Fed Keep Its Balance Sheet Large?" September 2. Located at: https://www.brookings.edu/blog/ben-bernanke/2016/09/02/should-the-fed-keep-its-balance-sheet-large/.

Bindseil, Ulrich (2016). "Evaluating Monetary Policy Operational Frameworks," *2016 Economic Policy Symposium Proceedings*. Jackson Hole: Federal Reserve Bank of Kansas City.

Bryant, John (1980). "A Model of Reserves, Bank Runs, and Deposit Insurance," *Journal of Banking and Finance* 4(4), pp. 335-344.

Bryant, John and Neil Wallace (1979). "The Inefficiency of Interest-Bearing National Debt," *Journal of Political Economy* 87(2), pp. 365-381.

Bryant, John and Neil Wallace (1980). "Open Market Operations in a Model of Regulated, Insured Intermediaries," *Journal of Political Economy* 88(1), pp. 146-173.

Caballero, Ricardo J. and Arvind Krishnamurthy (2009). "Global Imbalances and Financial Fragility," *American Economic Review* 99(2), pp. 584-88.

Cochrane, John H. (2005). "Money as Stock," *Journal of Monetary Economics* 52(3), pp. 501–528.

Cochrane, John H. (2014). "Monetary Policy with Interest on Reserves," *Journal of Economic Dynamics & Control* 49(C), pp. 74–108.

Curdia, Vasco and Michael Woodford (2011). "The Central-Bank Balance Sheet as an Instrument of Monetary Policy," *Journal of Monetary Economics* 58(1), pp. 54-79.

Dang, Tri Vi, Gary Gorton and Bengt Holmstrom (2009). "Opacity and the Optimality of Debt for Liquidity Provision," working paper.

Dang, Tri Vi, Gary Gorton and Bengt Holmstrom (2015). "The Information Sensitivity of a Security," working paper.

Dang, Tri Vi, Gary Gorton, Bengt Holmstrom and Guillermo Ordonez (2017). "Banks as Secret Keepers," *American Economic Review* 107(4), pp. 1005-1029.

Diamond, Douglas W. and Philip Dybvig H. (1983). "Bank Runs, Deposit Insurance, and Liquidity," *Journal of Political Economy* 91(3), pp. 401-419.

Duffie, Darrell and Arvind Krishnamurthy (2016). "Pass-Through Efficiency in the Fed's New Monetary Policy Setting," *2016 Economic Policy Symposium Proceedings*. Jackson Hole: Federal Reserve Bank of Kansas City.

Ennis, Huberto M. and Todd Keister (2008). "Understanding Monetary Policy Implementation," Federal Reserve Bank of Richmond *Economic Quarterly* 94(3), pp. 235-263.

Freixas, Xavier, Antoine Martin and David Skeie (2009). "Bank Liquidity, Interbank Markets and Monetary Policy," Federal Reserve Bank of New York *Staff Reports* No. 371.

Freixas, Xavier, Antoine Martin and David Skeie (2011). "Bank Liquidity, Interbank Markets and Monetary Policy," *Review of Financial Studies* 24(8), pp. 2656-2692.

Freixas, Xavier and Jean-Charles Rochet (2008). *Microeconomics of Banking, Second Edition*. Cambridge: MIT Press.

Gagnon, Joseph E., Matthew Raskin, Julie Remanche and Brian Sack (2011). "Large-Scale Asset Purchases by the Federal Reserve: Did They Work?" Federal Reserve Bank of New York *Economic Policy Review* 17(1), pp. 41-59.

Gagnon, Joseph E. and Brian Sack (2014). "Monetary Policy with Abundant Liquidity: A New Operating Framework for the Fed," *Policy Briefs* PB14-4, Peterson Institute for International Economics.

Gertler, Mark and Peter Karadi (2013). "QE 1 vs. 2 vs. 3...: A Framework for Analyzing Large-Scale Asset Purchases as a Monetary Policy Tool," *International Journal of Central Banking* 9(1), pp. 5-53.

Goodfriend, Marvin. (2002). "Interest on Reserves and Monetary Policy," Federal Reserve Bank of New York *Economic Policy Review* 8(1), pp. 13-29.

Gorton, Gary and George Pennacchi (1990). "Financial Intermediaries and Liquidity Creation," *Journal of Finance* 45(1), pp. 49-71.

Greenwood, Robin, Samuel G. Hanson and Jeremy C. Stein (2016). "The Federal Reserve's Balance Sheet as a Financial-Stability Tool," *2016 Economic Policy Symposium Proceedings*. Jackson Hole: Federal Reserve Bank of Kansas City.

Greenwood, Robin, Samuel G Hanson and Jeremy C Stein (2015). "A Comparative-Advantage Approach to Government Debt Maturity," *Journal of Finance* 70(4), pp. 1683-1722.

He, Zhiguo, Bryan Kelly and Asaf Manela (2017). "Intermediary Asset Pricing: New Evidence from Many Asset Classes," *Journal of Financial Economics* 126(1), pp. 1-35.

He, Zhiguo, In Gu Khang and Arvind Krishnamurthy (2010). "Balance Sheet Adjustments during the 2008 Crisis," *IMF Economic Review* 58(1), pp. 118-156.

He, Zhiguo and Arvind Krishnamurthy (2012). "A Model of Capital and Crises," *Review of Economic Studies* 79(2), pp. 735-777.

He, Zhiguo and Arvind Krishnamurthy (2013). "Intermediary Asset Pricing," *American Economic Review* 103(2), pp. 732-770.

Kashyap, Anil K. and Jeremy C. Stein (2012). "The Optimal Conduct of Monetary Policy with Interest on Reserves," *American Economic Journal: Macroeconomics* 4(1), pp. 266-282.

Keister, Todd, Antoine Martin and James McAndrews (2008). "Divorcing Money from Monetary Policy," Federal Reserve Bank of New York *Economic Policy Review* 14(2), pp. 41-56.

Klee, Elizabeth, Zeynep Senyuz, and Emre Yoldas (2016). "Effects of Changing Monetary and Regulatory Policy on Overnight Money Markets," Finance and Economics Discussion Series 2016-084. Washington: Board of Governors of the Federal Reserve System.

Krishnamurthy, Arvind (2002). "The Bond/Old-Bond Spread," *Journal of Financial Economics* 66(2-3), pp. 463-506.

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*, pp. 3-43.

Krishnamurthy, Arvind and Annette Vissing-Jorgensen (2012). "The Aggregate Demand for Treasury Debt," *Journal of Political Economy* 120(2), pp. 233-267.

Martin, Antoine, James McAndrews and David Skeie (2016). "Bank Lending in Times of Large Reserves," *International Journal of Central Banking* 12(4), pp. 193-222.

Nagel, Stefan (2016). "The Liquidity Premium of Near-Money Assets," *Quarterly Journal of Economics* 131(4), pp. 1927–1971.

Poole, William (1968). "Commercial Bank Reserve Management in a Stochastic Model: Implications for Monetary Policy." *Journal of Finance* 23(5), pp. 769-791.

Potter, Simon (2014). "Interest Rate Control during Normalization," Remarks at the SIFMA Conference on Securities Financing Transactions, October 7.

Reis, Ricardo (2016). "Funding Quantitative Easing to Target Inflation," *2016 Economic Policy Symposium Proceedings*. Jackson Hole: Federal Reserve Bank of Kansas City.

Rochet, Jean-Charles (2008). *Why Are There So Many Banking Crises? The Politics and Policy of Bank Regulation*. Princeton: Princeton University Press.

Rochet, Jean-Charles and Jean Tirole (1996). "Interbank Lending and Systemic Risk," *Journal of Money, Credit and Banking* 28(4, Part 2), pp. 733-762.

Sims, Christopher. (2013). "Paper Money," *American Economic Review* 103(2) pp. 563–584.

Sims, Christopher (2016). "Fiscal Policy, Monetary Policy and Central Bank Independence," *2016 Economic Policy Symposium Proceedings.* Jackson Hole: Federal Reserve Bank of Kansas City.

Stein, Jeremy C. (2012). "Monetary Policy as Financial Stability Regulation," *Quarterly Journal of Economics* 127(1) pp. 57–95.

Williamson, Stephen D. (2016). "Interest on Reserves, Interbank Lending, and Monetary Policy," working paper.

Wrightson ICAP (2008). *Money Market Observer*, October 6.

Yellen, Janet (2016). "Opening Remarks: The Federal Reserve's Monetary Policy Toolkit: Past, Present and Future," *2016 Economic Policy Symposium Proceedings.* Jackson Hole: Federal Reserve Bank of Kansas City.