Excess Reserves and Monetary Policy Normalization[*]

Roc Armenter
Federal Reserve Bank of Philadelphia

Benjamin Lester[†]
Federal Reserve Bank of Philadelphia

May 6, 2015

Abstract

PRELIMINARY AND INCOMPLETE.

1 Introduction

Prior to the financial crisis, the Federal Reserve implemented monetary policy in a fairly simple way. Each day, the Open Markets Trading Desk (the "Desk") at the Federal Reserve Bank of New York would estimate the (downward-sloping) aggregate demand of depository institutions for reserves. Then, they would engage in open market operations, adjusting the quantity of reserves available in the federal funds market so that the supply curve would intercept the demand curve at (or near) the target federal funds rate.

During the Great Recession, however, the Federal Reserve resorted to a number of unconventional policies that have left the Fed with an unprecedented quantity of assets on its balance sheet, and depository institutions awash with excess reserves. This environment poses a challenge for monetary policy implementation, since the size of the open market operations that would be required to lift the federal funds rate off the zero lower bound is neither feasible nor desirable in the medium term.

Recognizing this situation, the Federal Open Markets Committee (FOMC) has outlined precisely how it plans to influence interest rates in the federal funds market until borrowing and lending activity among depository institutions returns to pre-crisis levels. As detailed in the September 17, 2014 press release, "Policy Normalization Principles and Plans," the FOMC intends to rely on two tools to implement the desired policy rate.

First, the committee "intends to move the federal funds rate into the target range [...] by adjusting the interest rate it pays on excess reserve balances," or what is commonly called the IOER rate. In addition to varying the rate it pays depository institutions for their excess reserves, the committee also "intends to use an overnight reverse repurchase agreement facility [...] to help control the federal funds rate," though the plan is to use this latter tool, commonly called the ON RRP rate, "only to the extent necessary."[1]

As policymakers prepare to begin the process of normalization, many important questions linger. Most importantly, how will short-term rates respond to changes in the IOER and ON RRP rates—which are currently set at 25 basis points and 5 basis points, respectively—and how will adjusting these two key policy rates affect take-up at the ON RRP facility? While the Federal Reserve Bank of New York has started experimenting with an ON RRP facility, some scenarios simply cannot be tested until liftoff itself. In this context, it would seem helpful to have a theoretical framework that can be used to understand the factors affecting current short-term interest rates and take-up (or volume) at the ON RRP, and to help identify those factors that will affect rates and take-up in response to changes in either policy or the economic environment more generally.

This paper provides such a framework. The key ingredients of our model are as follows. There is a central bank that operates two facilities: one pays interest on excess reserves to qualified depository institutions (DIs), and another provides a positive rate of return for overnight reverse repurchase agreements. The latter (ON RRP) rate is lower than the former (IOER) rate, but is available to financial institutions with excess cash (who we call *lenders*) that do not qualify as

---

[1] All quotations are from "Policy Normalization Principles and Plans," Federal Reserve System Press Release, September 17, 2014. The FOMC also stated that it will "phase it [the ON RRP facility] out when it is no longer needed" and, in the longer run, intends to " hold no more securities than necessary to implement monetary policy efficiently and effectively... thereby minimizing the effect of Federal Reserve holdings on the allocation of credit across sectors of the economy."

DIs. Hence, there is an arbitrage opportunity: DIs should be willing to borrow cash at a rate below the IOER rate and pocket the difference. However, there are two potential frictions in this inter-bank market.

First, we assume that the market is not perfectly competitive, but rather characterized by search frictions in order to capture the "over-the-counter" nature of the fed funds market. In particular, we posit a *directed search* model. DIs post interest rates at which they wish to borrow. Lenders then decide which rate to pursue. Whether a lender successfully matches with a DI depends on the ratio of lenders-to-DIs seeking/offering a particular rate. Indeed, DIs face a trade-off between posting a high interest rate and attracting many lenders—thus matching with high probability but earning less revenue per match—or posting a low interest rate and matching with lower probability. Thus our framework endows DIs with some degree of *market power*, which depends on the (endogenously determined) equilibrium market tightness. In a successful match, the DI deposits the lender's cash at the Fed to earn the IOER, retaining some of the excess return for its own profit. Those lenders who do not match in the interbank market can attempt to access the ON RRP facility.

The second key friction in our model is that DIs incur *balance sheet costs* when they accept deposits from lenders; these costs capture both the direct costs of a DI expanding its balance sheet, like FDIC fees, as well as the indirect costs associated with requirements on capital and leverage ratios.[2] More specifically, each DI faces a potentially different balance sheet cost. In equilibrium, some DIs find that they cannot attract any lender at a rate that justifies their high balance-sheet costs, and thus remain out of the federal funds market. DIs with low balance sheet costs, on the other hand, can offer high rates and attract many lenders.

We provide a complete characterization of the equilibrium, and derive a rich set of testable predictions regarding the distribution of offered and realized trades in the interbank market, along with take-up at the ON RRP facility. We show how all of these important variables respond to various changes in the IOER and ON RRP rates, or to changes in the economic environment (such as a shock to the supply of funds). We also incorporate a cap on take-up at the ON RRP facility, and discuss both when the cap should be expected to bind and how a binding cap will affect equilibrium outcomes. We highlight that it is possible that the average traded rate drops below the ON RRP rate once the ON RRP facility is capped.

Finally we put our model to task to characterize conditions such that liftoff—the first rate increase from virtually zero rates—will be successfully implemented.

1.1 Related Literature

There is a long tradition of developing models of the federal funds market to study the implementation of U.S. monetary policy. The original contribution is Poole (1968), who posited a downward-sloping demand curve for reserves, and analyzed how the Federal Reserve can target the desired federal funds rate by manipulating the supply of reserves. While this approach abstracted from the actual trading mechanisms in place, it proved very useful and became the workhorse model for the federal funds market, being further developed by Ho and Saunders (1985) and Hamilton (1996), among many others. Indeed, the tractability of this approach has made it an attractive framework for embedding the federal funds market into a larger macroe-

---

[2]See, e.g., Potter (2014).

conomic model, as in Martin et al. (2013a) and Martin et al. (2013b).

Recently, a second generation of models have been developed in order to capture the actual micro-mechanics of trading in the federal funds market—in particular, its *over-the-counter nature*, which was first emphasized by Furfine (1999) and Ashcraft and Duffie (2007). One prominent example is Afonso and Lagos (2015), who develop a random search model to capture the idea that trade in this market is bilateral in nature and trading partners often take time to locate. They use their model to explore intra-day trading dynamics in the federal funds markets and the determinants of the federal funds rate.[3] Ennis and Weinberg (2013) also develop a search and matching model of the federal funds market to study the "stigma" associated with the use of the discount window and its implications for the demand of reserves. Bianchi and Bigio (2014) embed a simplified model of over-the-counter trading in a macro model designed to study the bank lending channel of monetary policy. Explicit models of over-the-counter trading have also been used to study monetary policy implementation outside the U.S.: see Berentsen and Monnet (2008) and Bech and Monnet (2014), among many others.

All of the models discussed above are based on the premise that most trades in the federal funds market occur *between* depository institutions adjusting their desired reserve holdings, which was a fairly accurate description of this market up until the crisis in 2007-2008. In such an environment, depository institutions were willing to lend to each other at a rate above the IOER rate and borrow from each other at a rate below the penalty rate associated with the discount window. Thus, these models imply that the federal funds rate would typically remain within a "corridor" defined by the IOER rate (as a floor) and the discount-window rate (as a ceiling). Currently, however, most depository institutions are holding large quantities of *excess* reserves—more than enough to satisfy their reserve requirements. Hence trade *between* depository institutions has become rare, and instead the supply of funds is now provided mostly by non-depository institutions seeking an overnight investment vehicle.[4] Moreover, these lenders are inherently different from depository institutions since they do not have access to the IOER rate. In this type of environment, the IOER rate no longer serves as an effective floor as Bech and Klee (2011) point out; indeed, the fed funds rate has traded below the IOER rate since 2008.

Our model is designed to capture the determinants of interest rates in this new economic environment, i.e., to study the implementation of monetary policy when depository institutions are holding excess reserves. To do so, we abstract from trades between depository institutions, and instead focus on federal funds trades between two distinct sides of the market: non-depository institutions on one side, seeking to obtain a positive yield on their overnight investments; and depository institutions on the other side, looking to arbitrage the interest rate differential between the IOER rate and the outside option of non-depository institutions.

We explicitly include and evaluate the key operational tools that have been recently introduced by the Federal Reserve—namely, the IOER rate, which is intended to serve as a ceiling on the federal funds rate, and the rates offered at the ON RRP facility, which are designed to serve as a floor in the current environment. In attempt to make our framework amenable to quantitative analysis, we also incorporate heterogeneous balance-sheet costs, which generate dispersion in traded rates and participation across depository institutions that can be mapped

---

[3]For an abridged version of their model, see Afonso and Lagos (forthcoming).

[4]See Afonso et al. (2013a).

to the data.[5]

Like many of the second-generation models discussed above, we utilize a model with search frictions to capture the over-the-counter nature of the federal funds market. However, unlike these models, we utilize a model of *directed* (instead of random) search.[6] This choice is motivated by several factors. First, this modeling device captures the idea that the market is currently segmented into two distinct sides, as we discussed above. Moreover, given the repeated daily interactions between these two sides, it seems less likely that market participants have no ex ante information about which depository institutions typically offer higher or lower interest rates on overnight loans, and thus contact depository institutions at random. A virtue of our directed search model is that the interest rate that a depository institution is willing to pay is known to all market participants, and those that offer relatively high interest rates will, in equilibrium, attract more overnight loans. Put differently, our model captures the bilateral and stochastic nature of meetings in the federal funds market without severing the tie between interest rates and allocations. Finally, models of directed search tend to have several attractive technical features as well. For one, they offer a framework in which sellers compete with one another in a setting that lies between the extreme cases of monopoly and Bertrand competition. Moreover, directed search models are highly tractable, they can easily incorporate various types of (observed or unobserved) heterogeneity, and the equilibrium in these models tend to be constrained efficient.[7]

2   Background on Monetary Policy Implementation

The Federal Open Market Committee (FOMC) pursues their mandated objectives of price stability and full employment by setting a target level or range for the overnight federal funds rate. A federal funds transaction is an unsecured loan of U.S. dollars between eligible entities, like depository institutions or government-sponsored enterprises.[8] There is no central repository for federal funds trades; instead, participants arrange transactions directly with each other or through brokers—what is commonly described as an "over the counter" market.[9]

The Trading Desk at the Federal Reserve Bank of New York (commonly known as simply the Desk) is tasked with implementing the FOMC's directive. Traditionally, the federal funds market was dominated by trades *between* depository institutions seeking to adjust their reserve holdings to desired levels: institutions with excess reserves would lend funds to institutions with reserves short of the required level. In this context, the Desk implemented the instructed rates simply

---

[5]See Martin et al. (2013a) and Potter (2014) for a discussion of the regulation and liquidity management concerns behind balance-sheet costs in overnight funds.

[6]For seminal contributions to the directed (or competitive) search literature, see Moen (1997), Burdett et al. (2001), and the references therein.

[7]In contrast, random search models typically require that any heterogeneity is observable, in order to avoid issues with bargaining under private information. Also, random search models are generically inefficient, as congestion externalities are not internalized when prices are determined by Nash bargaining except in a knife-edge case; see Hosios (1990).

[8]The Federal Reserve Regulation D establishes which institutions are eligible for federal funds, which are exempt from reserve requirements—a privilege shared by Eurodollar deposits since 1991.

[9]Ashcraft and Duffie (2007) emphasize the over-the-counter nature of the federal funds market in their empirical analysis; more recently, researchers have followed up by explicitly modeling the micro mechanics of federal funds transactions, as in Bech and Klee (2011) and Afonso and Lagos (2015), among others.

by fine-tuning the supply of reserves in the banking system via open market operations.[10] As described in Potter (2013), the Desk was able to reliably achieve the FOMC's policy directives, often without much need to conduct operations.

The current landscape in money markets is quite different. In response to the financial crisis and the following recession, the FOMC reduced its fed funds target to virtually zero in December 2008 and, crucially for the matter at hand, embarked on a series of liquidity and asset-purchase programs that resulted in an extremely high level of excess reserves. As all but very few depository institutions had any need for additional reserve, trade between banks dried up and the federal funds rate dropped substantially below the FOMC's target.[11] In an attempt to put a floor on short-term interest rates, the Federal Reserve started to pay interest on excess reserves (IOER) to depository institutions.[12]

The federal funds rate and other money market rates, though, have consistently traded below the IOER. The reason is that the IOER is only available to depository institutions holding balances at the Fed—and thus excluding key participants like, for example, government-sponsored enterprises (GSE) and money market funds. Indeed, the federal funds rate is now driven by GSEs, like the Federal Home Loans, seeking to place their funds with a depository institution, which then earn the IOER.[13] As argued by Potter (2014), depository institutions have been able to earn a spread between the cost of funds and the IOER for a variety of reasons like balance-sheet costs and market power.

The final piece of the institutional landscape is the overnight reverse repurchase (ON RRP) facility that the Desk has been testing since September 2013 with an expanded list of counterparties.[14] In a reverse repurchase, the Desk sells a security to a counter party with an agreement to buy the security back at a pre-specified date and price, with the interest rate of the repurchase computed from the difference between the purchase and the higher repurchase price. During the tests, the Desk announced an ON RRP rate as well as a maximum allotment.[15] As such, the ON RRP facility offers an investment alternative to institutions ineligible for the IOER and thus helps support the level of the federal funds rate.

It is important to emphasize that the list of counterparties for the ON RRP facility is designed to include nonbank institutions that are significant lenders in money markets. This includes GSE's but also key money market funds, which actually are not eligible to transact

---

[10]Permanent additions to the supply of reserves were slightly below the total demand, creating a "structural deficiency" in the system that allowed the Desk to fine-tune market conditions with additional temporary open-market operations. See Akhtar (1997) for a detailed description of open market operations.

[11]By September 2008, the Desk lost its capacity to offset the increase in the Fed's balance sheet and thus effectively lost control of the supply of reserves.

[12]The so-called "floor system" has been successfully implemented by a number of central banks, including the European Central Bank, the Bank of Japan, the Riksbank, the Bank of Canada, and the Bank of England among many others. It has also been extensively studied in the literature, see Ennis and Keister (2008), Whitesell (2006), and Berentsen and Monnet (2008) among others. The IOER remains an effective floor on interest rates in models capturing the over-the-counter nature of the federal funds market as Afonso and Lagos (2015) or Bianchi and Bigio (2014). An exemption is Bech and Klee (2011).

[13]See Afonso et al. (2013b) and Afonso et al. (2013a) for a description of the key participants in the federal funds market after the financial crisis.

[14]See Frost et al. (2015) for a complete discussion of the design of the ON RRP facility.

[15]If the demand for ON RRPs exceeds the maximum allotment, the FOMC directed the Desk to use an auction process to set the interest rate on ON RRPs, which can then be lower than the offering rate announced by the Desk.

federal funds and instead lend to depository institutions through the Eurodollar markets.[16] The rationale of the expanded counterparty list is to prevent the federal funds market to become disconnected from the rest of money-market rates: as noted by Potter (2013), the effectiveness of the ON RRP facility depends on including a sufficiently wide set of non-bank counterparties.[17] The Federal Reserve Bank of New York has indeed announced that it plans to publish an estimate of the overnight bank funding rate that would include transactions in Eurodollar markets by depository institutions.[18]

Going forward, the FOMC has already stated that 'intends to reduce the Federal Reserve's securities holdings in a gradual and predictable manner primarily by ceasing to reinvest repayments of principal on securities held" in the Fed's balance sheet, and expects to "cease or commence phasing out reinvestments after it begins increasing the target range for the federal funds rate."[19] Thus the FOMC expects to implement the process of interest rate "normalization" in the current context of excess reserves, relying on administered rates and other market actions to influence market rates. More precisely, the FOMC has stated that it intends to implement the desired target range for the federal funds rate primarily by adjusting the IOER rate, using the ON RRP facility and other tools as needed in a supplementary role.

3   Environment

We consider a two-period economy that is populated by three types of agents. First, there is a measure λ of non-depository financial institutions which we will refer to as "lenders," as this is representative of their primary role in the model. Second, there is a measure of depository institutions, normalized to one, which we will refer to as "DIs" for brevity. Finally, there is a central bank which, for obvious reasons, we will call the "Fed." All agents are risk neutral and do not discount between $t = 1$ and $t = 2$.

The Fed. The Fed is an institution that operates two facilities. The first is an "Overnight Reverse Repurchase" (ON RRP) facility, where an agent can use cash to purchase a security from the Fed at $t = 1$, and then resell the security to the Fed at $t = 2$ at a pre-specified price. We denote the net rate of return on this investment by r, and assume this is chosen by the Fed. Both lenders and DIs have access to the ON RRP facility.

The Fed also accepts cash deposits at $t = 1$, which earn a net interest rate of $R > r$ at $t = 2$. We refer to R as the Interest on Excess Reserves (IOER) rate, and assume that this rate is also chosen by the Fed. Importantly, only DIs have access to this second facility.

---

[16]The complete list of counterparties is available at `http://www.ny.frb.org/markets/expanded_counterparties.html`.

[17]Federal funds and Eurodollar deposits are nearly perfect substitutes for depository institutions since Eurodollar were also exempted from reserve requirements in 1991, as documented in Bartolini et al. (2008) and others. The other, large overnight money market rate is the "tri-party" repo market—however, depository institutions are not commonly very active in such market. See Demiralp et al. (2006) on the inter-linkages across money markets.

[18]"Statement Regarding Planned Changes to the Calculation of the Federal Funds Effective Rate and the Publication of an Overnight Bank Funding Rate," Federal Reserve Bank of New York, February 2, 2015.

[19]"Policy Normalization Principles and Plans," Federal Reserve System Press Release, September 17, 2014.

Lenders.   At $t = 1$, each lender is endowed with 1 unit of excess cash. As noted above, lenders cannot deposit their reserves directly at the Fed to earn the IOER rate, R. However, there exists an interbank market where they can lend their cash to a DI, who can then deposit it at the Fed to earn R, while retaining some of the excess return for its own profit. The interbank market is a frictional one, though: not all lenders will match with DIs, and not all DIs will match with a lender. We discuss the probability that a lender matches with a DI, and the interest rate that they earn if they do match successfully, in greater detail below.

Those lenders who do not match in the interbank market can attempt to access the ON RRP facility and earn the interest rate r. To start, we will assume that there are no limits on the volume of trade at the ON RRP facility, so that lenders can always access this facility if they fail to match in the interbank market. In this case, lenders are guaranteed a rate of return of at least r on their excess cash holdings. Later, we will consider what happens if policymakers impose a cap on the volume of trade at the ON RRP facility, so that lenders face some risk of having to store their cash holdings themselves, in which case we assume that they earn no interest.

Depository Institutions.   Each DI, which we index by j, can accept up to one unit of cash from a lender. However, doing so imposes a "balance sheet cost" on the DI, which we denote by c_j. DIs are heterogeneous with respect to their balance sheet costs: we denote by $G(c_j)$ the distribution of costs across DIs, and assume that $G(\cdot)$ is continuous and strictly increasing on the support $c_j \in [0, \infty)$. Given its balance sheet cost, each DI decides whether or not to enter the interbank market and, if they enter, they choose an interest rate that they will pay to borrow a unit of cash, which we denote by ρ_j.

Figure 1 summarizes the possible transactions that can occur between lenders, DIs, and the Fed, along with the interest rates that are paid in each type of transaction.

Matching in the Interbank Market.   Once DIs have made entry decisions and posted interest rates, each lender observes the interest rates that have been posted and chooses one to approach. We will often refer to the set of DIs that have chosen a particular interest rate as a "sub-market."[20] Conditional on choosing a sub-market, a lender may or may not be paired with a DI—there are *matching frictions* in the interbank market.[21] In particular, suppose a measure d of DIs post a particular interest rate and a measure ℓ of lenders choose to purse that rate. Then, letting $q = \ell/d$ denote the market tightness or "queue length" in that sub-market, the probability that each DI receives a deposit is $1 - e^{-q}$ and, symmetrically, the probability that each lender matches with a DI is $\frac{1-e^{-q}}{q}$.

---

[20]This nomenclature is borrowed from the competitive search literature, where each posted price (wage) is associated with a sub-market. In these models, sellers (firms) first choose a sub-market to sell their good (post their vacancy), and then buyers (workers) choose a sub-market in attempt to buy the good (get a job).

[21]As discussed in Section 2, these institutions interact through federal funds and eurodollar transactions. From the point of view of the model, the distinction is irrelevant: Lenders approach the DIs directly and their trade gets formalized as a federal funds or an eurodollar transaction without impacting the terms of trade. We return to this point in Section 6, when it becomes necessary to map the available data to the model in order to pursue a tentative calibration of the model.
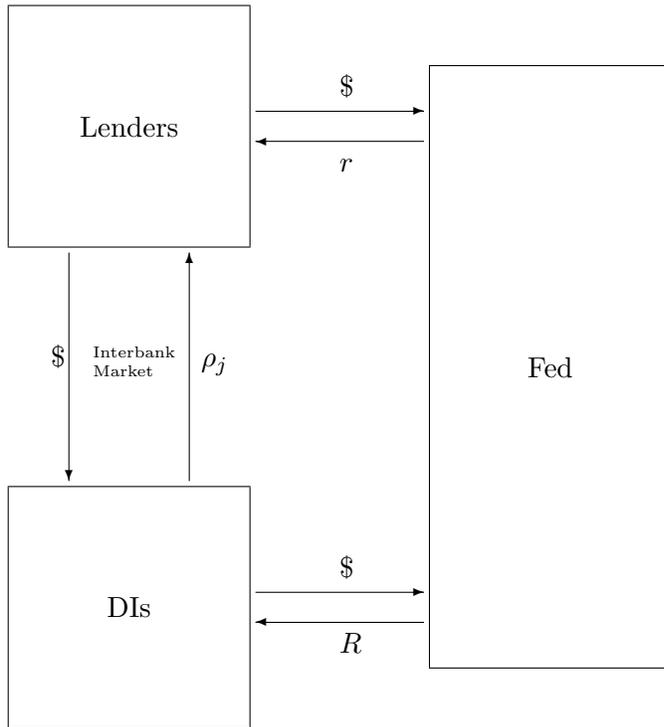
Figure 1: Summary of Possible Transactions

Summary of Timeline and Payoffs. The timeline in Figure 2 depicts the sequence of events. A DI who is matched with a lender at $t = 1$ incurs the balance sheet cost c_j and deposits the reserves at the Fed. Then, at $t = 2$, the DI earns the IOER rate R and pays the lender the promised interest rate ρ_j. Hence, a DI with balance sheet cost c_j that posts interest rate ρ_j and matches with a lender will earn a net profit of $R - c_j - \rho_j$. An unmatched DI, on the other hand, earns zero. Meanwhile, a lender earns a payoff of ρ_j if he successfully matches with a DI who has posted an interest rate ρ_j, and otherwise earns r from an ON RRP transaction with the Fed.

4 Equilibrium

In this section, we fully characterize the equilibrium in the benchmark model and explore its properties. We start by deriving the optimal behavior of DIs and lenders, which allows us to formally define an equilibrium. We then establish existence and uniqueness, and highlight several important properties of all equilibria. Finally, we conduct comparative statics, exploring how posted and traded rates, trading volume, and takeup at the ON RRP facility respond to policy changes and other exogenous shocks.
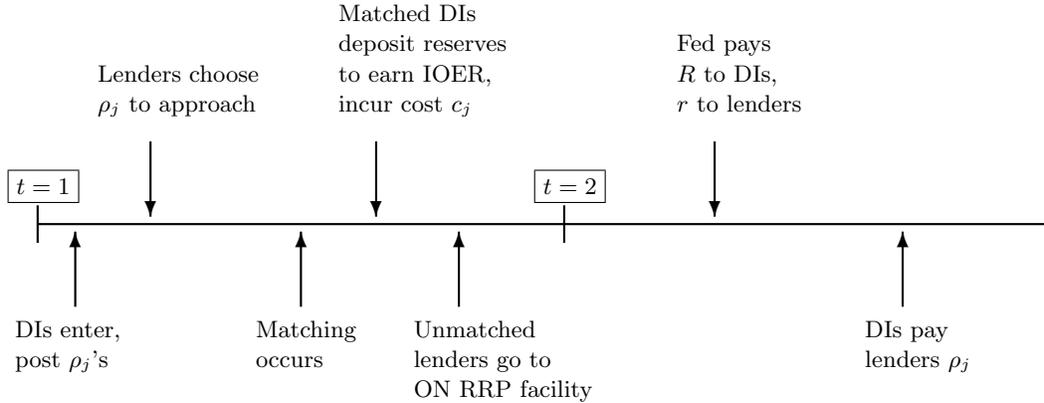
Figure 2: Timeline

4.1 Optimal Strategies and Definition of Equilibrium

Decision-making occurs at three different stages. First, DIs have to decide whether or not to enter the interbank market given their balance sheet cost, c_j. Second, those DIs that enter have to choose an interest rate, ρ_j. Finally, given the interest rates that have been posted, lenders have to choose which one to approach. In order to describe optimal behavior at each stage, we work backwards.

Optimal Search by Lenders. Once interest rates have been posted, each lender must choose the sub-market (or mix between sub-markets) that offers the maximum expected payoff, taking into account both the interest rate being offered in that sub-market *and* the probability of being matched. In particular, the expected payoff from a lender choosing a sub-market with interest rate ρ_j and queue length q_j is

$$u(\rho_j, q_j) = \left[\frac{1 - e^{-q(c_j)}}{q(c_j)} \right] \rho_j + \left[1 - \frac{1 - e^{-q(c_j)}}{q(c_j)} \right] r. \tag{1}$$

The first term captures the probability that the lender is matched with a DI, in which case he earns ρ_j, while the second term captures the probability that he fails to match, in which case the lender will approach the ON RRP facility and earn the rate r.

Let U denote the maximum expected payoff that a lender can obtain, or what we will call the "market utility." In equilibrium, then, any sub-market with $q_j > 0$ must satisfy

$$u(\rho_j, q_j) = U. \tag{2}$$

That is, in equilibrium, any DI that is able to attract lenders must deliver an expected payoff equal to the market utility: for example, if a DI posts a relatively low interest rate, lenders must be compensated with a high probability of being matched (i.e., a short queue length), and vice

10

versa. Using (1) to solve (2) yields

$$\rho_j = r + \left[\frac{q_j}{1 - e^{-q_j}}\right](U - r). \tag{3}$$

Optimal Interest Rate Posting by DIs. The lenders' indifference condition in equation (3) lays bare the trade-off facing DIs: they can post a low interest rate and match with low probability, or they can attract a longer queue by posting a higher interest rate, in which case they will match with higher probability. Taking this trade-off as given, a DI with balance sheet cost c_j who has entered the interbank market solves the following profit maximization problem:

$$\max_{\rho_j, q_j} \quad \left[1 - e^{-q_j}\right](R - c_j - \rho_j) \tag{4}$$

$$\text{sub to} \quad \rho_j = r + \left[\frac{q_j}{1 - e^{-q_j}}\right](U - r),$$

where the market utility U is taken parametrically by each DI. From the objective function, (4), it's clear that a DIs profits are equal to the product of the probability of being matched, $1 - e^{-q_j}$, and the revenue from accepting a deposit, $R - c_j - \rho_j$, taking as given the positive relationship between interest rates and queue lengths.

One can substitute the constraint into (4), which yields an objective function that is strictly concave over a single choice variable, q_j. Hence the first order condition delivers the optimal queue length for any c_j and any market utility U:

$$q_j \equiv q(c_j; U) = \log\left(\frac{R - c_j - r}{U - r}\right). \tag{5}$$

From (3), then, the optimal interest rate for a DI with balance sheet cost c_j, given U, is

$$\rho_j \equiv \rho(c_j; U) = r + \log\left(\frac{R - c_j - r}{U - r}\right)\left[\frac{(R - c_j - r)(U - r)}{R - c_j - U}\right]. \tag{6}$$

Optimal Entry by DIs. A DI enters the interbank market if, and only if, its expected profits from doing so are nonnegative.[22] One can easily show that a DI's profits are decreasing in c_j for any U, so that the optimal entry decision is determined by a cutoff rule: for any $U > r$, there exists a unique $\bar{c} > 0$ such that profits are nonnegative if, and only if, $c_j \leq \bar{c}$. Substituting (5) and (6) into (4) and solving reveals that this cutoff satisfies

$$\bar{c}(U) = R - U. \tag{7}$$

---

[22]In other words, we are assuming that a DI will stay out of the interbank market if it would not attract any lenders by entering. One could motivate this assumption by assuming that there was a cost $\epsilon > 0$ associated with posting an interest rate, where ϵ was arbitrarily close to zero.

Market Clearing. The analysis above describes the optimal decisions by lenders and borrowers, taking as given the market utility U. The final condition requires that markets clear:

$$\int_0^{\bar{c}(U)} q(c_j; U) dG(c_j) = \lambda. \tag{8}$$

In words, (8) requires that aggregating the queue lengths (or expected number of lenders per DI) across the active DIs yields the total measure of lenders in the market, λ.

Definition of Equilibrium. Given the results above, an equilibrium is a market utility $U^\star$, a cutoff $\bar{c} = \bar{c}(U^\star)$, queue lengths $q(c_j) = q(c_j; U^\star) > 0$ and interest rates $\rho(c_j) = \rho(c_j; U^\star) \in (r, R)$ for all $c_j < \bar{c}$ such that

1. Lenders are indifferent between all active DIs, i.e., (2) is satisfied.

2. Given lender's search behavior, those DIs that enter the market choose interest rates to maximize profits, i.e., (6) is satisfied.

3. DIs enter the market if, and only if, it is profitable to do so, i.e., (7) is satisfied.

4. Markets clear, i.e., (8) is satisfied.

4.2   Characterization of Equilibrium.

Let $s = R - r$ denote the spread between the IOER and ON RRP rates. Then, using (7), we can rewrite the market clearing condition (8) as

$$\int_0^{\bar{c}} \log\left(\frac{s - c_j}{s - \bar{c}}\right) dG(c_j) = \lambda. \tag{9}$$

Equation (9) reveals that characterizing the equilibrium boils down to solving one equation in one unknown, $\bar{c}$. Then, since $\bar{c}$ only depends on s, and not on the specific values of R and r, so too do the queue lengths. In particular, the equilibrium queue length at a DI with balance sheet cost $c_j \le \bar{c}$ is

$$q(c_j) = \log\left(\frac{s - c_j}{s - \bar{c}}\right). \tag{10}$$

Therefore, *allocations* only depend on policy through the spread between R and r. Intuitively, the spread s determines the total gains from trade between a lender and a DI. Absent any other changes, the share of these gains from trade that a DI can appropriate — i.e., the DI's profits — are constant.

Finally, given the cutoff and queue lengths, the remaining equilibrium objects follow immediately:

$$\rho(c_j) = r + \frac{q(c_j)}{1 - e^{-q(c_j)}}(s - \bar{c}) \text{ for all } c_j \le \bar{c}, \text{ and} \tag{11}$$

$$U = r + s - \bar{c}. \tag{12}$$

In the Appendix, we establish that there exists a unique solution to equation (9). In particular, if $s > 0$, then $\bar{c} > 0$ and there is a strictly positive measure of active DIs. If $s = 0$, of course, there are no gains from trade and the market shuts down, i.e., $\bar{c} = 0$. The following proposition summarizes.

Proposition 1. *There exists an equilibrium and it is unique. The cutoff $\bar{c} > 0$ if, and only if, $s > 0$.*

4.3   Properties of Equilibria

Before proceeding, we briefly summarize the relationship between a DI's balance sheet costs, the interest rates they post, the queue lengths they attract, and the corresponding profits they earn. Given an equilibrium cutoff $\bar{c} < s$, notice immediately that $q'(c_j) = -1/(s - c_j) < 0$ for all $c_j \leq \bar{c}$, so that

$$\rho'(c_j) = [1 - e^{-q(c_j)}(1 + q(c_j))]q'(c_j) < 0.$$

Hence, DIs with lower balance sheet costs post higher rates and attract longer queue lengths in equilibrium. As a result, a DI's profits $\Pi(c_j) = [1 - e^{-q(c_j)}][R - c_j - \rho(c_j)]$ are also strictly decreasing in c_j, with $\Pi(\bar{c}) = 0$.

Moreover, one can easily show that $r < \rho(\bar{c})$ and $\rho(0) < R$. The first inequality illustrates that the DI with the highest balance sheet costs still offers an interest rate greater than the ON RRP rate; in order to generate any demand, the DI with $c_j = \bar{c}$ must offer lenders an interest rate and a probability of matching that are at least as attractive as the terms being offered by banks with lower balance sheet costs. The second inequality illustrates that even banks with zero balance sheet costs offer an interest rate strictly less than the IOER rate. Intuitively, search frictions imply that queue lengths are not perfectly elastic, which allows DIs to extract some rents from lenders; that is, *DIs have some market power.*

4.4   A Numerical Example.

To further highlight some basic properties of the equilibrium, we consider a simple numerical example. We first use this example to illustrate some basic properties of the distributions of posted and traded rates, as well as takeup at the ON RRP. We then perturb some of the key parameters and show how the model can be useful for understanding how equilibrium outcomes might respond to changes in policy (e.g., the IOER and ON RRP rates), regulatory requirements (e.g., balance sheet costs), or market conditions (e.g., the supply of excess cash).

Benchmark.   In our benchmark example, we set the IOER rate at 25 basis points, $R = 0.25$, and the ON RRP rate at 5 basis points, $r = 0.05$, which are the rates that have been offered since 2014. We assume that the distribution of balance sheet costs is given by a Gamma distribution, with an average balance sheet cost of 15 basis points, $\mu_c = 0.15$, and a standard deviation of 5 basis points, $\sigma_c = 0.05$, which implies that roughly 95 percent of DIs have balance sheet costs below 25 basis points. Finally, we set the measure of lenders, λ, to 0.5. Our tentative choice of parameters generates an average traded rate of about 10 basis points, with trades ranging from approximately 8 to 12 basis points, and a bit more than one third of lenders ultimately relying on the ON RRP facility.
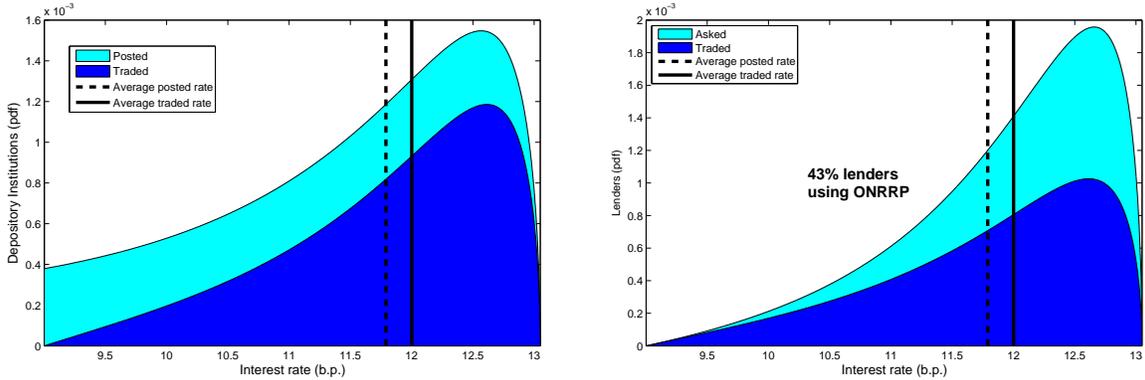
Figure 3: Posted, Chosen, and Traded Interest Rates

Distributions of Interest Rates   Given these parameter values, Figure 3 plots two distributions. The left panel depicts the distribution of interest rates that are posted by DIs in equilibrium and the distribution of interest rates that DIs actually pay. DIs that post higher interest rates are more likely to receive deposits, given that they attract larger queue lengths. Hence, the average interest rate paid by DIs exceeds the average posted interest rate. Also note that the distribution of interest rates inherits many properties from $G(c_j)$, the distribution of balance sheet costs. The right panel of Figure 3 plots the distribution of lenders across posted interest rates, and the distribution of interest rates actually received by lenders. Note that the mass of lenders that do not match with a DI deposit their funds at the ON RRP and earn r.

Policy Experiments.   In the analysis above, we established that the cutoff, $\bar{c}$, and the queue lengths at each type of DI, $q(c_j)$, are completely determined by s, λ, and $G(c_j)$. Hence, so long as the spread remains constant, changes in R and r alone have no effect on the aggregate market tightness, the number of lenders that match, and hence on the volume of lenders that use the ON RRP facility. Moreover, from (11), we know that increasing both the ON RRP and IOER rates by the same amount simply shifts the entire distribution of interest rates up one-for-one.

However, a policy change that effects the spread $s = R - r$ will have implications for DIs' entry decisions, queue lengths, and the volume of trade. As a result, this type of change in policy rates will also have nonlinear effects on the distribution of posted and traded rates. To see this, in the left panel of Figure 4, we plot the equilibrium distribution of traded interest rates in our benchmark model after increasing the IOER rate from 25 to 30 basis points, holding the ON RRP rate (and all other parameters) constant. This change causes an increase in $\bar{c}$, so that more DIs enter the market, market tightness falls, and the market becomes more competitive. As a result, interest rates shift up and the average traded interest rate rises. Moreover, since there are more DIs per lender, the fraction of lenders that ultimately end up depositing their funds at the ON RRP facility falls.

In the right panel of Figure 4, we plot the equilibrium distribution of traded interest rates after an increase in the ON RRP rate of 5 basis points, holding the IOER rate constant. Giving lenders a more attractive outside option reduces the rents available to DIs, and hence fewer DIs

14

enter. Ultimately, those DIs that do enter will offer higher interest rates, but fewer lenders will match, causing the volume at the ON RRP facility to rise.
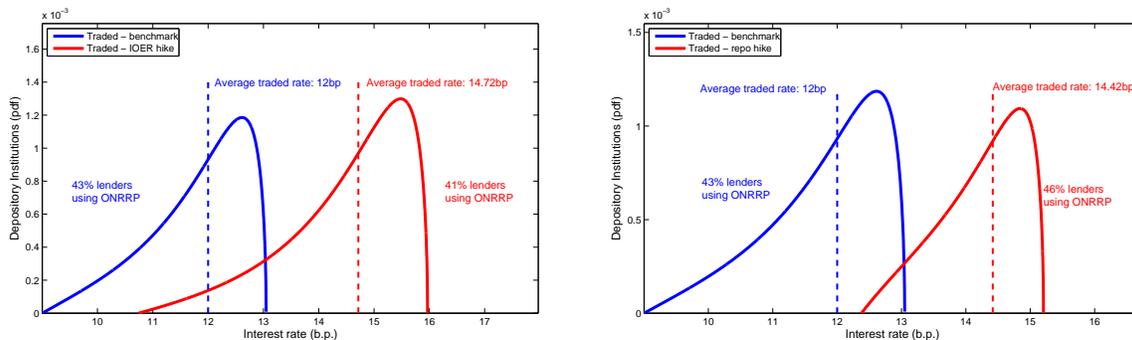

Figure 4: Increasing the IOER or ON RRP Rate

In a similar vein, Figure 5 plots the average traded interest rate and the share of lenders using the ON RRP for various values of r, holding R fixed. Several important points emerge. First, changing the spread between the IOER and ON RRP rates can have highly nonlinear effects. Second, as r increases, the rate at which DIs exit the market is informative about the shape of the distribution of balance sheet costs. Hence, variations in the ON RRP rate could contain valuable information about unobservable costs. Finally, as the ON RRP rate gets closer to the IOER rate, the average traded rate converges, but trade volume in the market also drops at an increasing rate as more and more deposits are done through the ON RRP facility.
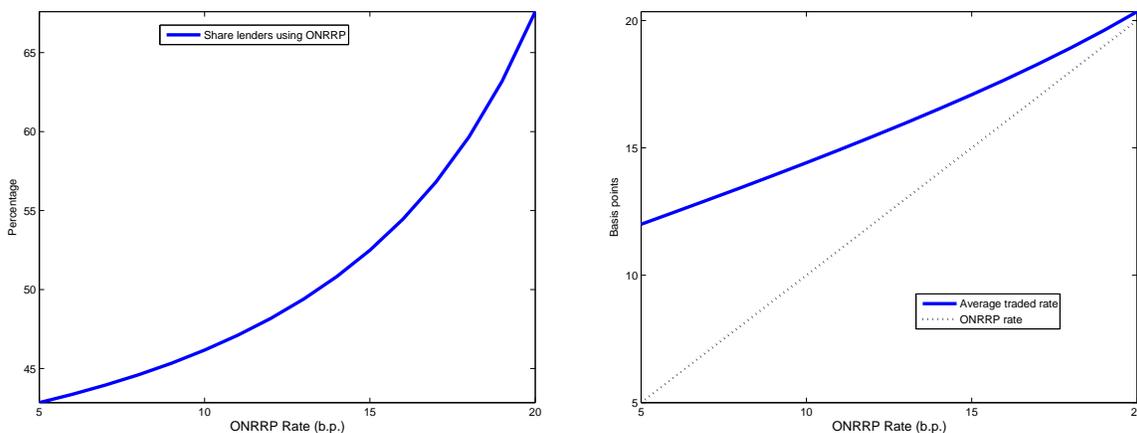

Figure 5: The ON RRP Rate, Take Up at the ON RRP Facility, and the Average Market Rate

Changes in Market Conditions. It is also important to understand how outcomes will respond to shocks to the economic environment. In Figure 6, for example, we plot the effect of

a 5 percent increase in the supply of lenders. When supply increases, $\bar{c}$ increases and more DIs enter. However, this increase in DIs is not enough to offset the increase in lenders, and overall market tightness increases. As a result, DIs set lower interest rates in equilibrium. In fact, the average traded interest rate falls by more than 5 percent: those DIs that participated under the benchmark parameterization face less elastic demand schedules, and hence lower the interest rates they offer; and the new entrants are those with relatively high balance sheet costs, which implies that they are offering relatively low interest rates. Also note that a higher fraction of lenders ends up utilizing the ON RRP facility.

Figure 6: A Shock to the Supply of Excess Cash
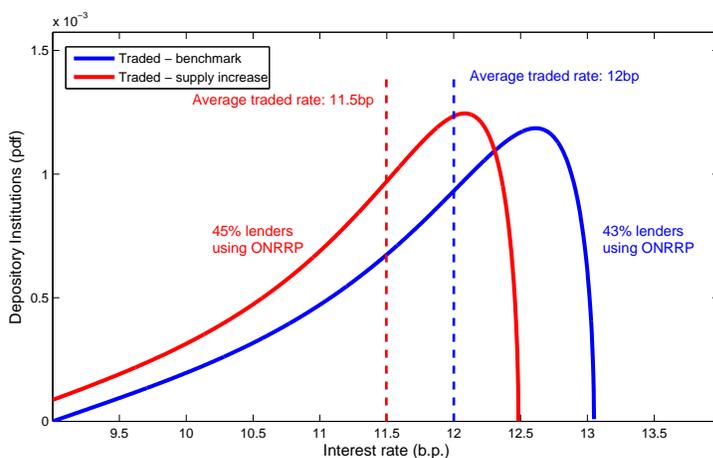
5 A Cap on ON RRP Activity

We have shown that a full-allotment ON RRP facility would be effective at implementing the desired monetary policy. There are, though, concerns that a large usage of the ON RRP facility could have undesired effects on the financial industry and perhaps even have implications for financial stability.[23] Introducing a cap on the ON RRP facility would clearly mitigate its footprint. In this Section we ask whether a cap on ON RRP activity can hinder monetary policy implementation.

5.1 Characterization of Equilibrium with a Cap

Suppose that policymakers place an upper bound on the amount of cash that could be accepted at the ON RRP facility. Formally, if κ denotes the upper bound set by policymakers and ϑ denotes the volume of cash that lenders attempt to exchange at the ON RRP facility, we assume

---

[23]The minutes from the June 2014 FOMC meeting noted that "a relatively large ON RRP facility had the potential to expand the Federal Reserve's role in financial intermediation and reshape the financial industry in ways that were difficult to anticipate." See Frost et al. (2015) for a detailed discussion of potential secondary effects associated with an ON RRP facility.

that each lender is able to use $\frac{\kappa}{\vartheta}$ units of cash in a reverse repo transaction if $\kappa < \vartheta$, and all of their cash otherwise.[24] Recall that any cash that is not deposited with a DI or exchanged at the ON RRP facility is simply stored by lenders—it neither appreciates nor depreciates.

In any candidate equilibrium with cutoff $\bar{c}$ and queue lengths $q(c_j)$ for $c_j \leq \bar{c}$, the measure of lenders that do not match with a DI is

$$\int_0^{\bar{c}} \left[1 - \frac{1 - e^{-q(c_j)}}{q(c_j)} \right] q(c_j) dG(c_j).$$

In words, there is a measure $q(c_j)dG(c_j)$ of lenders in each submarket, and each lender does not match with probability $1 - \frac{1-e^{-q(c_j)}}{q(c_j)}$. Since every lender is endowed with one unit of excess cash, it follows immediately that the supply of cash at the ON RRP facility is given by

$$\vartheta(\bar{c}) = \int_0^{\bar{c}} \left\{ q(c_j) - \left[1 - e^{-q(c_j)} \right] \right\} dG(c_j). \tag{13}$$

In this candidate equilibrium, the expected utility of a lender in a submarket with interest rate ρ_j and queue length q_j can then be written

$$\tilde{u}(\rho_j, q_j) = \left[\frac{1 - e^{-q(c_j)}}{q(c_j)} \right] \rho_j + \left[1 - \frac{1 - e^{-q(c_j)}}{q(c_j)} \right] \tilde{r}(\bar{c}),$$

where

$$\tilde{r}(\bar{c}) = \min \left\{ 1, \frac{\kappa}{\vartheta(\bar{c})} \right\} r. \tag{14}$$

Intuitively, $\tilde{r}$ is the effective (net) rate of return that lenders earn when they approach the ON RRP facility, or what we refer to below as the "effective repo rate."

Characterizing the remaining equilibrium objects follows the analysis in Section 4 very closely. First, imposing that lenders are indifferent between all active DIs, i.e., that $\tilde{u}(\rho_j, q_j) = U$ for all $c_j \leq \bar{c}$, one can solve for the profit-maximizing queue length at each DI, which yields

$$\tilde{q}(c_j; U) = \log \left(\frac{R - c_j - \tilde{r}(\bar{c})}{U - \tilde{r}(\bar{c})} \right). \tag{15}$$

Then, imposing $\tilde{q}(\bar{c}; U) = 0$ implies that

$$\bar{c}(U) = R - U, \tag{16}$$

as in the benchmark model, while the optimal interest rate set by each active DI can be written

$$\tilde{\rho}(c_j; U) = r + \log \left(\frac{R - c_j - \tilde{r}(\bar{c})}{U - \tilde{r}(\bar{c})} \right) \left[\frac{(R - c_j - \tilde{r}(\bar{c}))(U - \tilde{r}(\bar{c}))}{R - c_j - U} \right]. \tag{17}$$

.

---

[24]Alternatively, one could imagine that each lender gets to deposit their unit of cash with probability $\frac{\kappa}{\vartheta}$ if $\kappa < \vartheta$ and probability 1 otherwise; given our specification of linear utility, the two assumptions are identical.

Finally, the market clearing condition is essentially unchanged:

$$\int_0^{\bar{c}} q(c_j) dG(c_j) = \lambda. \tag{18}$$

An equilibrium, then, is a market utility $U^\star$, a cutoff $\bar{c} = \bar{c}(U^\star)$, queue lengths $q(c_j) = \tilde{q}(c_j; U)^\star$ and interest rates $\rho(c_j) = \tilde{\rho}(c_j; U^\star)$ for all $c_j \leq \bar{c}$, a volume of trade at the ON RRP facility $\vartheta = \vartheta\left(\bar{c}(U^\star)\right)$, and an effective interest rate $\tilde{r} = \tilde{r}\left(\bar{c}(U^\star)\right)$ such that (13)–(18) are satisfied.

5.2 The effect of a binding cap

In the benchmark model in which lenders can always access the ON RRP facility, the spread $s = R - r$ completely determines the gains from trade between lenders and DIs, and hence completely pins down DIs' entry decisions and the ensuing allocations. In particular, the spread between the ON RRP rate and the rate of return lenders earn by storing cash themselves – here assumed to be zero – is irrelevant since the latter option is never exercised. When the cap at the ON RRP facility binds, however, this result breaks down: both the spread between the IOER and ON RRP rates, $R - r$, *and* the spread between the ON RRP rate and the rate of return on lenders' own storage technology, $r - 0$, influence entry decisions and allocations.

To illustrate this point, Figure 7 plots the equilibrium distribution of interest rates, volume at the ON RRP facility, and the effective repo rate in the benchmark economy with no cap, and in an alternative economy with a cap equal to 20% of the total cash held by lenders. All other parameters unchanged from the benchmark model in Section 4.

By design, the cap is binding, that is, demand for the ON RRP facility exceeds its limit. As a result, the effective repo rate $\tilde{r}$ drops below the ON RRP rate in proportion to the fraction of cash that a lender can actually exchange at the ON RRP facility. The lenders' outside option worsens and DIs in the market thus offer lower interest rates. However, the average traded rate falls by less than $r - \tilde{r}$. Note that the relevant spread is now given by $R - \tilde{r}$, and thus the fall in the effective repo rate increases the gains from trade and leads additional DIs to become active in the market. This has two effects: first, it increases competition among DIs, which drives posted interest rates up and partially offsets the initial decline; second, it implies that more lenders will match with a DI, which explains the fall in demand at the ON RRP facility.

5.3 When the ON RRP rate fails to be a floor

In Section 4 we showed that in the absence of a cap on ON RRP activity, the average traded rate will always be above the repo rate, that is, the latter is an effective floor for the policy rate. The example above shows that a tight cap puts downward pressure on traded rates, though the average traded rate stayed above the ON RRP rate. This brings us to the next question: Is it possible for the average traded rate to fall below the ON RRP rate and thus outside the target range when the cap is tight? The answer is yes.

It is quite natural to frame the discussion of our results in terms of the measure of lenders, λ, that is, the supply of funds. After all, as in a standard market, an increase in the supply of funds brings a fall in prices or, in this case, the average traded rate, and thus it is the litmus test regarding a floor. For a full-allotment ON RRP facility, the average traded rate is decreasing

18

Figure 7: The Effect of a Cap
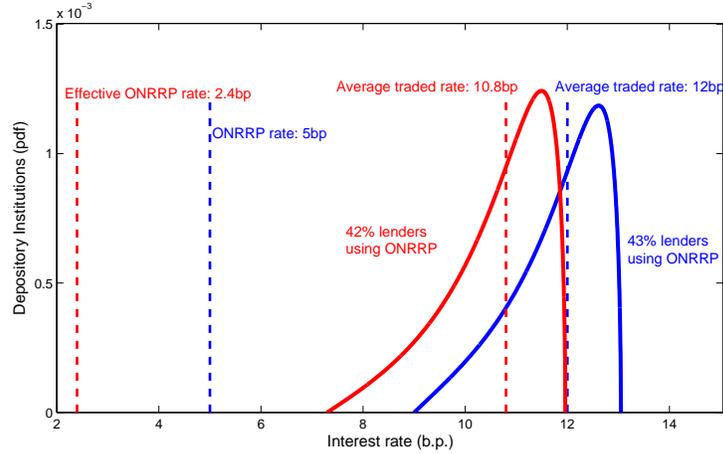
with the measure of lenders but never falls below the repo rate. However, if the ON RRP facility is capped, a large-enough increase in the measure of lenders will see the ON RRP cap bind; then further increases in the supply of funds will eventually see the average traded rate breach the floor of the target range.

More formally, consider any finite cap $\kappa > 0$ and IOER and ON RRP rates, $R > r > 0$. There are two thresholds for the measure of lenders, denoted λ^* and λ^{**}, such that $0 < \lambda^* < \lambda^{**}$ and:

- For all $\lambda^* \leq \lambda$, the cap is binding, $\kappa \leq \vartheta$,

- For all $\lambda^{**} < \lambda$, the average traded rate drops below the ON RRP rate, $\hat{\rho} < r$.

Note that since $\lambda^* < \lambda^{**}$, a binding cap at the ON RRP facility is a necessary, but not sufficient condition for $\hat{\rho} < r$. Recall from Section 4.3 that the average posted rate, and indeed the *minimum* post rate, are strictly above the repo rate, r: lenders must be compensated beyond their outside option to be lured from other DIs offering better rates. Thus when $\lambda = \lambda^*$, that is, take-up on the ON RRP facility exactly equals the cap, the average traded rate is strictly above the ON RRP rate, which is by construction equal to the effective repo rate, $\hat{\rho} > r = \tilde{r}$. As the supply of funds increases further, the effective repo rate drops—and the latter being the actual outside option for lenders, we eventually have $r > \hat{\rho} > \tilde{r}$. We should also emphasize that the exact threshold values λ^* and λ^{**} typically depend on all the parameters of the model.

Figure 8 shows the excess demand at the ON RRP facility (left figure) and the average traded rate (right figure) as a function of the measure of lenders, starting from the previous parameter choice of $\lambda = 1$. We plot three cases: no cap, a quite generous cap of 1.5 lenders, or one-and-a-half the initial number of lenders, and a relatively tight cap of 0.8 lenders. As the left-hand plot in Figure 8 shows, none of the caps is binding at $\lambda = 1$, when a bit more than 40 percent of the lenders bid at the ON RRP facility.[25] The tighter cap starts binding

---

[25]Rates and other parameters are identical to those previously used.

around $\lambda = 1.5$, while the looser cap does at $\lambda = 2.4$. Note that demand at the ON RRP facility continues to grow after the cap binds, albeit at a slower pace as the effective repo rate falls and ON RRP becomes less and less attractive.
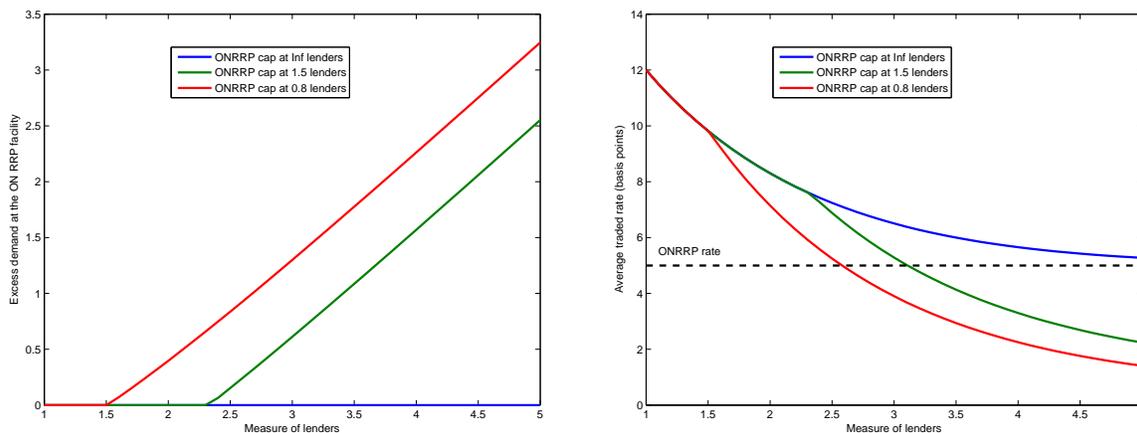

Figure 8: Excess take-up at ON RRP Facility and average traded rates

Let us turn our attention to average traded rates. As long the cap is not binding the average traded rate traces a smooth downward-sloping aggregate demand curve. Rates under a full-allotment ON RRP facility converge to the repo rate as the supply of funds grows unbounded—but never breach the floor as the ON RRP facility stands ready to absorb any take-up volume.

Turning to the finite cap scenarios, it is quite apparent the threshold for a binding cap from the right-hand plot in Figure 8, as the aggregate demand exhibits a kink and the average traded rate then falls at a faster rate with further increases in the measure of lenders. The reason is that the effective repo rate starts dropping with each additional increase in the supply of funds: Lenders not only see further competition to find a DI but also their outside option worsen. Eventually the average traded rate drops below the ON RRP rate and converges smoothly to zero.

As discussed before, the average trade rate is above the ON RRP rate when the cap starts binding. There is actually quite a bit of a difference between the increase in the supply needed to cap the ON RRP facility and to breach the floor set by the statutory repo rate. For example, the tighter cap of 0.8 lenders is binding after a 50 percent increase in the measure of lenders, but it takes thrice as many lenders, a 150 percent increase, for the average traded rate to breach the ON RRP rate.

Let us briefly remark on the role of DI entry in our results. As the measure of lenders expands, more DIs find it profitable to become active in the market as lower rates are more attractive and it is easier to match with a lender. Hence the aggregate tightness of the market, $\lambda/G(c^\star)$, increases by less than the measure of lenders does, sustaining traded rates a bit higher than they would have been so with no entry. Interestingly, once the cap is binding, the rate of entry by DIs speeds up: The lower effective repo rate effectively increases the gains of trade between DI and lender, enticing additional DIs to become active. Thus, the entry margin works

to support traded rates, more so when the cap is binding. The magnitude of this entry channel depends crucially on the balance-sheet costs of DIs that are initially inactive. If there is a large mass of DIs with balance-sheet costs just above the initial threshold $c^\star$, then the market will be able to absorb larger increases in the measure of lenders before the ON RRP cap is binding and the average traded rate drops below the ON RRP rate. In contrast, if infra-marginal DIs have very high balance-sheet costs, there would be little relief from the entry channel.

5.4   Rates and the floor

It is thus clear from the results above that a cap on the ON RRP facility can hinder the implementation of monetary policy. We shall argue next that a cap makes a higher target range for the FFR harder to implement, in the sense that a smaller surge in the supply of funds is needed for the average traded rate to drop below the ON RRP rate. The intuition is quite straightforward. Once the cap is binding, the effective repo rate is the weighted average between the ON RRP rate and zero, with the weight given by the relative demand at the ON RRP facility to the cap. The higher the ON RRP rate is, the faster the effective repo rate falls with an increase in the measure of lenders. Quite naturally, the rate at which the effective repo rate decreases with the supply of funds is mirrored by the average traded rate. Thus the average traded rate breaches the ON RRP rate with a smaller increase in the supply of funds the higher is the target range.

As an illustrative example, assume that the Fed seeks to implement a higher FFR by raising both the IOER and ON RRP by the same amount, that is, keeping the spread between them constant. Figure 9 shows how different values for the target FFR (as given by the range's midpoint, horizontal axis) and for the measure of lenders (vertical axis) map into one of three exclusive cases: ON RRP cap not binding, ON RRP cap binding but the FFR above the ON RRP rate, and the FFR below the ON RRP rate. In other words, Figure 9 plots the thresholds λ^* and λ^{**} as a function of the midpoint target range. We have set the cap value at 0.8 of the lenders.

Figure 9: Mapping a binding cap and a breached floor
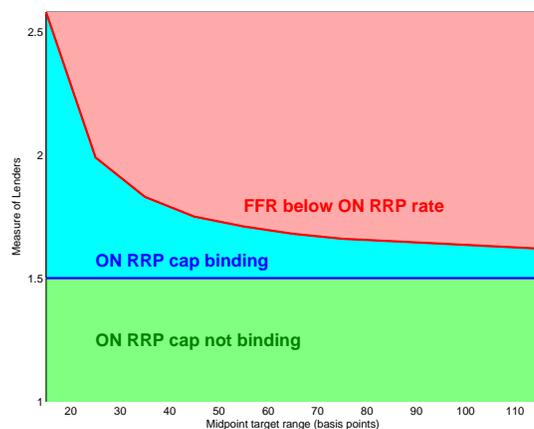
Notice first that the region such that the ON RRP cap is not binding does change with the midpoint of the target range. This follows from our results from Section 4.3: when the cap is not binding, allocations are determined exclusively by the spread between the IOER and the ON RRP, $R - r$, and not their level. By continuity, the threshold λ^*, that is, the smallest measure of lenders such that the ON RRP cap binds, does not depend on the level of rates either.

Things are clearly different for values $\lambda^* < \lambda$, that is, once the ON RRP cap is already binding. The higher the midpoint target range is, the larger is the region such that the FFR drops below the ON RRP rate. In other words, the threshold λ^{**} decreases with the level of rates. For example, a cap of 0.8 seems quite appropriate to implement a midpoint target of 25 basis points, with the FFR breaching the floor only if the measure of lenders doubles. However, if the Fed seeks to implement a midpoint target of 100 basis points, there is substantially less room for error: once the cap is binding, relatively small further increases in the measure of lenders would bring the FFR outside the intended target range.

Finally, we also note that the Fed can alleviate concerns regarding monetary policy implementation by setting a broader target range for the FFR. Figure 10 reproduces the threshold lines of Figure 9 for two range widths, 25 and 35 basis points, with a dashed and solid lines, respectively. The midpoint of the target range is given by the horizontal axis.

Figure 10: Mapping a binding cap and a breached floor: Wider target range
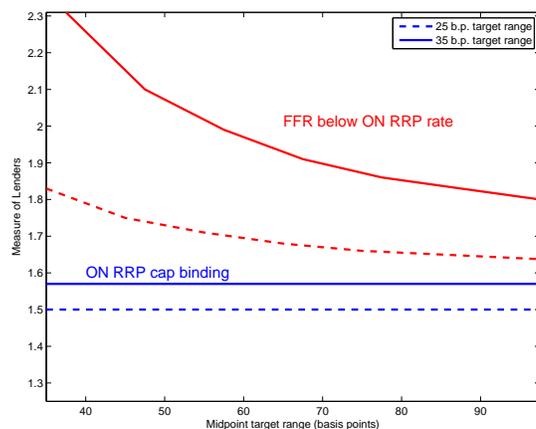
Clearly a wider target range improves the prospects of monetary policy implementation: the market can now absorb more lenders before the ON RRP facility is capped and the FFR drops below the ON RRP rate. The key observation is that by setting a wider range, the gains of trade increase. Additional DIs are now lured into the market, which decreases the demand at the ON RRP facility and explains why, despite having the same cap, it takes a larger surge in the measure of lenders to cap the ON RRP facility under a 35 basis points target range than under a narrower range. The entry margin is also key to understand why the threshold for the FFR to drop below the ON RRP rate is higher, though there is also the simple, mechanical effect that the ON RRP rate is further from the target midpoint under a wider target range.

We should note that our analysis of the width of the target range abstracts from several

considerations. A wider target range means the FOMC would have less control on the actual overnight rates. Indeed, the average traded rate may be further from the range's midpoint than under a narrow range. Day-to-day fluctuations may be larger, particularly around end-of-period dates associated regulatory reports. But perhaps most importantly, confidence by market participants on the implemention the desired policy makes the Desk's job much easier, as Potter (2013) notes. In this context, a wider target range could be interpreted simply as a loss of self-confidence.

6   Achieving liftoff

We have seen that a cap on the ON RRP facility may imperil monetary policy implementation. Yet we should not ignore the concerns regarding the footprint a permanent, full-allotment ON RRP facility would have. The key question thus becomes what is the smallest cap on ON RRP take-up that ensures the FOMC's dictates will be implemented, and thus the credibility of the Federal Reserve System preserved. It is, unavoidably, a quantitative question.

Nowhere the question is as pressing as it is for the very first interest-rate hike or, as it is commonly known, liftoff. There is no precedent for meaningful *positive* rates under the current level of excess reserves—never mind attempting to raise rates. Instructed by the FOMC, the New York Fed has conducted several tests on the ON RRP facility. However, those have been necessarily small-scale tests, as they must be confined within the target range of $0 - 25$ basis points for the FFR. The experiences in other countries are of limited use: corridor systems, where the IOER acts as a floor on rates, have been successful elsewhere but clearly do not work given the particularities of overnight markets in the U.S.

We put our model to the task. Using the limited data available, we discipline the model's parameters and ask under which conditions liftoff will be successful.

6.1   A tentative calibration

We next describe the parameter values chosen for our analysis of liftoff conditions. We have also used these parameters thorough the paper in the numerical illustrations.

We detail first how we map the model's agents and predictions to their counterparts in the data. Our primary interest lies on monetary policy implementation through the IOER rate and the ON RRP facility. It is thus crucial that we start by identifying the agents to have access to each policy tool. The DIs in the model are simply all the depository institutions which qualify to earn IOER—basically, all depository institutions but GSEs. On the lenders' side, the defining characteristic is access to the ON RRP facility, which is established by the Fed: it includes the primary dealers, GSEs, as well as prime money market funds. As discussed in Section 2, not all ON RRP counterparties can participate in the federal funds market and instead they interact through eurodollar contracts. There are currently no publicly available data regarding the rates paid between DIs and ON RRP counterparties in eurodollar contracts. To get around this, we assume that those rates do not differ from those traded in the federal funds market; that is, whether a trade is classified as a federal funds or an eurodollar contract has no impact on the terms of trade. This enables us to map the traded rates in the model to those captured in the data as federal funds transactions. We take some comfort that several studies have described

federal funds and eurodollars as nearly-perfect substitutes, and differences between the federal fund rate and broader eurodollar rates are minimal.[26]

Unfortunately, our calibration is necessarily crude due to the limited experience and data available for the federal funds market under the current situation of excess reserves. The Federal Reserve Bank of New York publishes daily data on the average federal funds rate and the ON RRP facility; the IOER and the ON RRP rates are also public information. However, there is little information regarding the distribution of traded rates and none regarding the traded volume.[27]

Let us start with the policy parameters. The IOER rate has been set to 25 basis points since 2008, so we set $R = 0.25$. The ON RRP facility has been capped at \$300 billions since September 2014, with an ON RRP rate has been set at 5 basis points, $r = 0.05$, until November 2014, when testing on the ON RRP facility began. Over the same period the federal funds rate averaged 12 basis points and ON RRP take-up \$150 billions.

In order to map the model's predictions to dollars, we need to define what a "unit of cash" is, in dollars, which we denote as x. We also choose a Gamma distribution to model the dispersion in balance-sheet costs, with mean and standard deviation denoted μ_c and σ_c, respectively. The parameters left to calibrate are thus $\{\lambda, x, \mu_c, \sigma_c\}$. Unfortunately we do not have data to separately identify λ and x.[28] We thus arbitrarily set $\lambda = 1$.

For the remaining parameters, we make quite straightforward choices. The mean of the distribution of balance-sheet costs is to match an average traded rate of 12 basis points, resulting in $\mu_c = .09$. The dispersion parameter $\sigma_c = .07$ generates a range of traded rates of 4 basis points, which encompasses most of the trade in the federal funds market. A standard deviation of 7 basis points for balance-sheet costs also appears reasonable given that DIs display a wide disparity in regulatory and cash management practices. Finally we set x such that the average ON RRP take-up matches the data.

We have also explored an alternative calibration that seeks to exploit the tests the Federal Reserve Bank of New York conducted, as instructed by the FOMC. Keeping the ON RRP cap at \$300 billion, the ON RRP rate was set at 3, 7, and 10 basis points for several business days each.[29] The tests did elicit some response in the average FFR and ON RRP take-up, though their short duration, combined with large day-to-day variation and seasonal factors, make it difficult to estimate with confidence the effects of the ON RRP rate in the market. In any case, we find that our model is able to fit the observed response by the FFR very well, and is broadly in line with the changes in ON RRP take-up. The parameter values are very close to those reported above, and the results concerning liftoff are basically unchanged.

---

[26]See Bartolini et al. (2008) and Demiralp et al. (2006) as well as the extended discussion in Section 2.

[27]The Federal Reserve Bank of New York also reports the standard deviation, maximum and minimum of the federal funds rates traded per day. We are hesitant to use that data because both the maximum and minimum rates are clearly outlayers. For example, the maximum traded rate is above the IOER rate—occasionally double the IOER rate. As a result, the standard deviation is not very informative of the underlying distribution either.

[28]Volume data by DIs and ON RRP counterparties would allow us to identify the "extensive" margin, λ, from the "intensive" margin, x. The model can be easily extended to allow for heterogeneity in the cash holdings per lender.

[29]The testing schedule was announced several weeks in advance.

6.2 Liftoff

We next describe the liftoff exercise. We assume the FOMC will move the target range for the FFR to $25 - 50$ basis points, setting the IOER rate at the top of the range and the ON RRP rate at the bottom of the range. We will consider three cap values on the ON RRP facility: the current level, \$300 bn.; doubling the current cap, \$600 bn.; and an intermediate value of \$450 bn.

The largest unknown regarding liftoff is how the supply of funds will react. We do not presume to know this, and will instead explore a very large range of possible changes in the supply of funds: from no change to an increase in excess of \$1 trillion. In doing so we provide the model's estimates for the increase in the supply of funds that would cap the ON RRP facility first and then bring the FFR under the floor.

We will assume that the surge in the supply of funds arises, in equal parts, from an increase in the numbers of lenders and an increase in the amount of cash per lender. It turns out this is not innocuous in our model: a surge in the supply of funds driven entirely by the number of lenders puts more downward pressure to the FFR in our model than a comparable supply surge driven exclusively by the amount of cash per lender. Search frictions constrain the number of matches made but, conditional on being matched, a DI and a lender can transact as large as an amount as they wish.[30] Given a constant number of DIs, the number of matches will increase less than proportionally with an increase in the number of lenders thus their probability of matching decreases.

6.3 Results

The key question is how large of a surge in the supply of funds the market can take until the ON RRP facility caps and then how much more until the FFR drops under the floor. Figure 11 answers both questions for the three caps considered. The surge in the supply of funds (in \$ billions) is in the horizontal axis, from zero, or no change, to an increase exceeding \$1 trillion. The vertical axis give the model's prediction for the FFR, in basis points. As none of the caps considered is binding for the baseline parameters, if there is no change in the supply of funds then the FFR will shift by the exact same amount of the target range, that is, increase by 25 basis points to 37 basis points. The kink in the rate schedule marks where the ON RRP cap starts binding, as it did in Figure 8. The dashed line indicates the level of the floor or ON RRP rate and thus the size of a supply surge needed for the FFR to breach the floor, for each cap value.

Table 1 summarizes the results. Under the current cap, it would take close to \$300 bn. of additional supplied funds for the ON RRP facility to become over-subscribed: Note this is close to twice the initial amount of spare capacity at the ON RRP facility of \$150 bn., as our model predicts that the federal funds market will be able to absorb a large fraction of the surge in the supply of funds as rates will be sustained by, among other things, entry of additional DIs seeking to make a profit. The resilience of the federal funds market is also apparent from the gentle decline of the FFR with the surge in supply until the cap binds, as shown in Figure 11. Staying with the current cap, we also see from Table 1, we also see it takes a very substantial

---

[30]Recall that balance-sheet costs are per unit of cash, so total balance-sheet costs do scale up with the amount of the transaction. However, the DI would not be in the market in the first place unless it made a net margin.

| | **Constant supply** | **Surge thresholds** | |
ON RRP Cap	Spare ON RRP capacity	ON RRP capped	Floor breached
$300 bn.	$150 bn.	+ $288 bn.	+ $632 bn.
$450 bn.	$300 bn.	+ $518 bn.	+ $880 bn.
$600 bn.	$450 bn.	+ $753 bn.	+ $1,085 bn.

Table 1: Thresholds and Liftoff

increase for the FFR to breach the floor, $632 bn., or a bit less than twice the initial supply value.

Figure 11: Average traded rate after liftoff
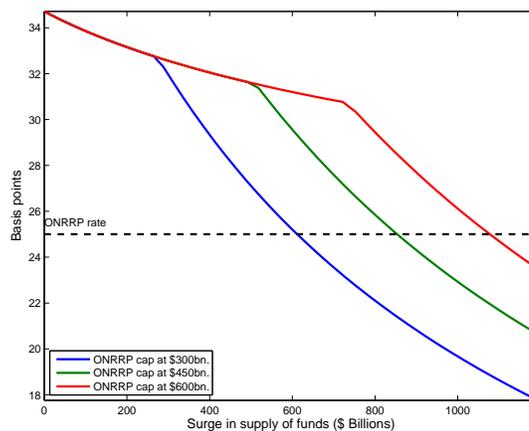
Not surprisingly, the larger the cap the larger the surge in the supply of funds that can be accommodated without compromising monetary policy implementation. It is worth noting, though, that larger caps do not provide much larger room for monetary policy implementation. For example, tripling the initial spare capacity at the ON RRP facility, from a cap of $300 bn. to $600 bn., less than doubles the largest supply surge that can be absorbed before the FFR breaches the floor. In any case, the thresholds are very large. At least at liftoff, even a modest cap on ON RRP facility should not interfere with monetary policy implementation.

Figure 12 provides further details for the intermediate cap value of $450 bn. The left-hand figure reproduces the FFR schedule as well as the effective repo rate, which is obviously equal to the ON RRP rate as long as the cap is not binding, as a function of the *total* supply of funds. The right-hand figure decomposes where the supply of funds ends: traded with a DIs, borrowed by the Fed through the ON RRP facility, or, once the cap binds, as excess demand at the facility.

As discussed above, we had to make several compromises regarding the calibration. To get a sense of the uncertainty regarding the model's predictions, we considered alternative parameter values and, in particular, changes in how the surge in the supply of funds would be split between an increase in the measure of lenders or an increase in the amount of cash per lender. Figure
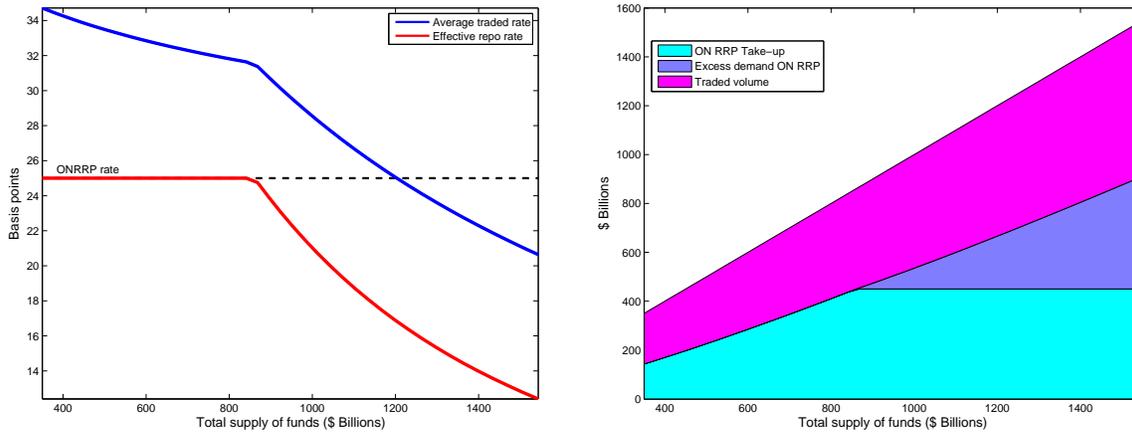
(a) Rates   (b) Volumes

Figure 12: ON RRP facility capped at $450 bn.: Details

13 shows, in a solid line, the 50-50 split we used above for the federal funds rate and ON RRP demand under a cap of $450 bn. The dashed lines indicate the best- and worst-case scenario. This shows that the ON RRP facility could be capped with "only" $300 bn. surge in the supply of funds, compared with the baseline prediction of more than $500 bn. Similarly, the floor could be breached with a relative small $450 bn. surge.

7   Conclusions

We present what we believe to be a competent model of the federal funds market, encompassing the most relevant features and institutions at the current juncture. The model can fit the data, identifying key underlying variables like balance-sheet costs; yet it remains parsimonious enough to allow researchers and policymakers to track the particular mechanisms at work—and extend the model if they wish to do so.

We have also explored the conditions needed for monetary policy implementation, with a focus on those around liftoff. The overall message, we believe, is quite positive: the IOER and the ON RRP facility combine to keep rates in the range of choice by the FOMC, which appears to be a key priority. A cap on the ON RRP facility may hinder implementation, but our best estimates suggest that only in extreme circumstances the FFR would drop outside the target range.
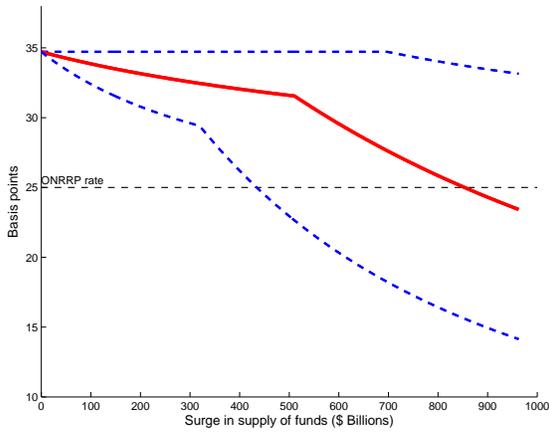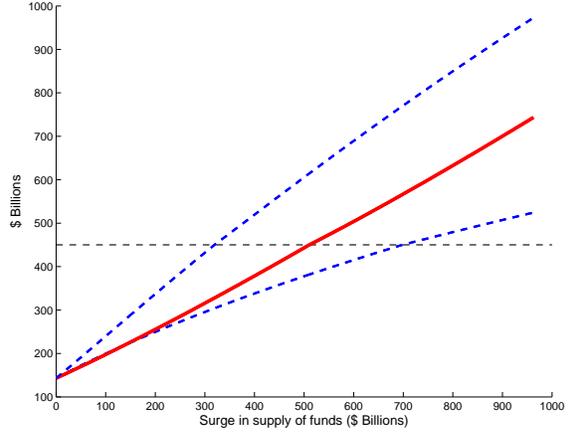
(a) Federal Funds Rate   (b) ON RRP Demand

Figure 13: ON RRP facility capped at \$450 bn.: Robustness analysis

Appendix

Proof of Proposition 1. Let $\Psi(\bar{c})$ denote the left hand side of (9). The result follows immediately from the observations that $\Psi(\bar{c})$ is continuous and strictly increasing over the domain $\bar{c} \in (0, s)$ — since $\Psi'(\bar{c}) = \frac{G(\bar{c})}{s - \bar{c}} > 0$ — with $\lim_{\bar{c} \to 0} \Psi(\bar{c}) = 0 < \lambda$ and $\lim_{\bar{c} \to s} \Psi(\bar{c}) > \lambda$ for any finite λ.

References

Afonso, G., Alex Entz, and Eric LeSeur, "Who is Lending in the Fed Funds Market?," Technical Report, Federal Reserve Bank of New York December 2013. Liberty Street Economics.

_ , _ , and _ , "Whos Borrowing in the Fed Funds Market?," Technical Report, Federal Reserve Bank of New York December 2013. Liberty Street Economics.

_ and R. Lagos, "The Over-the-Counter Theory of the Fed Funds Market: A Primer," *Journal of Money, Credit, and Banking*, forthcoming.

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

Akhtar, M.A., "Understanding Open Market Operations," Technical Report, Federal Reserve Bank of New York 1997.

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

Bartolini, Leonardo, Spence Hilton, and Alessandro Prati, "Money Market Integration," *Journal of Money, Credit and Banking*, 2008, *40* (1), 193–213.

Bech, M. and C. Monnet, "A search-based model of the interbank money market and monetary policy implementation," December 2014. Unpublished manuscript.

Bech, M.L. and E. Klee, "The Mechanics of a Graceful Exit: Interest on Reserves and Segmentation in the Federal Funds Market," *Journal of Monetary Economics*, 2011, *58*, 415–431.

Berentsen, Aleksander and Cyril Monnet, "Monetary policy in a channel system," *Journal of Monetary Economics*, September 2008, *55* (6), 1067–1080.

Bianchi, Javier and Saki Bigio, "Bank, Liquidity Management, and Monetary Policy," September 2014. Federal Reserve Bank of Minneapolis, Research Department Staff Report 503.

Burdett, Kenneth, Shouyong Shi, and Randall Wright, "Pricing and matching with frictions," *Journal of Political Economy*, 2001, *109* (5), 1060–1085.

Demiralp, Selva, Brian Preslopsky, and William Whitesell, "Overnight interbank loan markets," *Journal of Economics and Business*, 2006, *58* (1), 67 – 83.

Ennis, Huberto M. and John A. Weinberg, "Over-the-counter loans, adverse selection, and stigma in the interbank market," *Review of Economic Dynamics*, 2013, *16* (4), 601 – 616.

_ and Todd Keister, "Understanding Monetary Policy Implementation," *Economic Quarterly*, 2008, *94* (3), 235–263. Federal Reserve Bank of Richmond.

Frost, Joshua, Lorie Logan, Antoine Martin, Patrick McCabe, Fabio Natalucci, and Julie Remache, "Overnight RRP Operations as a Monetary Policy Tool: Some Design Considerations," Staff Reports 712, Federal Reserve Bank of New York February 2015.

Furfine, Craig H., "The Microstructure of the Federal Funds Market," *Financial Markets, Institutions & Instruments*, 1999, *8* (5), 24–44.

Hamilton, James D, "The Daily Market for Federal Funds," *Journal of Political Economy*, February 1996, *104* (1), 26–56.

Ho, Thomas S.Y. and Anthony Saunders, "A Micro Model of the Federal Funds Market," *The Journal of Finance*, 1985, *40* (3), 977–988.

Hosios, Arthur J., "On the Efficiency of Matching and Related Models of Search and Unemployment," *The Review of Economic Studies*, 1990, *57* (2), 279–298.

Martin, A., J. McAndrews, Ali Palida, and D. Skeie, "Federal Reserve Tools for Managing Rates and Reserves," September 2013. Federal Reserve Bank of New York Staff Reports, no. 642.

_ , **Jaime McAndrews, and David Skeie**, "Bank Lending in Times of Large Bank Reserves," June 2013. Federal Reserve Bank of New York Staff Reports, n. 497.

Moen, Espen, "Competitive Search Equilibrium," *Journal of Political Economy*, 1997, *105*, 385–411.

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

Potter, S., "Recent Developments in Monetary Policy Implementation," Technical Report, Federal Reserve Bank of New York December 2013.

Potter, Simon, "Interest rate control during normalization," Technical Report, Federal Reserve Bank of New York 2014.

Whitesell, William, "Interest rate corridors and reserves," *Journal of Monetary Economics*, September 2006, *53* (6), 1177–1195.