Can the US interbank market be revived?*

Kyungmin Kim[†], Antoine Martin[‡], and Ed Nosal[§]

November 5, 2018

Abstract

Large-scale asset purchases by the Federal Reserve as well as new Basel III banking regulations have led to important changes in U.S. money markets. Most notably the interbank market has essentially disappeared with the dramatic increase in excess reserves held by banks. We build a model in the tradition of Poole (1968) to study whether interbank market activity can be revived if the supply of excess reserves is decreased sufficiently. We show that it may not be possible to revive the market to pre-crisis volumes due to costs associated with recent banking regulations. Although the volume of interbank trading may initially increase as excess reserves continue to decline, the new regulations may engender changes in market structure that result in interbank trading being completely replaced by non-bank lending to banks when excess reserves become scarce. This non-monotonic response of interbank trading volume to reductions in excess reserves may lead to misleading forecasts about future fed funds prices and quantities when/if the Fed begins to normalize their balance sheet by reducing excess reserves.

Keywords: Interbank market, monetary policy implementation, balance sheet costs
JEL classification: E42, E58

*The views expressed in this paper are those of the authors and do not necessarily reflect those of the Federal Reserve Banks of Atlanta and New York, the Federal Reserve Board or the Federal Reserve System.

[†]Federal Reserve Board; kyungmin.kim@frb.gov

[‡]Federal Reserve Bank of New York; antoine.martin@ny.frb.org

[§]Federal Reserve Bank of Atlanta; ed.nosal@gmail.com

1 Introduction

Money markets in the U.S. have changed dramatically since the 2007-08 financial crisis. One of the most striking changes is the decrease in the size of the interbank market. Before the financial crisis the size of the interbank market was estimated to be about $100 billion per day. Today (2018) it is less than $5 billion. The change in the size of the interbank market can be attributed in part to both the Federal Reserve's monetary policy measures to stimulate the economy and new Basel III requirements that have been imposed on banks.

The reasons that underlie the significant decline in the size of the U. S. interbank market are well understood. Before the crisis the Fed relied on scarce reserves and reserve requirements to implement monetary policy.[1] As payment shocks affected banks' reserve holdings throughout the day they would trade with each other to ensure that they held just enough reserves to satisfy their reserve requirements and no more. Starting in 2009 the Fed increased the supply of reserves through large-scale asset purchases ultimately injecting almost $3 trillion of reserves into the banking system. As a consequence almost all banks held reserves that far exceeded what was required. This essentially eliminated the need for banks to trade with each other to offset their payment shocks.

In this paper we ask if and under what conditions the U.S. interbank market can be revived. Contrary to what is commonly believed, draining reserves by reducing the Fed's balance sheet will not necessarily revive the interbank market. We argue that although a small increase in the volume of interbank trading will likely appear at the early stage of reserve draining it is uncertain whether the market revival will continue with further draining of reserves. Indeed we show that in some circumstances interbank trading volumes could completely disappear as reserves are drained further. These results indicate that regulatory and monetary frameworks that need a revival of the interbank market—because they rely on benchmark rates generated largely from interbank activity—may no longer be viable even if the Fed's balance sheet shrinks to pre-crisis levels.

The key insight that underlies these results is related to stricter banking regulations which impose increased "balance sheet" costs on banks. We think of these costs as being primarily related to the size and not the composition of banks' balance sheets. Examples of such balance sheets costs are the Basel III leverage ratio and the Federal Deposit Insurance Corporation's (FDIC) assessment fee.

[1] See Ennis and Keister (2008) for a theoretical exposition.

The patterns of trade in money markets influence the size of banks' balance sheets which in turn affect those banks' balance sheet costs. To understand the link between balance sheets costs, the size of the balance sheets and money market activities, consider a simple example with one cash-rich non-bank and two banks. Suppose the non-bank does not renew a loan it made to bank 1 and instead lends to bank 2. Everything else being equal bank 1 will see its reserves decrease and bank 2 will see its reserves increase by the amount of the loan. The movement in reserves changes the size of the balance sheet of each bank but not the aggregate size of balance sheet of the banking system. Hence the non-bank loan to bank 2 (away from bank 1) does not affect the aggregate balance sheet cost of the banking system.

Consider now a different set of transactions that result in the same movement of reserves from bank 1 to bank 2. In this case the non-bank renews its loan with bank 1 and bank 1 makes an interbank loan to bank 2. This set of transactions increases the size of the balance sheet for the banking system. As above the balance sheet of bank 2 increases by the amount of reserves it has received. However the balance sheet of bank 1 does not decrease since the interbank loan replaces the reserves lent to bank 2 on the asset side of its balance sheet. Since the aggregate size of the balance sheet of the banking system increases by the amount of interbank trades, money market participants have an incentive to avoid these trades if they can collectively benefit from reducing balance sheet costs.

We develop a model in the spirit of Poole (1968) to formalize this argument. Consistent with actual market practice non-banks make loans to banks before interbank trading takes place. Interbank trades can partially offset payment shocks by redistributing reserves among banks. What is new compared to the basic Poole (1968) environment is the idea that non-banks can also delay lending to the banks until after the payment shock is realized but it is costly to do so. For simplicity we assume that this delayed lending happens at the same time as the interbank market operates. If banks' balance sheet costs are zero then interbank trading is costless and banks can completely offset payment shocks by relying solely on the interbank market. In this situation non-banks do not have an incentive to pay the cost associated with delaying a loan to a bank. When balance sheet costs are strictly positive interbank trades become costly. Since interbank trading increases the aggregate balance sheet cost of the banking sector, banks will not fully offset payment shocks just through interbank trades. Late loans from non-banks made after payment shocks can help offset the shocks more fully and thus are more valuable to banks than early loans made before the

shocks. Therefore banks are willing to pay higher interest for late borrowing. At the same time a higher rate provides an incentive for non-banks to incur the delay cost and provide late loans to the bank.

We show that if non-banks face a per dollar cost of loan delay some interbank trading can reappear when reserves become scarce. However the level of interbank trade will fall short of the pre-crisis levels due to both higher post-crisis regulatory costs and the existence of late loans. We also consider a case where there is a system-wide fixed cost associated with delaying loans. The fixed cost could for example represent the cost of establishing market standards or infrastructure that facilitates late day trading for non-banks with banks. In this case as the supply of reserves decreases interbank trade will reappear initially. However if the supply continues to decrease, the system-wide benefit of making late loans to banks will exceed the fixed cost at some point. If the fixed cost is paid, then the volume of interbank market trading will shrink to zero since banks can obtain needed reserves from non-banks.

Generally our analysis indicates that it may be difficult for the U.S. to have an active interbank market even if excess reserves are significantly drained. The main reason is that due to the new post-crisis regulations interbank trading activity has become expensive. We argue that it may not help policymakers to experiment with small changes in the level of reserves to assess whether the interbank market can be revived because interbank market volume can either respond non-monotonically to a draining of reserves or may can initially increase and remain relatively constant as reserves are further drained.

This result is important for monetary policy because it shows that implementation relying on the revival of the interbank market may not be feasible. In addition it is relevant for the ongoing debate regarding the desirability of interbank trading activity over trading between banks and non-banks for monetary policy implementation purposes. For example Bindseil (2016) argues that a central bank's "operational framework should not undermine incentives for active interbank [...] markets." Similarly the Norges Bank changed its monetary policy implementation framework in part to generate more interbank activity (see Norges Bank (2011)).

Our main contribution is to create a parsimonious framework that shows how a simple increase in balance sheet costs can affect activities in the fed funds market and broader money markets. As far as we know, we are the first to seriously evaluate the prospect of the revival of the interbank market with a theoretical model that can explain important

features of money markets before and after the 2008 financial crisis. Our paper is related to a growing literature on money markets and monetary policy implementation. Recent contributions include Chen et. al. (2016) which focus on monetary policy implementation in a model with preferred habitat features. Martin et. al. (2013) focus on tools available to the Federal Reserves to implement monetary policy with a large supply of reserves. Afonso and Lagos (2015) develop a search model to understand trade dynamics in the Federal Funds market. Finally, Armenter and Lester(2017) use a model with directed search to study the Federal Reserve's overnight reverse repurchase agreement facility.

The remainder of the paper is structured as follows. Section 2 describes the baseline model. Section 3 characterizes and discusses the equilibrium in the baseline model. Through the lens of the baseline model, section 4 discusses money market activity before and after the 2008 financial crisis as well as the markets' response to a reduction in the supply of reserves. In section 5 we extend the baseline model by allowing non-bank lenders to lend in the late market, and discuss its implications for the future of the interbank market. Section 6 concludes.

2 Baseline model

There are three types of agents—a non-bank or investor, two commercial banks and a central bank—and three periods that can be viewed as unfolding over a day. The periods are interpreted as morning, early afternoon and late afternoon/evening to resemble actual U.S. money markets. In period 1 (the morning) the investor decides how to allocate resources between the two banks over periods 1 and 2 by making deposits with the banks. In period 2 (the early afternoon) banks trade in the interbank market and may receive additional deposits from the investor. In period 3 (the late afternoon and extending into the beginning of the next day) banks borrow from the discount window if needed, receive interest on reserves from the central bank and repay interbank loans.

We first describe the endowments, feasible actions and behavior of each type of agent. Then we characterize the shocks that hit the economy.

2.1 Investor

The investor is endowed with M units of funds. At the beginning of period 1 the investor decides how to allocate M between the two banks by making deposits in periods 1 and 2.[2] The investor can, for example, allocate all of M to period 1 bank deposits or can allocate some of M to period 1 deposits and the remainder to period 2 deposits. Since a deposit is a loan we will use these two terms interchangeably.

The deposit markets in periods 1 and 2 are competitive. The returns to the investor's deposits are realized at the end of period 3. There is a cost associated with delaying a bank deposit until period 2. In the baseline model we assume this cost is prohibitive which implies that the investor deposits the entire endowment M in period 1. We relax the prohibitive cost assumption in section 5.

The investor is risk neutral and allocates funds to maximize its net expected payoff.

2.2 Banks

There are two banks indexed by $i = 1, 2$. Regulation requires that bank i hold at least R_i reserves at the end of period 3. Reserves can only be held by the banks and for simplicity we assume that the only assets that banks hold are reserves and interbank loans. Banks buy reserves by borrowing in competitive deposit markets in both periods 1 and 2 and by borrowing in the competitive interbank market in period 2. Denote the period 1 bank deposit rate as r_D. We denote the period 2 interbank rate r_R. Below we show that the period 2 interbank rate must be equal the period 2 deposit rate. In period 3 a bank can borrow at the central bank's discount window. A bank must borrow from the discount window if its reserve holdings at the beginning of period 3 falls short of what is required R_i or in other words if its excess reserves are negative. The discount window borrowing rate is r_W and the interest paid on positive excess reserves held at the Fed is r_E.

Banks are risk neutral and maximize profits. A bank's profit is given by the total return on its assets minus the sum of the cost of its liabilities and balance sheet costs that result from regulations. We assume that balance sheet costs depend only on the size of the bank's balance sheet and not for example on the composition of the its assets and liabilities. Modeled in this way balance sheet costs are similar to the costs incurred by banks that face the leverage

[2]The period 2 deposit has to physically rest somewhere in period 1. We can think of the funds as sitting at the two banks but not be renumerated for period 1. When the funds are allocated in period 2 they will receive the period 2 deposit rate.

ratio constraint imposed by recent Basel III regulations or the revised FDIC assessment fee that domestic deposit-taking institutions pay. For simplicity we assume that if bank i's total asset holdings are A_i, then it incurs a balance sheet cost equal to $c_B A_i$ where c_B is a positive constant. In this formulation balance sheet costs are linear where marginal and average balance sheet costs are equal to one another.

2.3 Central bank

The central bank does not make any optimizing decisions. The central bank determines the supply of reserves, sets interest on excess reserves r_E and the cost of borrowing reserves from the discount window r_W where $r_E < r_W$. We assume that the economy has M reserves at the beginning of period 1 which is equal to the endowment of the investor. This assumption is not essential but it does simplify the presentation. None of our results are affected if we instead assume that the supply of reserves differs from the investor's endowment.[3]

2.4 Payment/reserve shocks

Banks are hit by shocks that affect the distribution of reserve holdings between periods 1 and 2 and periods 2 and 3. Between periods 1 and 2 bank i receives a shock η_i to its reserve holdings. We assume that η_i is uniformly distributed over $[-\bar{\eta}, \bar{\eta}]$. This assumption greatly simplifies the analysis but relaxing it does not qualitatively affect our results.[4] We interpret the η shock as a movement in reserves between banks 1 and 2 which means that $\eta_1 = -\eta_2$.[5] In a richer model one could imagine that agents initiate wire transfers between their accounts at the two banks for idiosyncratic reasons at the end of period 1 and the transfers are settled in between periods 1 and 2.

Between periods 2 and 3 banks receive another reserve shock ν_i. This shock is similar to the shock to reserves in the Poole (1968) model. We assume that ν_i is uniformly distributed

[3] If the supply of reserves exceeds M the difference can be held by banks in the form of endowed equity. If the supply is smaller than M bank liabilities/assets exceed M and the difference can be held as securities by banks.

[4] The assumption that the support of η_i has finite lower and upper bounds is sufficient for most results. This assumption is realistic because it is hard to imagine a shock that is too large relative to the stock of reserves M.

[5] Relaxing this assumption complicates the model while somewhat preserving the results. We can think of $\eta_1 + \eta_2$ as a shock to total reserve supply and $\eta_1 - \eta_2$ as a differential flow from bank 2 to bank 1. The correlation between $\eta_1 + \eta_2$ and $\eta_1 - \eta_2$ is generally not zero so it is not possible to cleanly separate their effects.

over $[-\bar{\nu}, \bar{\nu}]$. These assumptions greatly simply the equilibrium expressions and deliver clear interpretations of the model.[6]

From the banks' point of view there is a conceptual difference between the η shocks and the ν shocks. Since the interbank and loan markets operate after the η shock is realized banks are able to trade with one another or raise additional deposits to offset this shock. In contrast the money market is closed after the ν shock is realized. As a result banks respond to ν shocks passively by either borrowing reserves at the discount window or holding excess reserves at the Fed.

3 Equilibrium in the baseline model

The equilibrium is described and characterized by solving the model backward starting with the last period.

3.1 Period 3: late afternoon/evening

A bank's profit depends in part on its excess reserve holdings after the ν shock which is realized between periods 2 and 3. The excess reserve holdings of bank i at the beginning of period 3 are the sum of four components and denoted by e_i:

1. Excess reserves held at the end of period 1 denoted x_i;
2. The interbank shock η_i realized between periods 1 and 2;
3. Reserves borrowed in the period 2 interbank market denoted y_i; and
4. The ν_i shock to reserve holdings realized between periods 2 and 3.

Hence $e_i \equiv x_i + \eta_i + y_i + \nu_i$. Notice that $y_i < 0$ means that bank i lends in the interbank market.

If excess reserves are negative then the bank must borrow reserves equal to $-e_i$ from the central bank. Bank i's total reserve holdings will then equal what is required, R_i. Notice

[6] Our assumptions eliminate heterogeneity and nonlinearity in marginal values making the two banks' preferences basically identical. The uniform distribution assumption introduces linearity in the marginal value of reserves because the density of ν_i is constant over $[-\bar{\nu}, \bar{\nu}]$.

the bank has no incentive to borrow more than $-e_i$ since $r_E < r_W$. The final payoff or profit for bank i that enters period 3 with negative excess reserves $e_i < 0$ is given by

$$R_i r_{RR} - (R_i + x_i + \eta_i + \nu_i) r_D - y_i r_R + e_i r_W - (R_i + [-y_i]^+) c_B \tag{1}$$

where r_{RR} represents the interest earned on required reserves, $[x]^+ = x$ if $x \geq 0$ and $[x]^+ = 0$ if $x < 0$.[7] The first term $R_i r_{RR}$ represents the interest income the bank earns from its required reserve holdings. The next three terms $-(R_i + x_i + \eta_i + \nu_i) r_D$, $-y_i r_R$ and $e_i r_W$ represent the cost of the bank's liabilities and depend on whether the liability is a deposit, an interbank loan or a discount window loan respectively. If y_i is negative, meaning that the bank lends in the interbank market, then $-y_i r_R$ represents the interest income from an interbank loan which is an asset. The final term $-(R_i + [-y_i]^+) c_B$ represents the bank's balance sheet cost since the only assets that banks hold are reserves and interbank loans. If $e_i < 0$ bank i will ultimately hold R_i reserves—since it will borrow $-e_i$ at the central bank's discount window. If $y_i < 0$ then bank i has an interbank loan which is an asset and generates a balance sheet cost.

If a bank's excess reserves are positive then it earns r_E on its excess reserves $e_i > 0$. The final payoff (or profit) for bank i is

$$R_i r_{RR} + e_i r_E - (R_i + x_i + \eta_i + \nu_i) r_D - y_i r_R - (R_i + e_i + [-y_i]^+) c_B. \tag{2}$$

The first two terms represent the interest income the bank earns on the required and excess reserves respectively. The next two terms are the costs associated with the bank's liabilities (if $y_i < 0$ then $-y_i r_R$ represents interbank interest income). The last term is the bank's balance sheet cost: Holding excess reserves $e_i > 0$ adds to the bank's balance sheet costs. Notice that (2) does not have a term associated with the discount window rate r_W since the bank does not borrow from the central bank.

The terms in (1) and (2) that include R_i, η_i or ν_i are from the banks' point of view exogenous. For simplicity and without loss of generality we suppress them.[8] The banks' period 3 payoff functions (1) and (2)—net of the exogenous terms—can be compactly expressed as

$$-x_i r_D - y_i r_R + [e_i]^+ r_E - [-e_i]^+ r_W - ([e_i]^+ + [-y_i]^+) c_B. \tag{3}$$

[7] Currently in the U.S. the interest on required reserves is equal to the interest on excess reserves. But in principle the two rates could be different.

[8] Prices or rates are not exogenous to the model but are exogenous to banks because markets are assumed to be competitive.

3.2 Period 2: interbank market

In period 2 banks trade with each other in an interbank market. Bank i chooses the amount of reserves it wants to borrow y_i so as to maximize its expected period 3 payoff. Bank i enters period 2 with excess reserves $x_i + \eta_i$: It obtained excess reserves equal to x_i in period 1 via investor deposits and experienced a payment shock equal to η_i between periods 1 and 2. We now describe a bank's optimal behavior.

With the help of (3) bank i's period 2 interbank borrowing/lending problem can be described by

$$\max_{y_i} \Big\{ - x_i r_D - y_i r_R - [-y_i]^+ c_B + E\{[e_i]^+ (r_E - c_B) - [-e_i]^+ r_W\} \Big\}$$

where $e_i = x_i + \eta_i + y_i + \nu_i$. Since $-x_i r_D$ is exogenous from a period 2 perspective—it was determined in period 1—it is irrelevant for the bank's period 2 decision problem. It will be convenient to write the bank's problem as

$$\max_{y_i} [-y_i r_R - [-y_i]^+ c_B + v_i(z_i)], \tag{4}$$

where $z_i \equiv x_i + \eta_i + y_i$, and

$$v_i(z_i) \equiv \int_{-\bar{\nu}}^{\bar{\nu}} \{[z_i + \nu_i]^+ (r_E - c_B) - [-(z_i + \nu_i)]^+ r_W\} (2\bar{\nu})^{-1} d\nu_i.$$

Since ν_i is uniformly distributed on $[-\bar{\nu}, \bar{\nu}]$ its probability density function is $(2\bar{\nu})^{-1}$ on $[-\bar{\nu}, \bar{\nu}]$ and zero otherwise. We assume that ν_1 and ν_2 have the same distribution, which implies that the functional form of $v_i(z)$ does not depend on i. Therefore we can drop the subscript i from $v_i(z)$ and let $v(z_i)$ represent the expected benefit of having z_i excess reserves at the end of period 2.

Notice that $v(z_i)$ is strictly concave over $z_i \in [-\bar{\nu}, \bar{\nu}]$ since $v'(z_i)$ is linear and decreasing. To see this take the derivative of $v(z_i)$ with respect to z_i to get

$$v'(z_i) = (r_E - c_B) \cdot \Pr(\nu \geq -z_i) + r_W \cdot \Pr(\nu \leq -z_i).$$

The probability $\Pr(\nu \geq -z_i)$ is strictly increasing and $\Pr(\nu \leq -z_i)$ is strictly decreasing in z_i over $[-\bar{\nu}, \bar{\nu}]$. Since ν is uniformly distributed over $[-\bar{\nu}, \bar{\nu}]$, (3.2) can be rewritten as

$$v'(z_i) = \begin{array}{ccc} r_E - c_B & & \text{if } z_i \geq \bar{\nu} \\ \frac{r_E - c_B + r_W}{2} - s z_i & \text{if} & -\bar{\nu} \leq z_i \leq \bar{\nu} \\ r_W & & \text{if } z_i \leq -\bar{\nu} \end{array}, \tag{5}$$

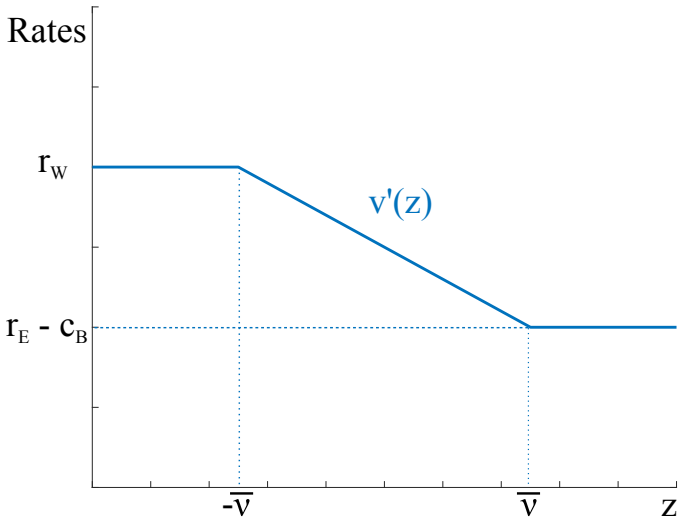

Figure 1: demand for excess reserves in period 2.

where $s = (r_W - r_E + c_B)/(2\bar{\nu})$. It is obvious from (5) that $v'(z_i)$ is linear and decreasing in z_i when $-\bar{\nu} \leq z_i \leq \bar{\nu}$. Figure 1 illustrates $v'(z_i)$. The middle expression in (5) describes a situation that we define as scarce excess reserves in period 2. Here scarcity means there is a strictly positive probability that the bank's excess reserves will be negative at the beginning of period 3 and a strictly positive probability that they will be positive.[9] Intuitively as z_i approaches $-\bar{\nu}$ there is a high probability that bank i will have to borrow reserves from the discount window and as a result has a marginal valuation of reserves that is close to discount window rate r_W. Similarly as z_i approaches $\bar{\nu}$ there is a high probability that bank i will have positive excess reserves and as result has a marginal valuation of reserves that is close to the rate on excess reserves net of marginal balance sheet costs $r_E - c_B$.

After the banks receive their η shocks (between periods 1 and 2) one of the banks will have more excess reserve holdings than the other. We assume without loss of generality that it is bank 1. That is $x_1 + \eta_1 \geq x_2 + \eta_2$. This inequality implies that bank 1 is the potential lender in the interbank market and bank 2 is the potential borrower.

Lemma 1. *In any equilibrium bank 1 lends to bank 2,* $y_1 \leq 0$.[10]

Moreover bank 2 does not borrow too much from bank 1 in the following sense:

[9] We can say that a bank's excess reserves are "really scarce" if $z_i < -\bar{\nu}$ and "abundant" if $z_i > \bar{\nu}$. If reserves are really scarce then there is a zero probability that the bank will have positive excess reserves after the ν shock; if they are abundant then there is a zero probability that the bank will have negative excess reserves after the ν shock.

[10] All proofs can be found in the appendix.

Lemma 2. *In any equilibrium bank 2's excess reserve holdings never exceed those of bank 1 after interbank trading,* $x_1 + \eta_1 + y_1 \geq x_2 + \eta_2 - y_1$.

Lemmas 1 and 2 imply that

$$x_2 + \eta_2 \leq x_2 + \eta_2 - y_1 \leq x_1 + \eta_1 + y_1 \leq x_1 + \eta_1.$$

After interbank market closes the excess reserve holdings of both banks lie between their pre-interbank-market excess reserves $x_2 + \eta_2$ and $x_1 + \eta_1$.

Given lemmas 1 and 2 excess reserves will always be scarce in period 2 if for all $\eta_i \in [-\bar{\eta}, \bar{\eta}]$, $-\bar{\nu} \leq x_i + \eta_i \leq \bar{\nu}$ for $i = 1, 2$. Hence we can now define excess reserve scarcity in period 2 as

Definition 1. *Excess reserves are said to be scarce in period 2 if*

$$-\bar{\nu} + \bar{\eta} \leq x_i \leq \bar{\nu} - \bar{\eta} \tag{6}$$

which implies that $v''(x_i + \eta_i) < 0$ *for all* $\eta_i \in (-\bar{\eta}, \bar{\eta})$.

Unless otherwise specified we shall assume condition (6) prevails.[11]

Bank 1's lending decision y_1 is given by the solution to the problem in (4) which is

$$v'(x_1 + \eta_1 + y_1) + c_B \geq r_R$$

with equality if $y_1 < 0$. Bank 2's borrowing decision $y_2 = -y_1$ is given by

$$v'(x_2 + \eta_2 - y_1) \leq r_R \tag{7}$$

with equality if $y_1 < 0$. Notice that we have imposed the condition $y_2 = -y_1$ in (7).

If $y_1 < 0$ in equilibrium then we have an active interbank market characterized by

$$v'(x_2 + \eta_2 - y_1) - v'(x_1 + \eta_1 + y_1) = c_B. \tag{8}$$

This outcome is possible if and only if $v'(x_2 + \eta_2) - v'(x_1 + \eta_1) > c_B$. Hence the interbank market will be inactive in equilibrium and $y_1 = y_2 = 0$ whenever

$$v'(x_2 + \eta_2) - v'(x_1 + \eta_1) \leq c_B. \tag{9}$$

This inequality has a nice interpretation: The interbank market will shut down if the marginal balance sheet cost c_B exceeds the marginal gain from borrowing and lending

[11] Notice that x_i is a period 1 endogenous variable. Below we will restate the definition of scarcity using only model parameters and/or exogenous variables.

$v'(x_2 + \eta_2) - v'(x_1 + \eta_1)$. When the interbank market is inactive the interbank rate r_R is indeterminate. In fact equilibrium will be consistent with any interbank rate that satisfies

$$v'(x_2 + \eta_2) \leq r_R \leq v'(x_1 + \eta_1) + c_B. \tag{10}$$

When the interbank rate is indeterminate we will for convenience define it to be the average of the two limiting values in (10), $r_R \equiv 0.5[v'(x_1 + \eta_1) + v'(x_2 + \eta_2) + c_B]$

If the interbank market is active, $y_1 < 0$, then the interbank rate satisfies

$$r_R = v'(x_2 + \eta_2 - y_1) = v'(x_1 + \eta_1 + y_1) + c_B. \tag{11}$$

Clearly (11) and (8) plus some simple arithmatic implies that

$$r_R = \frac{1}{2}[v'(x_2 + \eta_2 - y_1) + v'(x_1 + \eta_1 + y_1)] + \frac{1}{2}c_B. \tag{12}$$

Since $\eta_1 + \eta_2 = 0$ and v' is linear—see (5)—(12) can be simplified to read

$$r_R = \frac{1}{2}[v'(x_2) + v'(x_1)] + \frac{1}{2}c_B = v'(\frac{x_1 + x_2}{2}) + \frac{1}{2}c_B. \tag{13}$$

This equation is rather interesting: The interbank rate depends on the period 1 excess reserve holdings of the banks and the marginal balance sheet costs but not on the size of η shocks. The latter is an implication of assumed absense of aggregate shocks to reserve supply $\eta_1 + \eta_2 = 0$ and the linearity of the marginal value function $v'(z)$. Note that the expression for r_R in (13) equals what we defined the interbank rate to be when the market was inactive $y_1 = y_2 = 0$.

We can use (13) along with (5) to get an explicit expression for interbank trade volume $-y_1$. In particular interbank trade volume is given by

$$-y_1 = [\frac{1}{2}(x_1 - x_2) + \eta_1 - \frac{c_B}{2s}]^+, \tag{14}$$

where s is the slope of v' over its decreasing region. Notice that trade volume increases with the difference in excess reserves holdings by banks at the end of period 1, $x_1 - x_2$, and the interbank shock η_1 and decreases with balance sheet costs c_B but is independent of the amount of (scarce) excess reserves that are in the economy.[12]

Figure 2 provides an illustration of equilibrium interbank trade. The lender's marginal value curve lies c_B above the borrower's curve. The difference between $x_1 + \eta_1$ and $x_2 - \eta_1$

[12]We shall see that in equilibrium $x_1 = x_2$. Hence, from (14), a necessary condition for strictly positive interbank trade volume is $c_B < 2s\bar{\eta}$.

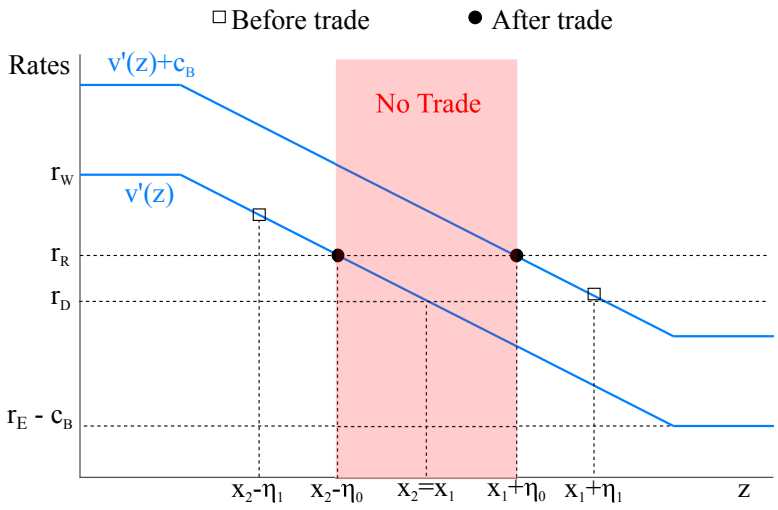

Figure 2: interbank trade.

is sufficiently large so that the interbank market is active as indicated by the squares on the piecewise linear curves where $v'(x_2 - \eta_1) > v'(x_1 + \eta_1) + c_B$. Bank 1 extends a loan to bank 2 of size $-y_1$ such that $v'(x_1 + \eta_1 + y_1) + c_B = v'(x_2 - \eta_1 - y_1) = r_R$. Notice that the average after-interbank-trade excess reserve holdings of the banks is equal to $(x_1 + x_2)/2$ and that the geometric construction that underlies figure 2 implies that $r_R = v'[(x_1+x_2)/2]+(1/2)c_B$. Define $\eta_0 = c_B/(2s)$ to be an η shock such that banks are indifferent between trading and not trading in the interbank market as illustrated in figure 2. (In figure 2 $\eta_0 \equiv \eta_1 + y_1$.) Then if $\eta_1 > \eta_0$ there is a positive trade and for any $\eta_1 < \eta_0$ the interbank market will be inactive since $v'(x_2 + \eta_2) - v'(x_1 + \eta_1) < c_B$.

3.3 Period 1: demand for deposits

In period 1 banks accumulate reserves by issuing deposits to the investor. Bank i chooses an amount of deposits R_i+x_i taking the deposit rate r_D as given. When making this decision banks anticipate that they will be hit by payment shocks between periods 1 and 2—that they can partially offset in the interbank market—as well as another round of payment shocks between periods 2 and 3—that they cannot offset in a market.

Bank i's choice of excess reserves in period 1 x_i can be compactly expressed as the solution to the following problem

$$\max_{x_i} u(x_i) - x_i r_D$$

where $u(x_i) = Ew(x_i, \eta_i, r_R)$ and

$$w(x_i, \eta_i, r_R) \equiv \max_{y_i}[-y_i r_R - [-y_i]^+ c_B + v(x_i + \eta_i + y_i)]. \tag{15}$$

The function $u(x_i)$ is the expectation of the bank's maximized period 2 objective function $w(x_i, \eta_i, r_R)$—or equivalently (4)—where y_i is chosen as a function of η_i and r_R and the expectation of $w(x_i, \eta_i, r_R)$ is taken over the distribution of η_1. Note that $Ew(x_i, \eta_i, r_R)$ generally depends on both x_i and the label i, but in equilibrium we can ignore the dependence on i.[13] In making their decisions banks take the interbank rate r_R as given. Notice from (13) that the interbank rate r_R is not a random variable even from period 1 perspective since it does not depend on either η or ν.[14] In light of (15) it is not surprising that $u(x_i) = Ew(x_i, \eta_i, r_R)$ inherits some properties of $v(z_i)$. One important property that $u(x_i)$ inherits is strict concavity.[15]

Before we proceed, we restate our definition of scarcity of reserves with a new one:

Definition 2. *Excess reserves are said to be scarce if*

$$-\bar{\nu} + \bar{\eta} \leq X/2 \leq \bar{\nu} - \bar{\eta} \tag{16}$$

where $X \equiv M - R_1 - R_1$.

Notice that definition 2 simply substitutes per bank excess reserves supplied by the central bank $X/2$ for a bank's excess reserve holdings x_i in definition 1 and that it implies definition 1 in equilibrium. This definition of scarcity is described solely by exogenous model parameters. Given (16) we can demonstrate that the period 1 equilibrium is unique and is described by

Lemma 3. *Period 1 equilibrium is characterized by*

$$u'(x_i) = v'(x_i) = v'(\frac{X}{2}) = r_D \tag{17}$$

for $i = 1, 2$.

[13] $Ew(x_i, \eta_i, r_R)$ depends on i due to the presence of η_i. It will not depend on η_i if r_R is symmetric around zero as a function of η_1, $r_R(\eta_1) = r_R(-\eta_1)$. In the appendix we show that $x_1 = x_2$ in equilibrium and thus $r_R(\eta_1) = r_R(-\eta_1)$.

[14] Generally r_R needs to be treated as a function of η_1. However in the baseline model it can be shown that r_R is constant in equilibrium.

[15] Strictly speaking $u(x_i)$ is only concave. However $u(x_i)$ is strictly concave near the equilibrium x_i.

Lemma 3 is quite intuitive. Banks are essentially identical at the beginning of period 1 each bank and will borrow one half of the aggregate excess reserves in equilibrium via deposits.[16] Since the function $v'(z)$ describes the value to a bank of borrowing an additional unit of reserves, the equilibrium period 1 deposit (borrowing) rate will be given by $v'(z)$ evaluated at the bank's expected excesss reserve holdings in equilibrium.[17]

Since $x_1 = x_2 = X/2$, (13) implies that the equilibrium period 2 interbank rate is given by

$$r_R = v'(\frac{X}{2}) + \frac{1}{2}c_B. \tag{18}$$

Comparing (17) with (18) it can be seen that there is a rather simple relationship between the period 1 deposit rate and the period 2 interbank rate. In particular

$$r_R = r_D + \frac{c_B}{2}.$$

Notice that the deposit rate is lower than the interbank rate because bank reserves that are lent or borrowed in the interbank market will incur a balance sheet cost. The existence of balance sheet costs means that the banks cannot "arbitrage away" the difference between the two rates. Finally the equilibrium deposit and interbank rates can be expressed in terms of model parameters by using equation (5). The equilibrium period 2 interbank rate is given by

$$r_R = \frac{r_W + r_E}{2} - s\frac{X}{2}, \tag{19}$$

and the equilibrium period 1 deposit rate is given by $r_D = r_R - c_B/2$.[18]

3.4 Summary of the benchmark model

Equilibrium in our model is described by two interest rates r_D and r_R and four borrowing levels x_1, x_2, y_1, and y_2. More specifically

- The period 1 deposit rate r_D equals $v'(X/2)$. Banks hold the same amount of excess reserves $x_1 = x_2$ and the deposit market clears $R_1 + x_1 + R_2 + x_2 = M$.

- The period 2 interbank rate r_R equals $v'(X/2) + c_B/2$. The interbank market clears $y_1(\eta_1) + y_2(\eta_1) = 0$ for all η_1 and depending on the size of the shock η the interbank

[16] Banks care about excess reserves. So even though it may be the case that $R_1 \neq R_2$ banks are *ex ante identical in how they treat excess reserves* x_i.

[17] This also requires that $v'(z)$ be linear over a relevant interval which was assumed earlier.

[18] Recall that $s = (r_W - r_E + c_B)/(2\bar{\nu})$.

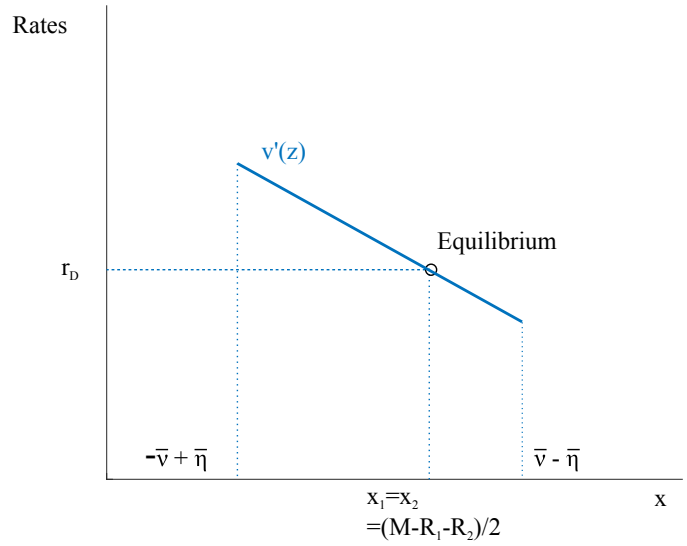

Figure 3: deposit market equilibrium when reserves are scarce.

market can be either active $y_1 < 0$ or inactive $y_1 = 0$. Intuitively, if shocks are small the interbank market will be inactive and if they are big it will be active.

The equilibrium for our economy can be neatly described in two diagrams: one for equilibrium in the period 1 deposit market and another for equilibrium in the period 2 interbank market. Equilibrium in the period 1 deposit market is illustrated in figure 3. The per bank aggregate excess reserves $X/2$ lie in between $[-\bar{\nu}+\bar{\eta}, \bar{\nu}-\bar{\eta}]$ and the function $v'(x)$ is linear and strictly decreasing over the interval. The function $v'(x)$ should be interpreted as a borrowing bank's demand for excess reserves for either period 1 or period 2. In equilibrium each bank borrows $x_1 = x_2 = X/2$ excess reserves from the investor and the market clearing deposit rate r_D is given by the intersection of $v'(x)$ and the perpendicular emanating from $x = X/2$.

Equilibrium for the period 2 interbank market is illustrated in figure 2. The two downward sloping lines $v'(z)$ and $v'(z)+c_B$ can be interpreted as the demand curves for excess reserves by a borrowing bank and a lending bank respectively.[19] Suppose that η_1 is sufficiently large so that the interbank market is active, $y_1 < 0$, as illustrated in figure 2. In particular the interest rate at which the potential lender is willing to lend when its excess reserves are $z = X/2 + \eta_1$, $v'(z) + c_B$—illustrated by small square on the upper demand curve—is strictly less than the rate at which the potential borrower is willing to borrow when its excess

[19]Recall that in equilibrium $v'(z) = r_R$ for a borrowing bank and $v'(z) + c_B = r_R$ for a lending bank if there is a nonzero trade.

reserves are $z = X/2 - \eta_1$, $v'(z)$—illustrated by the small square on the lower demand curve. This configuation indicates that there are gains from trade in the interbank market. The equilibrium trade volume is given by $-y_1 = \eta_0 - \eta_1$ in figure 2 at the interbank rate r_R that satisfies

$$r_R = v'(\frac{X}{2} + \eta_1 + y_1) + c_B = v'(\frac{X}{2} - \eta_1 - y_1) = v'(\frac{X}{2}) + \frac{1}{2}c_B. \tag{20}$$

As defined earlier for η_1 shocks that exceed $\eta_0 \equiv \eta_1 + y_1$ there is a positive trade. Figure 2 illustrates that if $\eta_1 > \eta_0$ then the interbank market is active and the interbank rate is given by equation (20). If instead $0 < \eta_1 < \eta_0$ the interbank market is inactive. Even in this case the interbank rate given by equation (20) is consistent with equilibrium but it is not the unique rate consistent with equilibrium.

4 Interbank market pre- and post-crisis

We now use the model to study trade volume in the interbank market and other money market outcomes under different scenarios. The first two cases correspond to the pre-crisis and the current interbank markets. The former is characterized by scarce excess reserves and negligible balance sheet costs while the the latter is characterized by abundant excess reserves and significant balance sheet costs.[20] We show that in each of these cases our model delivers predictions that are consistent with stylized money market facts.

A third case which we consider in the subsequent section is hypothetical but one that could be relevant in the future. This case is characterized by significant balance sheet costs and scarce excess reserves. This case is relevant if in the future the size of the Federal Reserve balance sheet is sufficiently reduced so that excess reserves become scarce. In this case we show that interbank market trade volume can be much smaller than in the pre-crisis case and may completely disappear in spite of excess reserves being scarce.

4.1 Pre-crisis period

In the pre-crisis period the federal funds market was primarily an interbank market and the eurodollar market was a place where non-banks lent to banks. In our model it is best to interpret the interbank rate r_R as the fed funds rate and the deposit rate r_D as the eurodollar rate.

[20] We formally define abundant excess reserves later.

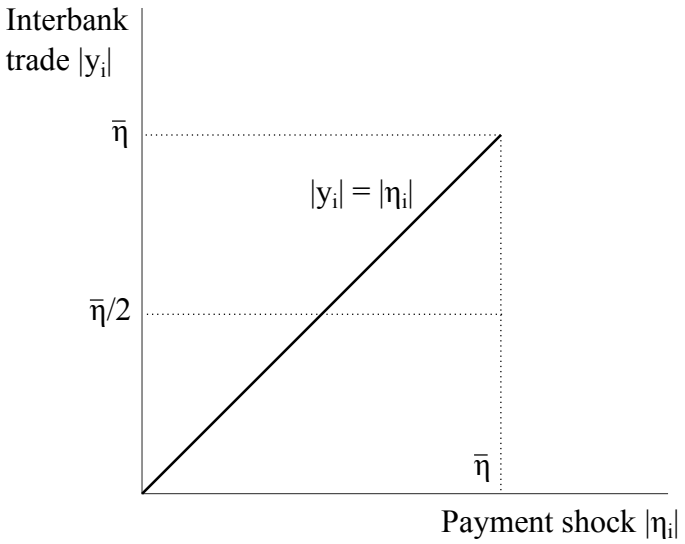

Figure 4: interbank trade volume.

In the pre-crisis period banks did not receive interest on reserves deposited at the Fed and their balance sheet costs were quite small.[21] For example the pre-crisis period's formula for the FDIC assessment fee depended primarily on deposit liabilities and not on the size of the bank's balance sheet. And while the U.S. did have leverage ratio requirements during this time the ratios were lower than current ratios and the base upon which the ratio was calculated was narrower than it is now. Moreover anecdotal evidence suggests that the pre-crisis behavior of banks was consistent with very low or no balance sheet costs. Finally and importantly, the hallmark of the pre-crisis period was scarcity of excess reserves. In terms of model parameters the pre-crisis period is best characterized by $-\bar{\nu} + \bar{\eta} \leq X/2 \leq \bar{\nu} - \bar{\eta}$, $r_E = 0$ and $c_B = 0$.

In the pre-crisis period the fed funds and the eurodollar rates were typically very close to one another. This is consistent with our model when $c_B = 0$. When excess reserves are scarce, we have $r_R = r_D + (1/2)c_B$ and with $c_B = 0$, $r_R = r_D$.

Pre-crisis interbank trading volumes in the federal funds market was quite high. Our model is consistent with a very active federal funds market when $c_B = 0$ since all realizations of the shock η_i give rise to a nontrivial amount of interbank trade. In particular $-y_i = \eta_i$. The average volume of trade in our model is $\bar{\eta}/2$ and actual volume is given by $|\eta_i|$; see figure

[21] The Fed only received authority to pay interest on reserves in 2008. Congress voted to give the Federal Reserve the authority to pay interest on reserves in 2006 but this authority was supposed to take effect five years later in 2011. The authority was accelerated during the financial crisis so as to give the Fed additional tools to maintain interest rate control while trying to stabilize financial markets.

4.

In the baseline model we assume that the investor incurs an extremely large cost when it delays its bank lending until period 2. But notice that for our pre-crisis model parameters even if the cost of delay were arbitrarily small, the investor would never have an incentive to delay lending until period 2 because $r_D = r_R$. In other words since the investor does not earn a higher rate by lending late in period 2 it will never pay any positive cost to do so. This observation is consistent with the relatively clear distinction between the timing of the federal funds market and the eurodollar market that existed in the pre-crisis period: The federal funds market was primarily interbank and trading happened mostly late in the day while eurodollar lenders were non-banks that traded mostly in the morning.

Finally we discuss pre-crisis monetary policy implementation through the lens of our model. When excess reserves are scarce the interbank rate in the model responds to small changes in reserves M since, by (19), we have $dr_R/dM < 0$. In practice changes in reserves were achieved through open market operations. In the pre-crisis period the target for the federal funds rate was set δ percentage points below the discount window rate r_W where δ was equal to 1 percentage point. We can use equation (19) to characterize the supply of reserves M or excess reserves X that is consistent with the target federal funds rate by assuming that $r_W = r_R + \delta$. In particular

$$X = 4\bar{\nu}\frac{\delta}{r_R + \delta} - 2\bar{\nu}.$$

Notice that aggregate excess reserves X are a decreasing function of r_R. Hence to implement a higher target interbank rate the central bank needs to reduce the amount of excess reserves in the banking system—in practice by either undertaking an open market sale or a reverse repo—and by increasing r_W. Finally by setting $r_W = r_R + \delta$ any interbank rate $r_R > 0$ can be implemented by choosing the appropriate amount of reserves.[22]

4.2 Post-crisis period

The Federal Reserve started paying interest on excess reserves in 2008 at the beginning of the financial crisis. Hence we have that $r_E > 0$ for the post-crisis period. The post-crisis period is also characterized by financial market regulations that impose significant costs on

[22]Implementing r_R close to 0 or r_W may require increasing or decreasing M beyond the range consistent with assumption 2. However since c_B is zero the relationship between M and r_R does not change as long as $-\bar{\nu} \leq X/2 \leq \bar{\nu}$. This will be seen more clearly from our discussion on abundant reserves.

banks that are related to the size of their balance sheets. Hence $c_B > 0$. Finally since 2009, the amount of excess reserves in the U.S. banking system has been very large or *abundant*. In terms of our model we define abundance as

Definition 3. *Excess reserves are said to be abundant if*

$$\frac{X}{2} \geq \bar{\nu} + \bar{\eta}.$$

If excess reserves are abundant in the sense of definition 3 and banks hold the same amount of excess reserves at the end of period 1, $x_1 = x_2 = X/2$, then each bank's period 3 reserve holdings will always exceed what is required R_i for any shocks η_i and ν_i when there is no interbank trade, $y_i = 0$. Abundant reserves along with balance sheet costs and interest on reserves have interesting implications for equilibrium in our model.

Although we shall see that $x_1 = x_2 = X/2$ is an equilibrium—as in the case with scarce reserves—this equilibrium allocation of date 1 excess reserves is not unique. In fact any period 1 bank excess reserve holdings x_1 and x_2 such that $x_1, x_2 \geq \bar{\nu} + \bar{\eta}$ and $x_1 + x_2 = X$ is an equilibrium. All of these equilibria share a common feature: Independent of the realizations of η_i and ν_i banks never need to borrow from the discount window in period 3. We now examine these equilibria in greater detail.

When excess reserves are abundant in the sense of definition 3 the bank's period 1 deposit-taking and period 2 interbank borrowing/lending problems are drastically simplified. Assume for time being that $x_1, x_2 \geq \bar{\nu} + \bar{\eta}$. An implication of this assumption is that each bank's beginning of period 3 excess reserve holdings is strictly positive when it does not borrow or lend in the period 2 interbank market. Therefore if a bank borrows a unit of reserves in the period 2 interbank market then that loan generates an additional balance sheet cost for the borrowing bank with probability one. Hence a bank will only borrow in the interbank market if $r_R \leq r_E - c_B$. However a bank that has with probability one strictly positive excess reserves in period 3 will never lend at a rate $r_R < r_E$ because its marginal value for reserves equals r_E, the rate that it earns from the Fed. Therefore when excess reserves are abundant and $x_1, x_2 \geq \bar{\nu} + \bar{\eta}$, the period 2 interbank market will be inactive, $y_i = 0$. In this situation any $r_E - c_B \leq r_R \leq r_E$ will be consistent with equilibrium in the period 2 interbank market. When $r_E - c_B \leq r_R \leq r_E$ banks do not have an incentive to either borrow or lend in the interbank market.

Also it can be shown that the equilibrium deposit rate r_D is equal to $r_E - c_B$.[23] This

[23] See appendix.

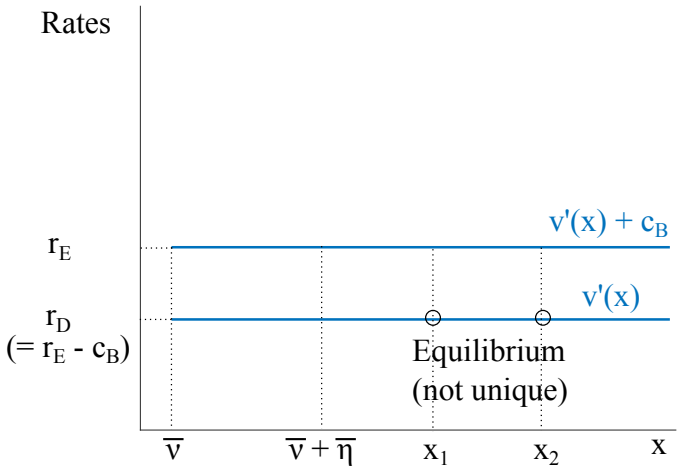

Figure 5: equilibrium with abundant reserves.

is intuitive because the banks always end up with positive excess reserves in equilibrium: Any additional unit of early deposit brings in r_E in interest on excess reserves but is subject to balance sheet cost c_B. To see that $x_1, x_2 \geq \bar{\nu} + \bar{\eta}$ holds in equilibrium, suppose instead that $x_2 < \bar{\nu} + \bar{\eta}$ implying $x_1 > \bar{\nu} + \bar{\eta}$. Then for some large negative η shocks, bank 2's marginal benefit of borrowing in period 2 will exceed $r_E - c_B$ because interbank trades do not fully offset η shocks due to balance sheet costs. Therefore bank 2's expected marginal benefit of borrowing in period 2 from a date 1 perspective also exceeds $r_E - c_B$. However bank 1 always ends up with positive excess reserves and thus its expected marginal benefit of borrowing is $r_E - c_B$. This cannot be an equilibrium because the marginal benefits are different. Therefore equilibrium requires that $x_i \geq \bar{\nu} + \bar{\eta}$ for banks $i = 1, 2$.

Figure 5 diagramatically describes the equilibrium outcome. When reserves are abundant, the relevant part of the borrower's and lender's marginal valuation of reserves are simply horizontal lines beyond $x_i \geq \bar{\nu} + \bar{\eta}$ at heights $r_E - c_B$ and r_E respectively. Figure 5 illustrates an equilibrium where $x_1 \neq x_2$. In contrast to the case where reserves are scarce, equilibrium with abundant reserves is not unique. Although every equilibrium is characterized by $r_D = r_E - c_B$ and $r_E - c_B \leq r_R \leq r_E$, any period 1 deposits $x_1, x_2 \geq \bar{\nu} + \bar{\eta}$ such that $x_1 + x_2 = X$ is an equilibrium.

In reality the volume of interbank trade in the U.S. is not zero but it is very close to it. Potter (2016) notes that during "the first half of 2016 less than 5 percent of fed funds transactions were interbank transactions based on the FR2420 data or about $3 billion per day on average." Our model also predicts that banks would never choose to lend at

a rate below the interest rate paid on excess reserves (IOER) r_E in the interbank market. Indeed with excess reserves that fall a bit short of abundance the interbank rate will be r_E.[24] Although there are some interbank trades with rates below IOER, interbank rates are generally above rates on fed funds lent by non-banks and closer to IOER.[25]

This can be seen on figure 2 of Potter (2017) where the upper tail of the distribution of rates—consisting of mostly interbank trades—in the fed funds market is above the IOER. In contrast fed funds trades between Federal Home Loan Banks and commercial banks and non-bank to bank trades in the eurodollar market occur at a rate below the IOER as in the model. Note that Federal Home Loan Banks do not earn IOER from the Federal Reserve so they can be regarded as non-banks in the money market. Therefore their fed funds lending is considered deposits in our model along with eurodollar trades.

As in the pre-crisis regime of section 4.1 the investor has no incentive to pay any cost for the option of lending to the bank in period 2. Investors understand that the period 2 borrowing rate will equal $r_E - c_B$ which is the rate they receive for lending in period 1.[26]

5 U. S. interbank market in the future

In this section we examine how the U. S. interbank market may evolve in response to a reduction in excess reserves. We examine how equilibrium outcomes will be affected when reserves declined from being abundant—with no need for interbank trading in period 2—to scarce. Generally speaking interbank trading volume will increase when excess reserves are lowered from an abundant level. Then depending on the specification of the investor's cost associated with delaying bank lending to period 2, interbank trade volume can stop increasing and remain at a relatively low level even with continued reserve drain or can disappear at some point, being completely replaced by the investor's period 2 loans. These outcomes suggest that it may be difficult or misleading to assess the implications associated with large scale draining of excess reserves by experimenting with small changes in reserves

[24]With abundant reserves there are no interbank transactions. However there is an intermediate region between reserves abundance and scarcity where deposit rate r_D stays below IOER while a small volume of interbank trades occur at IOER. This will become clearer in the next section when we discuss the future of the interbank market.

[25]Therefore lending banks sometimes make losses by lending. This may be understood as being part of various business relationships between the lending and the borrowing bank.

[26]Period 2 borrowing rate is indeterminate in equilibrium only because there is no trade in period 2. If any trade occurs the rate needs to be no higher than $r_E - c_B$ to induce borrowing.

when excess reserves are still large.[27]

The analysis here suggests that if excess reserves are lowered sufficiently, open market operations will be able to affect interest rates, such as the deposit rate, in a predictable way. However, the stance of monetary policy may not be able to be transmitted by a rate that relies on active interbank trading because interbank trades may disappear.

We first examine the situation where reserves have been drained from the banking system to the point where excess reserves X are scarce, as in definition 2. We now assume that the investor's of cost of delaying bank lending to period 2 is no longer prohibitive and consider two types of costs: a constant marginal cost per unit of period 2 loan and a fixed cost. The cost may represent the burden of maintaining operational readiness, creating new types of contracts or infrastructures to execute late loans or even the fear of not being able to invest resources profitability if one waits for too long.[28]

Let h represent the amount of funds that the investor lends to banks in period 2 and h_i the amount of reserves that bank i borrows from the investor in period 2. The total cost of lending h in period 2 is characterized by either $c_h h$, where c_h represents the constant marginal cost c_h associated with late loans, or a one-time fixed cost C_h. We will interpret the fixed cost as the cost for constructing a public good infrastructure that facilitates the late loans.

We first analyze a bank's problem taking the amount of resources that the investor lends in period 2, h, as given. Then we analyze the investor's choice h of period 2 lending under the two cost functions.

5.1 Period 3: the bank's payoff

The expression for excess reserves e_i is almost identical to that of the baseline model except now it includes period 2 loans h_i from the investor:

$$e_i \equiv x_i + y_i + h_i + \eta_i + \nu_i.$$

The period 3 payoff for bank i, net of the exogenous terms having R_i, η_i or ν_i, is given by

$$-x_i r_D - (y_i + h_i) r_R - [-y_i]^+ c_B + [e_i]^+ (r_E - c_B) - [-e_i]^+ r_W.$$

[27] By 'large' we mean that excess reserves are neither scarce nor abundant. We formalize this notion below.

[28] In our model the investor is always has the opportunity to lend so this last interpretation comes from outside the model.

Notice that the period 2 interbank market rate r_R equals the rate at which banks borrow from the investor in period 2. This is an equilibrium outcome. To see this suppose that the rates are different. Then borrowers would avoid the higher borrowing rate and the type of trades associated with that rate would be characterized by an excess supply of funds.

5.2 Period 2: banks' borrowing decisions

Equilibrium in period 2 is characterized by a borrowing rate r_R that clears both the interbank market $-y_1 = y_2$ and the market where banks borrow from the investor $h = h_1 + h_2$.

The bank's period 2 problem is only slightly different from the baseline model and reflects the bank's ability to borrow directly from the investor. The bank's decision problem is given by

$$\max_{y_i, h_i} -(y_i + h_i)r_R - [-y_i]^+ c_B + v(x_i + \eta_i + y_i + h_i)$$

where

$$v(x_i+\eta_i+y_i+h_i) = \int_{-\bar{\nu}}^{\bar{\nu}} \{[x_i+\eta_i+y_i+h_i+\nu_i]^+(r_E-c_B)-[-(x_i+\eta_i+y_i+h_i+\nu_i)]^+ r_W\}(2\bar{\nu})^{-1} d\nu. \tag{21}$$

As in the baseline model we identify bank 1 as the potential lender which means that $x_1+\eta_1 \geq x_2+\eta_2$ and $\eta_1 = -\eta_2 \geq 0$—anticipating $x_1 = x_2$ in equilibrium. Bank 1 never lends in the interbank market at the same time it borrows from the investor. To see this take the first order conditions of (21) with respect to h_i and y_1 to get $r_R = v'(z_1)$ and $r_R = v'(z_1)+c_B$ respectively. Clearly it is not possible for both equations to simultaneously hold. Hence there are 3 general cases to consider:

1. Both banks borrow from the investor;
2. Bank 2 borrows from the investor and bank 1 neither borrows nor lends;
3. Bank 2 borrows from both the investor and bank 1.

Case 1 is characterized by $v'(x_2 - \eta_1 + h) < v'(x_1 + \eta_1)$. If bank 2 receives all of the investor's loans h then its marginal value of borrowing reserves is less than bank 1's marginal value. This implies that there exists a $h_2 > 0$ such that

$$v'(x_2 - \eta_1 + h_2) = v'(x_1 + \eta_1 + h - h_2).$$

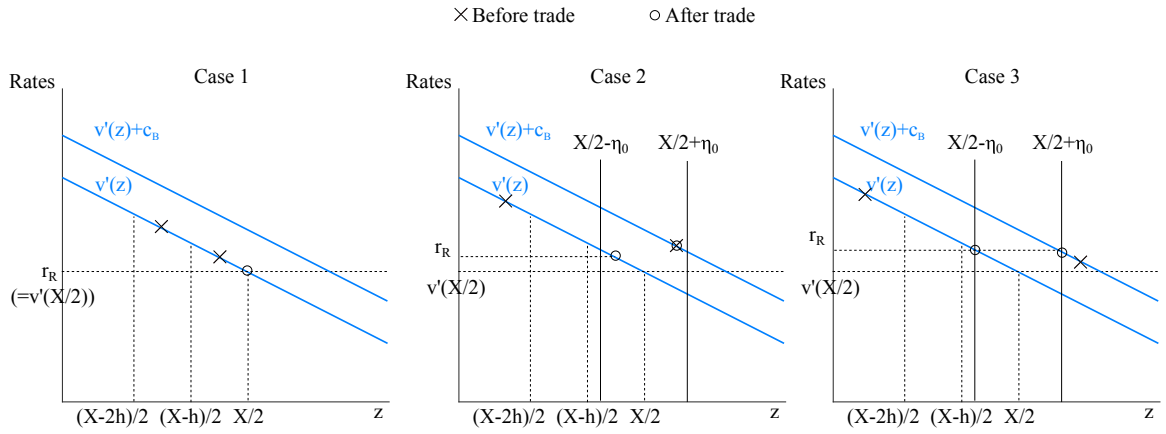

Figure 6: late market with non-bank lending.

In this situation the investor extends loans to both banks. When the period 2 borrowing market closes, both banks will hold the same amount of excess reserves which in equilibrium equals $X/2$. Therefore the period 2 borrowing rate r_R must satisfy

$$r_R = v'(\frac{X}{2}).$$

Case 1 arises when the η shock is relatively small. Anticipating that in equilibrium we have $x_1 = x_2$, case 1 occurs whenever $\eta_1 \in [0, h/2]$. If the investor lends h in total to the banks then each bank will hold excess reserves $(X - h)/2$ in period 1; see the left panel in figure 6. Note that period 1 bank excess reserve holdings exclude h reserves (which may sit at either bank before getting lent out in period 2). For any shock $\eta_1 \in [0, h/2]$ bank 1's excess reserves will be less than or equal to $X/2$ which means that the investor can lend to both banks and both banks' post trade excess reserves can equal $X/2$. It is clear from figure 6 that if $\eta_1 = h/2$ bank 1 will have exactly $X/2$ excess reserves before the date 2 loan market opens and bank 2 will have $X/2$ excess reserves after it borrows h from the investor. For an arbitrary $\eta_1 \in (0, h/2)$, illustrated by the 'x's' on the bank's borrowing demand curve in figure 6, both banks borrow from the investor and bank 2 borrows $2\eta_1$ more reserves than bank 1. Both banks end up holding $X/2$ reserves and the fed funds rate equals $v'(X/2)$, as indicated by the circle on the lower demand curve in figure 6.

In case 2 bank 2 borrows h from the investor while bank 1 neither borrows nor lends.

Case 2 therefore must be characterized by $v'(x_2 + \eta_2 + h) \geq v'(x_1 + \eta_1)$ and

$$v'(x_2 + \eta_2 + h) \leq v'(x_1 + \eta_1) + c_B.$$

The first inequality says that bank 2's marginal valuation of borrowing reserves exceeds that of bank 1's (marginal valuation of borrowing) when it receives all of the depositor's loans h. This implies that bank 1 does not borrow from the investor and becomes a potential lender. The second inequality says that bank 1 does not lend in the interbank market because any gains from any trade are more than offset by the balance sheet costs. Therefore we have that $h_2 = h$ and $h_1 = y_1 = y_2 = 0$. From bank 2's decision problem the equilibrium period 2 borrowing rate must satisfy

$$r_R = v'(x_2 + \eta_2 + h).$$

Once again, anticipating that $x_1 = x_2 = (X - h)/2$ we can identify the set of η's that are relevant for case 2. Notice that when $\eta_1 = h/2$ the equilibrium allocation of excess reserves after period 2 trading with the investor is identical to that of the baseline model when $\eta_1 = 0$. Therefore the set of η's that are consistent with case 2 simply shifts the set of η's that are consistent with an equilibrium in the baseline model that has $y_1 = 0$ by $h/2$:

$$\eta_1 = -\eta_2 \in [\frac{h}{2}, \frac{h}{2} + \eta_0]$$

with $\eta_0 \equiv c_B/(2s)$. (Recall that $s = (r_W + c_B - r_E)/2\bar{\nu}$ is the absolute value of the slope of $v'(z)$ for $z \in [-\bar{\nu}, \bar{\nu}]$). The middle panel of figure 6 illustrates this set. Notice that by construction the marginal valuation of borrowing for bank 2 for $z = X/2 - \eta_0$ and of lending for bank 1 for $z = X/2 + \eta_0$ are equal and given by

$$v'(\frac{X}{2} - \eta_0) = v'(\frac{X}{2} + \eta_0) + c_B.$$

If we assume that $x_1 = x_2 = (X - h)/2$ then the excess reserve holdings for case 2, after the investor lends h to bank 2, lies somewhere in the interval $[X/2, X/2 + \eta_0]$ for bank 1 and $[X/2 - \eta_0, X/2]$ for bank 2 depending upon the magnitude of the η shock. Notice that the bank 2's marginal valuation of borrowing after receiving h from the investor, indicated by the circle on the borrower's demand curve in the middle panel of figure 6, is less than bank 1's marginal valuation of lending, indicated by the 'x' and the circle on the lender's demand curve. The borrowing rate r_R is given by bank 2's marginal valuation of borrowing after it borrows h from the investor and is illustrated by the circle on the borrower's demand

curve. The borrowing rate r_R is characterized by $r_D^* \leq r_R \leq r_R^*$, where $r_D^* = v'(X/2)$ and $r_R^* = r_D^* + c_B/2$. r_R^* will be defined as the period 2 rate in case 3 and r_D^* as the deposit rate in the bank's period 1 problem.

Finally in case 3 bank 2 borrows from both the investor and bank 1. This case is characterized by

$$v'(x_2 + \eta_2 + h) > v'(x_1 + \eta_1) + c_B,$$

which means that even after bank 2 borrows h from the investor, there are gains from trade in the interbank market. The equilibrium interbank rate is

$$r_R = v'(x_2 + \eta_2 + h - y_1) = v'(x_1 + \eta_1 + y_1) + c_B. \tag{22}$$

Anticipating that $x_1 = x_2 = (X - h)/2$, case 3 arises when

$$\eta_1 > \frac{h}{2} + \eta_0.$$

In case 3 the allocation of reserves held by banks 1 and 2 after period 2 trading will be identical to the the excess reserves held by these banks when η_1 precisely equals $h/2 + \eta_0$. In particular when $\eta_1 > h/2 + \eta_0$, bank 2 borrows $h_2 = h$ from investor and

$$y_2 = -y_1 = \eta_1 - \frac{h}{2} - \eta_0$$

from bank 1. We can therefore rewrite (22) as

$$r_R = v'(x_2 + \frac{h}{2} - \eta_0) = v'(x_1 + \frac{h}{2} + \eta_0) + c_B. \tag{23}$$

Since ν is uniformly distributed—which implies that v' is linear—the equations in (23) can be simplified to read

$$r_R = \frac{1}{2}[(v'(x_2 + \frac{h}{2}) + v'(x_1 + \frac{h}{2})] + \frac{1}{2}c_B.$$

Anticipating that $x_1 = x_2 = (X - h)/2$ this equation can be further simplied to

$$r_R^* = v'(\frac{X}{2}) + \frac{1}{2}c_B.$$

The last panel in figure 6 illustrates this case. The 'large' η shock results in excess reserve holdings indicated by the 'x' on bank 1's lending demand curve and bank 2's borrowing demand curve. After bank 2 borrows h from the investor, its excess reserves will still be less than $X/2 - \eta_0$. Therefore bank 2's marginal value for borrowing reserves is greater than bank

1's marginal value for lending reserves; hence there are gains from trade. Bank 1 lends to bank 2 until the marginal values are equated, with bank 1 holding $X/2 + \eta_0$ excess reserves and bank 2 holding $X/2 - \eta_0$ as indicated by the circles on the borrower's and lender's demand curves. Finally, notice that when η_1 is large in the sense of case 3, the period 2 borrowing rate equals the interbank rate in the baseline model when there is strictly positive volume.

5.3 Period 1: banks' demand for deposits

5.3.1 Bank i's demand for deposits

As in the baseline model, the bank's choice of deposits is given by

$$\max_{x_i} u(x_i) - x_i r_D \tag{24}$$

where $u(x_i) = Ew(x_i, \eta_i, r_R)$ and

$$w(x_i, \eta_i, r_R) = \max_{y_i, h_i} [-(y_i + h_i) r_R - [-y_i]^+ c_B + v(x_i + \eta_i + y_i + h_i)].$$

As in the baseline model it can be shown that $u(x_i)$ is concave and that $w'(x_i, \eta_i, r_R) \equiv \partial w / \partial x_i(x_i, \eta_i, r_R) = v'(z_i)$ where $z_i \equiv x_i + \eta_i + y_i + h_i$ for optimal y_i and h_i. The former implies that the solution to (24) which is $u'(x_i) = r_D$ is unique under broad assumptions. As a result $x_1 = x_2$ as anticipated.[29]

To characterize $u'(x_i) = r_D$ we can evaluate $w'(x_i, \eta_i, r_R)$ for each of the three cases examined in section 5.2 and then compute its expectation over η_1. However we can simplify the problem further by taking advantage of the fact that η_i has a symmetric distribution around zero.

Since η_i is symmetric around zero, η_i can be $\hat{\eta}_1$ or $-\hat{\eta}_1$ with equal probability, for any $0 \leq \hat{\eta}_1 \leq \bar{\eta}$. Since the two banks are identical prior to period 2, $x_1 = x_2$, it is as if bank i with $\eta_i = \hat{\eta}_1$ were trading against bank i with $\eta_i = -\hat{\eta}_1$ in period 2. Writing the optimal choices of y_i and h_i as functions of η_i implies in equilibrium that

$$(x_i + \hat{\eta}_1 + y_i(\hat{\eta}_1) + h_i(\hat{\eta}_1)) + (x_i - \hat{\eta}_1 + y_i(-\hat{\eta}_1) + h_i(-\hat{\eta}_1)) = X.$$

This equation simply means that the excess reserves are distributed between the two banks. Since v' is linear, we have

$$E_{\eta_i \in \{\hat{\eta}_1, -\hat{\eta}_1\}} w'(x_i, \eta_i, r_R) = E_{\eta_i \in \{\hat{\eta}_1, -\hat{\eta}_1\}} v'(z_i) = v'(\frac{X}{2}).$$

[29] As in the baseline model the appendix shows that the dependence of $Ew(x_i, \eta_i, r_R)$ on the label i can be ignored in equilibrium.

The expression for the conditional expectation holds for any $\hat{\eta}_1$ and thus in equilibrium we have

$$r_D^* = E[w'(x_i, \eta_i, r_R)] = v'(\frac{X}{2}).$$

Notice the equilibrium period 1 deposit rate is identical to that of the baseline model which has $h \equiv 0$. We have characterized equilibrium behavior for the banks for a given level of period 2 investor loans h. We now examine the period 1 equilibrium behavior of the investor.

5.4 Investor's supply of deposits

It is costly for the investor to withhold funds to lend to banks in period 2. We consider two types of costs: a per unit constant marginal cost and a fixed cost. We examine the case of a constant marginal cost first.

5.4.1 Per unit withholding costs

In period 1 the investor chooses the amount of resources h to lend to banks in period 2. The period 1 problem that the lender solves is simple:

$$\max_h (M - h)r_D + h[E(r_R) - c_h] \tag{25}$$

where c_h is the marginal cost associated with lending in period 2. If $c_h = 0$ then h must be consistent with the first order condition $r_D = E(r_R)$. Since $r_R \geq r_D$ for any η, $r_D = E(r_R)$ implies that $r_D = r_R$.[30] Hence we have

$$r_R = r_D^* = v'(\frac{X}{2}).$$

In this situation the investor is indifferent between lending in period 1 and period 2. In order to ensure that the investor has sufficient period 2 resources to be consistent with $r_R = r_D^*$ for all possible η shocks, it must be the case that $h \geq 2\bar{\eta}$.[31]

Suppose now that $c_h > 0$. If $h > 0$ is optimal then from (25), $E(r_R) = r_D + c_h$. Since in any equilibrium the maximum ex post value of r_R is $r_D^* + c_B/2$, a necessary but not sufficient condition for $h > 0$ is that $c_h < c_B/2$. Also in any equilibrium where $c_h > 0$ it must be that $h < 2\bar{\eta}$. If this was not the case then $r_R = r_D^*$ and the investor could increase its payoff by setting $h = 0$.

[30] $r_R \geq r_D$ because r_R equals the marginal value of the bank with fewer excess reserves while r_D equals the marginal value of the bank with half of total excess reserves.

[31] Otherwise for $\eta_1 > h/2$ bank 2 will end up with fewer than $X/2$ reserves even after trading in period 2 and r_R will be greater than r_D^*.

When $h < 2\bar{\eta}$ there is a strictly positive probability that cases 1 and 2 (from section 5.2) will prevail in equilibrium. This implies that $E(r_R) > r_D^*$. Whether or not case 3 prevails with a strictly positive probability depends on the size of h. In particular if

$$h < 2(\bar{\eta} - \eta_0)$$

then case 3 will prevail with a strictly positive probability. If $h > 2(\bar{\eta} - \eta_0)$ then date 2 lending by the investor will be sufficiently large so that there will be no gains from trade in the interbank market. Recall that in case 3 we have that $r_R = r_D^* + c_B/2$.

For an arbitrary $h < 2\bar{\eta}$ the expected return to lending $h > 0$ in period 2, $E(r_R|h)$, is given by

$$E(r_R|h) = \frac{1}{\bar{\eta}}\{\frac{h}{2}r_D^* + \int_{h/2}^{\min\{\bar{\eta}, h/2+\eta_0\}} v'(\frac{X+h}{2} - \eta)d\eta + [\bar{\eta} - h/2 - \eta_0]^+(r_D^* + \frac{1}{2}c_B)\}.$$

Note that $E(r_R)$ can be treated as a function of h because once h is chosen, the equilibrium conditions for period 2 determine $E(r_R|h)$. When the investor deposits all of his resources at the banks in period 1, the expected return on a marginal period 2 loan is $E(r_R|h = 0)$. Hence if we define $c_B^* \equiv E(r_R|h = 0) - r_D$ then the necessary and sufficient condition for $h > 0$ is $c_h < c_B^*$. When this condition is met, the equilibrium level of h is simply determined by $E(r_R|h) - r_D^* = c_h$.

If $0 < c_h < c_B^*$ then the expected trading volume in the interbank market will increase when excess reserves are reduced from an abundant level to a scarce level. Specifically the expected interbank trading volume is zero when excess reserves are abundant and are positive when excess reserves are scarce as long as c_h is not too small; if it is too small the equilibrium value of h can be large enough to prevent any interbank trading. However as indicated in figure 7 the expected trading volume will decline compared to the pre-crisis period where $h \equiv 0$ for two reasons. First expected trading volume will fall because c_B has increased due to recent regulations. An increase in c_B (from zero) will decrease trading volumes for any given η shock. This decline in volume is illustrated in figure 7 as the downward shift in interbank trade curve from $(c_B = 0, h = 0)$ to curve $(c_B > 0, h = 0)$. Second an increase in the supply of loans h in period 2 by the investor will further reduce interbank trade volume for any given c_B and η. This decline in volume is illustrated in figure 7 as the downward shift in interbank trade curve from $(c_B > 0, h = 0)$ to curve $(c_B > 0, h > 0)$.

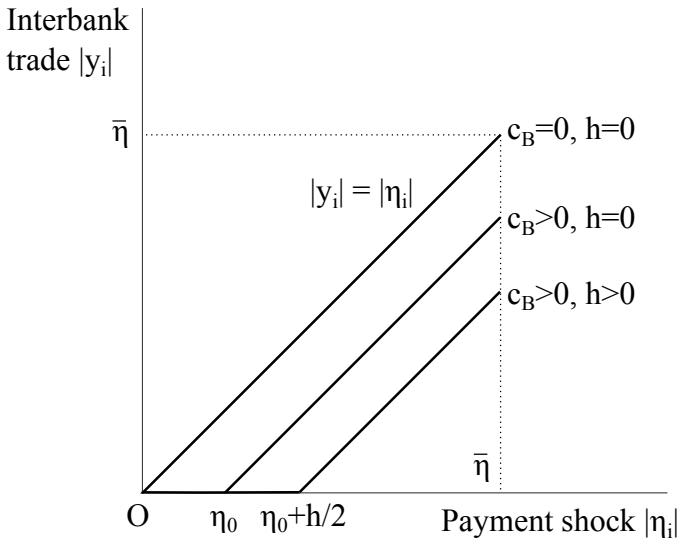

Figure 7: interbank volume under different conditions.

5.4.2 Fixed cost

Suppose a public good-type infrastructure is needed to facilitate date 2 investor loans to banks. One can think of the public good as a trading platform or a set of legal contracts that allow the investor to lend freely in period 2. Once the public good is constructed at cost C_h it is costless to use.

If the public good is constructed then the investor's outcome is identical to the case where the marginal cost of a period 2 loan is zero. In that situation the allocation of excess reserves at the end of period 2 is equal to $z_1 = z_2 = X/2$ and the deposit and period 2 borrowing rates are equal, $r_D^* = r_R = v'(X/2)$. The investor holds back sufficient resources equal to at least $2\bar{\eta}$ to ensure that banks' holdings of excess reserves can be equalized for any η shock. In this situation the interbank market shuts down.

If the public good is not constructed then the outcome is identical to the baseline model where $h \equiv 0$. In particular in period 2 if $\eta_1 \in (0, \eta_0)$ then $y_1 = y_2 = 0$ and no additional balance sheet costs due to interbank trades are incurred; as before $\eta_0 \equiv c_B/(2s)$ is the minimal shock necessary for an interbank trade to happen. If $\eta_1 \in (\eta_0, \bar{\eta})$ then $-y_1 = y_2 > 0$, additional balance sheet costs equal to $-y_1 c_b$ will be incurred by the banking sector and end of period 2 excess reserve holdings will be $z_1 = X/2 + \eta_0$ and $z_2 = X/2 - \eta_0$.

Collectively the banks and the investor have an incentive to incur the cost of providing the public good if the surplus generated by late investor loans exceeds its cost. For simplicity we assume that the public good is constructed if the surplus exceeds the cost. Conceptually

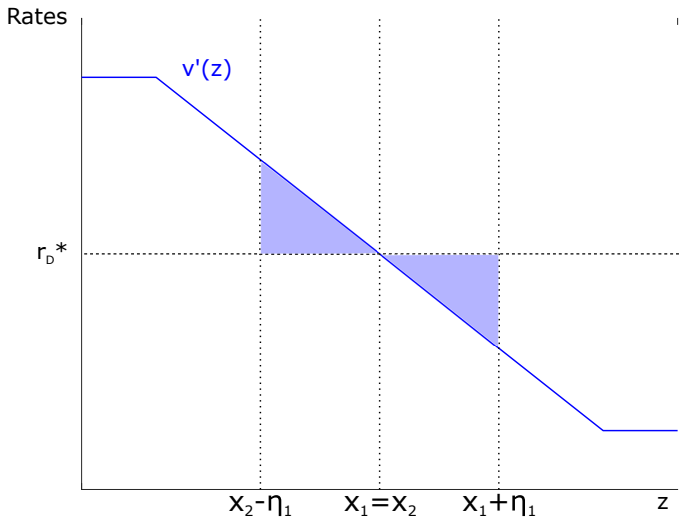

Figure 8: illustration of surplus.

the surplus is the sum of the benefits associated with costless period 2 trades compared to the baseline model. There are two components to this surplus. The first is the balance sheet cost savings from replacing interbank trades with non-bank lending. The second is the additional trades that take place via period 2 loans that equalize the two banks' marginal values for reserves.

The ex post surplus is a function of η_1. If $\eta_1 \in (0, \eta_0)$ then the interbank market is inactive in the baseline model and the surplus comes from equalizing the marginal values for reserves between the two *borrowing* banks. The surplus generated by date 2 lending for this case is illustrated in figure 8: It is the sum of the areas of the two triangles. More formally the ex post surplus generated by letting the investor lend to banks in period 2 without any extra cost, $S(\eta_1)$, is

$$\begin{aligned} S(\eta_1) &= \int_0^{\eta_1} [[v'(\frac{X}{2} - \eta) - r_D^*] + [r_D^* - v'(\frac{X}{2} + \eta)]] d\eta \\ &= \int_0^{\eta_1} [v'(\frac{X}{2} - \eta) - v'(\frac{X}{2} + \eta)] d\eta \\ &= \int_0^{\eta_1} 2s\eta d\eta \\ &= s\eta_1^2. \end{aligned}$$

If instead $\eta_1 \in (\eta_0, \bar{\eta})$ then the surplus is augmented by the elimination of balance sheet costs that would have been generated by interbank trading in the baseline model. More formally the surplus $S(\eta_1)$ is

$$S(\eta_1) = S(\eta_0) + c_B(\eta_1 - \eta_0).$$

The first term reflects the increase in efficiency associated with period 2 loans equating banks' marginal excess reserve valuations and the second term reflects the reduction in balance sheet costs associated with interbank trading. Suppose that the Fed reduces excess reserves from an abundent level to a scarce level. The expected total surplus $E[S]$ associated with *any* level of excess reserves that are scarce is given by

$$E[S] = \frac{1}{\bar{\eta}} \int_0^{\bar{\eta}} S(\eta) d\eta.$$

The public good infrastructure will be undertaken if $E[S] > C_h$. An implication of $E[S] > C_h$ is that the interbank market is inactive, i.e., it effectively shuts down.

5.5 Future path of money markets

To explore how money markets may evolve in response to possible draining of excess reserves, we perform a simple exercise. We extend the equilibrium analysis to all possible values of excess reserves X. In particular, we imagine that the Fed reduces excess reserves from abundant to scarce. We then document the volumes of interbank trade, period 2 non-bank lending and periods 1 and 2 lending rates at each level of excess reserves that is between abundant to scarce. We do this for three scenarios regarding the cost of period 2 lending by the investor. In the first scenario we assume that the cost of period 2 lending is prohibitive; in the second we assume that there is only a finite, non-prohibitive marginal cost to period 2 lending; and in the third we assume that there is zero marginal cost but a non-prohibitive fixed cost to period 2 lending. Here we provide the basic intuition behind the results and relegate the formal analysis to the appendix.

We first consider the baseline case where the cost associated with investor period 2 loans is prohibitive (meaning that the investor never lends to banks in period 2). Intuitively there are four qualitatively different regions of excess reserves per bank as in illustraged in figure 9:

1. $X/2 \in [\bar{\nu} + \bar{\eta}, \infty)$: Excess reserves are abundant. In equilibrium $r_D = r_E - c_B$ and interbank trade volume is zero.

2. $X/2 \in [\bar{\nu} + \bar{\eta} - c_B/s, \bar{\nu} + \bar{\eta})$: If η_1 is sufficiently large, then $X/2 - \eta_1 < \bar{\nu} + \bar{\eta}$ and the marginal value of reserves for bank 2 exceeds $r_E - c_B$. However by construction the marginal value is always less than r_E and therefore there is no interbank trade. In

particular, highest possible marginal valuation of excess reserves for a borrower in the interval is always less than $v'(\bar{\nu} - c_B/s) = r_E$. Since the marginal value of borrowing in period 2 exceeds $r_E - c_B$ for some η shocks, the date 1 deposit rate r_D—which is the expected marginal value of borrowing in period 2—exceeds $r_E - c_B$ and is decreasing in aggregate excess reserves.

3. $X/2 \in [\bar{\nu} - c_B/(2s), \bar{\nu} + \bar{\eta} - c_B/s)$: Since $v'(\bar{\nu} - c_B/s) = r_E$, there exists $\eta_1 < \bar{\eta}$ such that $v'(X/2 - \eta_1) > r_E$ which implies that expected interbank trade volume will be positive. The interbank rate will never exceed r_E since, by construction, $v'(\bar{\nu} - c_B/(2s)) + c_B/2 = r_E$. Hence the interbank rate is equal to r_E for all shocks η_1 that generate interbank trading volume. The date 1 deposit rate, r_D, is decreasing in aggregate excess reserves.

4. $X/2 \in [0, \bar{\nu} - c_B/(2s))$: Qualitatively speaking, rates and volumes in this region correspond to the baseline model when excess reserves are assumed to be scarce. In particular, $r_D = v'(X/2)$ and $r_R = v'(X/2) + c_B$.

The equilibrium in case 4 is characterized by the same equations as the baseline equilibrium in section 3. Assuming that $\bar{\eta} > c_B/(2s)$, the constraints on excess reserves in case 4 is less restrictive than our scarcity assumption under definition 2, $-\bar{\nu} + \bar{\eta} \leq X/2 \leq \bar{\nu} - \bar{\eta}$. Hence, we can relax the definition of scarcity to be $-\bar{\nu} + c_B/(2s) \leq X/2 \leq \bar{\nu} - c_B/(2s)$. Note that we have implicitly assumed that $\bar{\eta} \geq c_B/(2s)$ throughout the paper; otherwise the volume of interbank trade would be zero for any η realization since the marginal balance sheet cost c_B were 'too high.'

Next we consider the case where there is a constant, non-prohibitive marginal cost associated with period 2 investor lending. It is useful to generalize our definition of c_b^* from section 5.4.1. Specifically, define $\bar{c}_h(X) \equiv E(r_R | h = 0, X) - r_D$. Notice that $\bar{c}_h(X) = c_b^*$ for $X/2 \leq \bar{\nu} - c_B/(2s)$, i.e., when excess reserves are scarce. If excess reserves are abundant then equilibrium borrowing rate is equal to $r_E - c_B$ for both period 1 and period 2 lending for any given level of h. As a result the investor is not willing to lend in period 2 if the marginal cost is positve. Hence $\bar{c}_h(X) = 0$ for $X/2 \geq \bar{\nu} + \bar{\eta}$.

When $\bar{\nu} - c_B/(2s) \leq X/2 \leq \bar{\nu} + \bar{\eta}$, $\bar{c}_h(X)$ is a decreasing function of X. Intuitively, for any given η_1 the difference $r_R(\eta_1) - r_D$ decreases as X increases. When excess reserves are between scarce and abundant, the the investor will lend in period 2 only if marginal costs are characterized by $c_h(X) < c_h$. Figure 10 illustrates an example where period 2 lending by

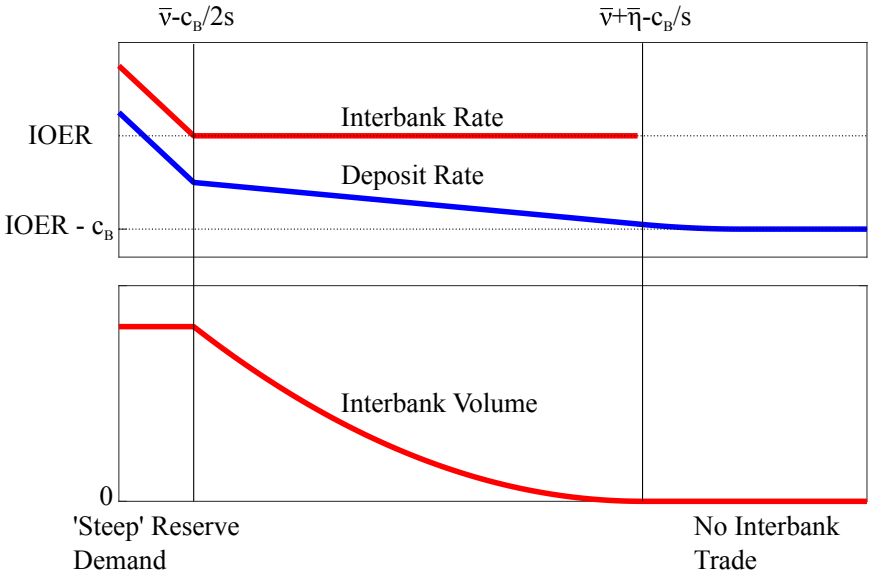

Figure 9: future path with no late non-bank lending

the investor occurs for $X \leq \hat{X}$, where $\bar{c}_h(\hat{X}) = c_h$ and $\bar{\nu} - c_B/(2s) \leq \hat{X}/2 < \bar{\nu} + \bar{\eta} - c_B/s$.

When $X > \hat{X}$, $h = 0$ and the equilibrium is described by cases 1, 2 and 3 of the previous example. When $\bar{\nu} - c_B/(2s) < X/2 < \hat{X}/2$ two interesting results stand out. First interbank volume remains constant as excess reserves are reduced. When excess reserves decrease banks are more willing to borrow in period 2 but now the higher demand for borrowing is met by increased period 2 lending by the investor rather than by increased interbank trading. This is in contrast to the above baseline example where interbank trading volume increases as excess reserves decrease in this region. Second for any given $X < \hat{X}$ the period 2 borrowing rate is not constant at r_E as in the above baseline example but varies with the magnitude of the η payment shock. In particular r_R will be higher when there is a larger demand for period 2 borrowing and this occurs with larger $|\eta_i|$'s.

Notice in figure 10 that interbank trading volume is dramatically reduced compared to the baseline model (where period 2 lending by the investor is zero). Clearly the magnitude of the reduction in interbank trading volume, as well as the size of period 2 lending by the investor, depends critically on the value of c_h.

Finally consider the case where the investor can lend in period 2 if a fixed cost is undertaken, e. g., for a public good infrastructure required to facilitate period 2 loans. The fixed cost will be undertaken if the total expected surplus to the private players, the banks and the investor, associated with the period 2 loans $E(S)$ exceeds the cost C_h of infrastructure.

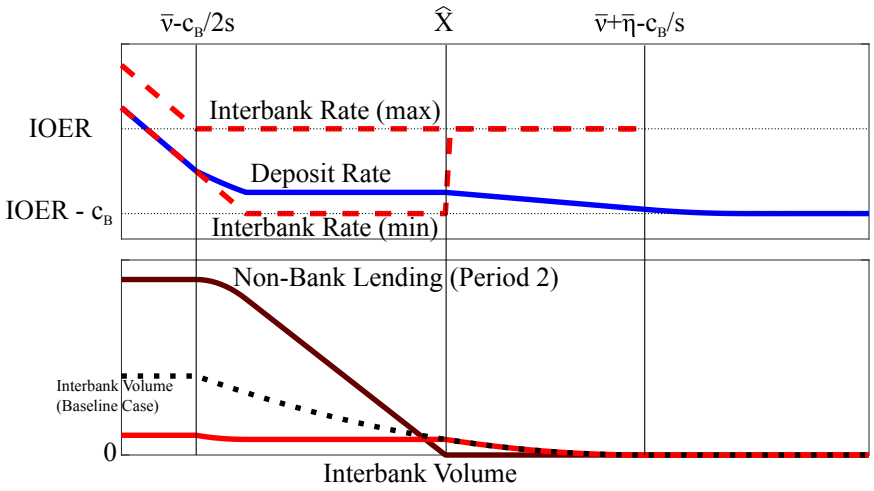

Figure 10: future path with a constant marginal cost.

We define a critical value of fixed cost C_h^* where

$$C_h^* = E(S|0 < X/2 < \bar{\nu} - c_B/(2s)).$$

The fixed cost C_h^* is such that the expected surplus of period 2 loans is zero when excess reserves are scarce.[32] If $C_h > C_h^*$ then the infrastructure will never be undertaken as the cost will always be larger than the surplus.

Now suppose that $0 < C_h < C_h^*$. Then there exists a critical level of aggregate excess reserves, $\bar{\nu} - c_B/(2s) < X^*/2 < \bar{\nu} + \bar{\eta}$, such that $C_h = E(S|X = X^*)$. Therefore, for all $X > X^*$ the fixed cost of public good infrastructure is not undertaken and the equilibrium outcomes are identical to the baseline case. If however $X \leq X^*$, then the fixed cost is incurred and the public good infrastructure is constructed. In this situation the interbank trading volume is always zero and h is always large enough to make $r_D = r_R = v'(X/2)$, for example $h = 2\bar{\eta}$; see figure 11. In contrast to the case where period 2 lending entails a constant marginal cost, the interbank market completely shuts down.

The foregoing discussion is based on the idea that the expected surplus is decreasing in the level of excess reserves and becomes zero when excess reserves are abundant. The intuition behind this relationship is clear: Banks have less need to borrow in period 2 when they hold larger excess reserves and, as a result, get less surplus from the ability to borrow from the investor. The negative equilibrium relationship between expected surplus and excess reserves is demonstrated in the appendix.

[32]The surplus is identical for all values of excess reserves that are scarce. Note that we are now using the less restrictive definition of scarce excess reserves.

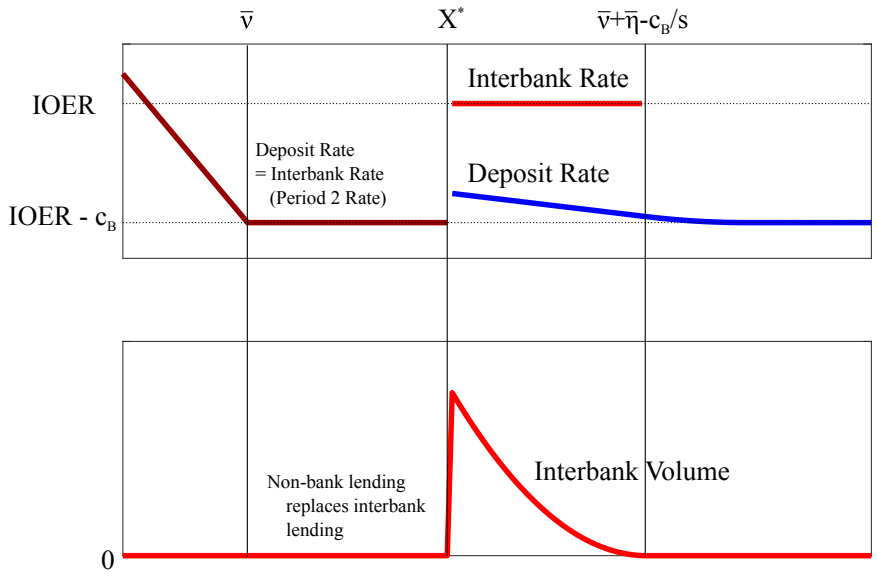

Figure 11: future path with a fixed cost.

The price and quantity responses associated with a reserve draining depend on the nature of costs faced by non-banks that provide loans to banks which substitute for interbank borrowing.[33] Figures 9, 10, and 11 describe outcomes when these costs are prohibitive, marginal and fixed, respectively. The response of interbank trading volumes and prices to decreases in excess reserves depend critically on the magnitude and the nature of costs that (institutional) lenders to banks face.

6 Conclusion

We study how interbank trading volumes and rates are affected by the size of aggregate excess reserves supplied to the banking sector and by recent regulations. Our analysis indicates that the new regulations increase the cost of interbank trading which, holding all else constant, results in decreased trading volume. These regulations also provide an incentive for non-banks to compete with the interbank market by lending reserves directly to banks even though there may be direct costs associated with competition. As a result interbank trading volumes may not return to pre-crisis levels even if the Fed returns to a monetary implementation framework that is based on scarce excess reserves. This could pose a potential problem since that framework partly relies on the existence of an active interbank market.

[33] In practice these non-banks can be institutions such as Federal Home Loan Banks, money market funds and corporate treasuries.

7 Appendix

This appendix formalizes some of the ideas and arguments presented in the main text.

7.1 Formal definition of the equilibrium

Here we define the equilibrium of our model. We define all choice variables and prices as functions of η_1 over the domain $\eta_1 \in [-\bar{\eta}, \bar{\eta}]$.

The equilibrium is defined by a set of prices r_D and $r_R(\eta_1)$, investor's choice h, and the two banks' choices x_i, $y_i(\eta_1)$, and $h_i(\eta_1)$, $i = 1, 2$ that clear market when the investor and the two banks maximize profits. In particular, h solves

$$\max_{0 \leq h \leq M} (M - h) r_D + E[h(r_R(\eta_1) - c_h)].$$

Taking x_i as given, $y_i(\eta_1)$ and $h_i(\eta_1)$ solve for each η_1

$$\max_{y_i, h_i} -(y_i + h_i) r_R(\eta_1) - [-y_i]^+ c_B + v(x_i + \eta_i + y_i + h_i).$$

Note that the function v is defined in the main text and we have omitted terms that enter the bank's final payoff but do not affect the choice of y_i or h_i. To formulate bank i's choice of x_i define a function $w(x_i, \eta_i, r_R)$ as

$$w(x_i, \eta_i, r_R) = \max_{y_i, h_i} -(y_i + h_i) r_R - [-y_i]^+ c_B + v(x_i + \eta_i + y_i + h_i).$$

Then x_i solves

$$\max_{x_i} -x_i r_D + E[w(x_i, \eta_i, r_R)].$$

The market clearing conditions are

$$\begin{aligned} x_1 + x_2 &= M - h - R_1 - R_2, \\ y_1(\eta_1) + y_2(\eta_1) &= 0, \\ h_1(\eta_1) + h_2(\eta_1) &= h \end{aligned}$$

for all η_1.

7.2 Proofs for section 3

Proof of lemma 1: The lemma says that in the baseline model of section 3, $y_1 \leq 0$ if $x_1 + \eta_1 \geq x_2 + \eta_2$. Suppose instead that $y_1 > 0$. Then, bank 1's profit-maximizing implies

$$\begin{aligned} v(x_1 + \eta_1 + y_1) - y_1 r_R &\geq v(x_1 + \eta_1). \\ v(x_1 + \eta_1 + y_1) - v(x_1 + \eta_1) &\geq y_1 r_R. \end{aligned}$$

In equilibrium $y_2 = -y_1$. Therefore, bank 2's profit-maximizing implies

$$v(x_2 + \eta_2 - y_1) + y_1 r_R - y_1 c_B \geq v(x_2 + \eta_2).$$
$$v(x_2 + \eta_2) - v(x_2 + \eta_2 - y_1) \leq y_1 r_R - y_1 c_B.$$

Since $x_1 + \eta_1 \geq x_2 + \eta_2$ and v is concave, we have $v(x_2 + \eta_2) - v(x_2 + \eta_2 - y_1) \geq v(x_1 + \eta_1 + y_1) - v(x_1 + \eta_1)$. All these inequalities together imply that

$$y_1 r_R \leq v(x_2 + \eta_2) - v(x_2 + \eta_2 - y_1) \leq y_1 r_R - y_1 c_B,$$

a contradiction.

Proof of lemma 2: The lemma says that if $x_1 + \eta_1 \geq x_2 + \eta_2$ then $x_1 + \eta_1 + y_1 \geq x_2 + \eta_2 + y_2 = x_2 + \eta_2 - y_1$. If $y_1 = 0$, then the lemma holds. Suppose that $y_1 < 0$ and contrary to the lemma $x_1 + \eta_1 + y_1 < x_2 + \eta_2 - y_1$. Since v is concave and v' continuous, bank 1's profit-maximizing implies

$$v'(x_1 + \eta_1 + y_1) \leq r_R - c_B.$$

Bank 2's profit-maximizing implies

$$v'(x_2 + \eta_2 - y_1) \geq r_R.$$

Since $x_1 + \eta_1 + y_1 < x_2 + \eta_2 - y_1$, we have that $v'(x_1 + \eta_1 + y_1) \geq v'(x_2 + \eta_2 - y_1)$. Therefore the profit-maximizing conditions imply

$$r_R \leq v'(x_2 + \eta_2 - y_1) \leq r_R - c_B,$$

a contradiction.

Concavity of $u(x_i)$: We define $u(x_i) = Ew(x_i, \eta_i, r_R)$, where the expectation is taken over all possible realizations of η_1. To show that u is concave it is enough to show that $w(x_i, \eta_i, r_R)$ is concave with respect to x_i for any fixed η_i and $r_R(\eta_i)$. For simplicity, write $w(x_i)$ as a function of only x_i. Recall that

$$w(x_i) = \max_{y_i} [-y_i r_R - [-y_i]^+ c_B + v(x_i + \eta_i + y_i)].$$

Since v is concave, we can apply the envelope property to show that w is concave. Formally, let a and b $(a < b)$ such that $v'(a) = r_R$ and $v'(b) = r_R - c_B$. For $x_i < a - \eta_i$, the optimal choice of y_i is $y_i = a - \eta_i - x_i$. Therefore, $w(x_i) = (x_i + \eta_i - a) r_R + v(a)$, and $dw/dx = r_R$. Similar arguments show that for $x_i > b - \eta_i$, $dw/dx = r_R - c_B$.

For $a - \eta_i < x_i < b - \eta_i$, the optimal choice of y_i is simply $y_i = 0$, and $dw/dx = v'(x_i)$, which is decreasing in x_i and is between $r_R - c_B$ and r_R. Therefore, $w(x_i)$ is concave, and thus, $u(x_i)$ is concave.

If $v'(a) < r_R$ for all a we can simply define $a = -\infty$. Similarly we can define $b = \infty$ if $v'(b) > r_R - c_B$ for all b. Note that it is not possible in equilibrium that $v'(a) > r_R$ for all a or $v'(b) < r_R - c_B$ for all b because either condition makes both banks borrow or lend infinite amounts of reserves violating market clearing conditions. If $v'(a) = r_R$ for multiple values of a we can pick any a and the arguments will still work. Similarly we can pick any b if $v'(b) = r_R - c_B$ for multiple values of b.

Proof of lemma 3: This lemma effectively characterizes the equilibrium. It states that

$$u'(x_i) = v'(x_i) = v'(\frac{X}{2}) = r_D.$$

This equation determines x_i. Once x_i is determined we can solve for period 2 borrowing y_i and rate r_R and characterize the equilibrium fully. We will also show that $u(x_i)$ does not depend on i in equilibrium as part of the proof.

First we show $x_1 = x_2 = X/2$ in equilibrium if excess reserves are scarce under definition 2. Suppose that this is not the case. Then without loss of generality we can assume that there exists an equilibrium in which $x_1 > x_2$. Furthermore let $r_R(\eta_1)$ be the interbank rate in that equilibrium. For any η_1 such that bank 1 lends a positive amount to bank 2, $y_1 < 0$, the following holds:

$$v'(x_1 + \eta_1 + y_1) + c_B = v'(x_2 - \eta_1 - y_1) = r_R(\eta_1).$$

Note that the solution $x = b$ to the equation $v'(x) + c_B = v'(X - x)$ is unique. First it exists because the $v'(x)$ decreases from r_W to $r_E - c_B$ as x increases and $r_W - (r_E - c_B) > c_B$. If the latter inequality is not satisfied then $y_1 = 0$ for all η_1 and the interbank market shuts down. Second it is unique because for any solution x to the equation, x or $X - x$ is on the steep part of v', i.e., where $v''(x) < 0$.

Since $x = x_1 + \eta_1 + y_1$ is also a solution to $v'(x) + c_B = v'(X - x)$ it follows that $x_1 + \eta_1 + y_1 = b$. Therefore for any η_1 such that bank 1 lends a positive amount to bank 2, $r_R(\eta_1)$ is the same, denoted $r_0 \equiv v'(b) + c_B$. Similarly for any η_1 such that bank 1 borrows a positive amount from bank 2, $r_R(\eta_1) = r_0$.

It can be seen from the previous discussion that if $x_1 > b - \eta_1$ then $y_1 < 0$ and $\partial w(x_1, \eta_1)/\partial x_1 = r_0 - c_B$ in equilibrium: $r_R(\eta_1)$ must be equal to r_0 for such η_1 in equi-

librium. Also, if $x_1 < a - \eta_1$ where $a \equiv X - b$, then $y_1 > 0$ and $\partial w(x_1, \eta_1)/\partial x_1 = r_0$ in equilibrium. Finally, if $b - \eta_1 < x_1 < a - \eta_1$, then $y_1 = 0$ and $\partial w(x_1, \eta_1)/\partial x_1 = v'(x_1 + \eta_1)$ in equilibrium. These conditions hold for x_2 as well due to the symmetry between the two banks.

Recall that $u_1(x_1) = Ew(x_1, \eta_1)$. Therefore $u_1'(x) = E[\partial w(x_1, \eta_1)/\partial x_1]$. We can expand this equation using $\partial w(x_1, \eta_1)/\partial x_1$ derived earlier,

$$u_1'(x) = \int_{-\bar{\eta}}^{\bar{\eta}} h(v'(x_1 + \eta))g(\eta)d\eta,$$

where $g(\eta)$ is the probability density function of η and $h(z)$ is a function such that $h(z) = r_0$ if $z > r_0$, $h(z) = r_0 - c_B$ if $z < r_0 - c_B$, and $h(z) = z$ otherwise. Note that the value of $u_1'(x_1)$ depends only on the distribution of $x_1 + \eta_1$. Since η_2 has the same distribution as η_1, $u_2'(x_2) = u_1'(x_2)$. Therefore in equilibrium $u_i(x_i)$ does not depend on the label i and we can drop the dependence on i, simply writing it as $u(x_i)$.

Let a' be the supremum of the set of x such that $v'(x) = r_0$, and let b' be the infimum of the set of x such that $v'(x) = r_0 - c_B$. Clearly $a' \geq a$ and $b' \leq b$. It is not always the case that $a' = a$ or $b' = b$ because a or b can fall on the flat region of v', i.e., where $v''(x) = 0$. Note that $h(v'(x+\eta))$ is strictly decreasing if $a' < x + \eta < b'$, flat at r_0 if $x + \eta < a'$, and flat at $r_0 - c_B$ if $x + \eta > b'$. Therefore, $u''(x) < 0$ if $a' - \bar{\eta} < x < b' + \bar{\eta}$ and $u''(x) = 0$ otherwise.

In an equilibrium $u'(x_1) = u'(x_2)$. Furthermore $u'(x)$ is monotonically decreasing which means that $u'(x) = u'(x_1)$ for any $x_2 \leq x \leq x_1$. At the same time we have been assuming that $x_1 > x_2$. Therefore $u'(x_1) = u'(x_2)$ can happen only if $x_1 \leq a' - \bar{\eta}$ or $x_2 \geq b' + \bar{\eta}$. If $x_1 \leq a' - \bar{\eta}$ then $x_1 + x_2 < 2(a' - \bar{\eta})$. Note that $a < X/2$ by construction and therefore if $a = a'$ then $x_1 + x_2 < X/2$ violating a market clearing condition. If $a < a'$ then a' is the upper limit of the flat region $v'(x) = r_W$ and $a' = -\bar{\nu}$. Thus $x_1 + x_2 < 2(-\bar{\nu} - \bar{\eta})$ contradicting assumed scarcity under definition 2.

Similarly we can show that $x_2 \geq b' + \bar{\eta}$ leads to a contradiction. Therefore we have established $x_1 = x_2$ in equilibrium. Note that the scarcity of excess reserves has been used only to contradict $x_1 + x_2 < 2(-\bar{\nu} - \bar{\eta})$ and $x_1 + x_2 > 2(\bar{\nu} + \bar{\eta})$. This means that to have excess reserves equally distributed between the two banks, $x_1 = x_2 = X/2$, we only need to assume $-\bar{\nu} - \bar{\eta} < X/2 < \bar{\nu} + \bar{\eta}$; excess reserves do not need to be scarce but only need not to be 'really' scarce or abundant.

We now suppose that $x_1 = x_2 = X/2$ and show $u'(x_1) = v'(x_1)$. Given the scarcity

assumption $-\bar{\nu}+\bar{\eta} \leq X/2 \leq \bar{\nu}-\bar{\eta}$, we have that $v'(x_1+\eta)$ is strictly decreasing over $-\bar{\eta} \leq \eta \leq \bar{\eta}$. If $a \leq x_1-\bar{\eta}$, then $y_1 = y_2 = 0$ for every η_1, and $v'(x_1+\eta_1)$ lies strictly between $r_0 - c_B$ and r_0, except possibly at the endpoints $\eta_1 = \pm\bar{\eta}$. Therefore

$$u'(x_1) = \int_{-\bar{\eta}}^{\bar{\eta}} h(v'(x_1+\eta))g(\eta)d\eta = \int_{-\bar{\eta}}^{\bar{\eta}} v'(x_1+\eta)g(\eta)d\eta = v'(x_1).$$

The last equality follows from the linearity of v' and the symmetry of η's distribution.

If $a > x_1 - \bar{\eta} \geq \bar{\nu}$, then a lies on the steep part of $v(x)$. Furthermore this implies that $b = X - a < x_1 + \bar{\eta}$ and b lies on the steep part of $v(x)$ as well. Therefore $h(v'(x_1+\eta)) = r_0$ if $\eta \leq a - x_1$, $h(v'(x_1+\eta)) = r_0 - c_B$ if $\eta \geq b - x_1 = -(a - x_1)$, and $h(v'(x_1+\eta))$ linearly decreases from r_0 to $r_0 - c_B$ over $\eta \in [a-x_1, -(a-x_1)]$. This symmetric shape of $h(v'(x_1+\eta))$ implies that $u'(x_1) = v'(x_1)$.

Strict concavity of $u(x_i)$: From the proof of lemma 3 we can see that with scarce excess reserves, $u(x_1)$ is strictly concave around the equilibrium value of x_1, defined under the equilibrium functional form of $r_R(\eta_1)$. It is sufficient to show that there is a positive probability such that $(\partial^2/\partial x_1^2)w(x_1, \eta_1)$ is strictly negative. Using earlier notation this partial derivative is strictly negative if $a < x_1 + \eta_1 < b$. We have seen in the proof of the lemma that this inequality holds with a positive probability. Since $x_1 = x_2$ this result applies to x_2 as well.

Relationship between definitions 1 and 2: We have seen that if reserves are scarce under definition 2, then $x_i = X/2$ in equilibrium and thus $-\bar{\nu}+\bar{\eta} \leq x_i \leq \bar{\nu}-\bar{\eta}$. Therefore reserves are scarce under definition 1 as well in equilibrium.

7.3 Model with large excess reserves

This section supports the discussion on post-crisis (section 4.2) and future (section 5.5) money markets. We solve the model with reserves no longer scarce under definition 2, $-\bar{\nu}+\bar{\eta} \leq X/2 \leq \bar{\nu}-\bar{\eta}$. In doing so we can rely on many of previous results for scarce reserves and modify them as required.

First we solve the model for abundant excess reserves $X/2 \geq \bar{\nu}+\bar{\eta}$. This condition implies that there are enough reserves for both banks to always end up with strictly positive excess reserves at the end of period 3. Since $x_1 + x_2 = X \geq 2(\bar{\nu}+\bar{\eta})$, we assume $x_1 \geq \bar{\nu}+\bar{\eta} \geq x_2$ without loss of generality. For any η_1 such that there is a nontrivial interbank trade, bank 1 lends and never borrows. Bank 1 does not borrow because it borrows only if

$r_R \leq v'(x_1 + \eta_1) = r_E - c_B$. However bank 2 will lend only if $r_R \geq v'(x_2 + \eta_2) + c_B \geq r_E$ which implies that bank 1 will never borrow in equilibrium.

If bank 1 lends, $y_1 < 0$, the interbank rate r_R is determined by $r_R = v'(x) + c_B$, where x is the solution to $v'(x) + c_B = v'(X - x)$ as discussed in the proof of lemma 3. This implies that $x \geq X/2 \geq \bar{\nu} + \bar{\eta}$ and thus $v'(x) = r_E - c_B$ and $r_R = r_E$. Note that given this r_R bank 1 chooses $y_1 \leq 0$ so that $v'(x_1 + \eta_1 + y_1) + c_B \leq r_R = r_E$. Since v' is bounded from below by $r_E - c_B$ it simply means that $v'(x_1 + \eta_1 + y_1) = r_E - c_B$ for all η_1. Therefore $u_1'(x_1) = r_E - c_B$ as well. Also note that $r_R = r_E$ is consistent with no trade, $y_1 = 0$, so in equilibrium $r_R = r_E$ is a constant, $u_i(x_i)$ does not depend on i and we can simply write it as $u(x_i)$.

Equilibrium requires that $u'(x_2) = u'(x_1) = r_E - c_B$. This implies that bank 2 always has enough excess reserves at the beginning of period 2 in the sense that $x_2 + \eta_2 \geq \bar{\nu}$. If $x_2 + \eta_2 < \bar{\nu}$ with a positive probability then $u'(x_2) > r_E - c_B$ with a positive probability. The reason is that with $x_2 + \eta_2 < \bar{\nu}$, the marginal value of reserves to bank 2 is either $v'(x_2 + \eta_2) > r_E - c_B$ or the interbank rate, $r_R = r_E$, and is greater than $r_E - c_B$ in either case. Therefore $x_2 \geq \bar{\nu} + \bar{\eta}$.

We have shown that in equilibrium, $x_1, x_2 \geq \bar{\nu} + \bar{\eta}$. Also we can see that any such x_1 and x_2 constitute an equilibrium as long as $x_1 + x_2 = X$. Since $x_i + \eta_i \geq \bar{\nu}$ for both banks, there is no interbank trade in period 2, and the interbank rate is indeterminate as $r_E - c_B \leq r_R \leq r_E$. The deposit rate is simply determined as $r_D = u'(x_1) = u'(x_2) = r_E - c_B$.

Next we characterize the model in the intermediate case with excess reserves between $\bar{\nu} - \bar{\eta}$ and $\bar{\nu} + \bar{\eta}$. Let s be the slope of v' in its steep part. We assume that $\bar{\eta} > c_B/(2s)$; otherwise there is no interbank trading under any parameters. For any level of total excess reserves X that is not abundant or 'really' scarce, $-\bar{\nu} - \bar{\eta} < X/2 < \bar{\nu} + \bar{\eta}$, we have $x_1 = x_2$. This has been discussed in the proof of lemma 3.

If $\bar{\nu} + \bar{\eta} - c_B/s \leq X/2 < \bar{\nu} + \bar{\eta}$ there is no interbank trade because the difference in v' after η shocks between the two banks is always less than c_B. However note that $v'(x_i + \eta_i)$ can rise above $r_E - c_B$ for small enough negative shocks such that $\eta_i < -\bar{\eta} + c_B/s$. Thus $r_D > r_E - c_B$. Also it is obvious that r_D is decreasing in X because v' is decreasing.

If $\bar{\nu} - c_B/(2s) < X/2 < \bar{\nu} + \bar{\eta} - c_B/s$ then the volume of interbank trade increases as X decreases. Let us assume $\eta_1 \geq 0$ without loss of generality making bank 1 as the potential lender. There will be a positive interbank volume if and only if $x_2 - \eta_1 < \bar{\nu} - c_B/s$. The volume of trade $-y_1$ is determined by $v'(x_2 - \eta_1 - y_1) = v'(x_1 + \eta_1 + y_1)$. Since $x_1 + x_2 > 2\bar{\nu} - c_B/s$

it is always the case that $x_2 - \eta_1 - y_1 = \bar{\nu} - c_B/s$ and $x_1 + \eta_1 + y_1 > \bar{\nu}$. Therefore $r_R = r_E$.

At $X/2 = \bar{\nu} - c_B/(2s)$ the banks would act as if excess reserves were scarce. The banks do not care about that v' is flat if η_i is greater than $c_B/(2s)$ because any extra excess reserves beyond $X/2 + c_B/(2s)$ will be lent to the other bank in the interbank market. Therefore we can broaden definition 2 as $-\bar{\nu} + c_B/(2s) \leq X/2 \leq \bar{\nu} - c_B/(2s)$. With $X/2 = \bar{\nu} - c_B/(2s)$ the rates are $r_D = r_E - c_B/2$ and $r_R = r_E$. Any further reduction in excess reserves will increase r_D and r_R as in the scarce excess reserves case.

7.4 Model with constant marginal cost for late deposits

We first describe how the equilibrium can be derived, supplementing sections 5.1, 5.2, 5.3, and 5.4.1. Then we explain how the equilibrium can be derived for different levels of excess reserves, supplementing 5.5.

We solve for the equilibrium in the main text assuming $x_1 = x_2$. Intuitively this makes sense; it can be proved by following steps similar to those in the proof of lemma 3, which we now explain. We assume $\eta > c_B/(2s)$ to ensure that there is some interbank trade. Otherwise the model is less interesting but the equilibrium can still be characterized in the essentially same way. Also we assume $c_h > 0$. Otherwise the model becomes much simpler because in equilibrium the investor will lend enough in period 2 so that the marginal values of the two banks are always the same. In that case there is no need for $x_1 = x_2$ as long as h is sufficiently large to ensure that both banks end up with $X/2$ reserves after the η shock and period 2 lending.

To solve for equilibrium we fix the value of h and look for an equilibrium consistent with that value. In equilibrium h and X determine $r_R(\eta_1)$ and $v'(x_i + \eta_i + y_i + h_i)$ in period 2. These can be used to determine r_D in period 1. Therefore for a given X we can determine $E[r_R(\eta_1)] - r_D$ as a function of h. Then we can find an equilibrium by finding the value of h such that $E[r_R(\eta_1)] - r_D = c_h$.

First let us assume $h = 0$ and solve for outcomes in the period 2 market. This can be seen by just following arguments used in the proof of lemma 3 which is possible because $h = 0$. Following an earlier notation let $x = b$ be the solution to $v'(x) + c_B = v'(X - x)$. Then obviously $b > X/2$.

If $\eta_1 < X/2 - b$ or $\eta_1 > b - X/2$ there is a nonzero interbank trade and $r_R = v'(X - b) = v'(X) + c_B/2$. If $X/2 - b \leq \eta_1 \leq b - X/2$ then $r_R = max\{v'(X/2 + \eta_i) | i = 1, 2\}$. Here

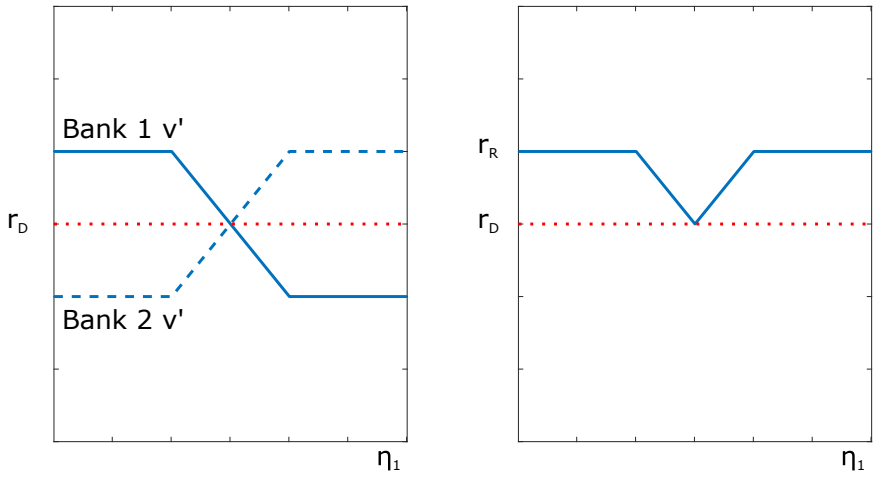

Figure 12: marginal values and interbank rate.

we treat r_R as uniquely determined even if $h = 0$; this is technically incorrect because we have indeterminancy in r_R as in the baseline model. However we need to use this formula for r_R to evaluate the investor's willingness to pay the delay cost, $E(r_R) - r_D$, at $h = 0$. If we imagine that h is very small instead of being zero, the existence of non-bank lending in period 2 pins down the rate in period 2 as determined by this previous formula.

$E[r_R]$ is generally above r_D because $E[r_R]$ is the expected marginal value of the bank with the higher marginal value in period 2 while r_D is the expected marginal value of a given bank. Figure 12 illustrates this graphically. The difference $E[r_R] - r_D$ can be computed directly as the area between $r_R(\eta_1)$ and r_D:

$$E[r_R] - r_D = \frac{c_B}{2}[1 - \frac{1}{4}\frac{c_B}{s\bar{\eta}}],$$

where $s = (r_W - r_E + c_B)/(2\bar{\nu})$ is the slope of v' over its steep part. Therefore, for any c_h greater than this value of $E[r_R] - r_D$, the equilibrium level of the investor's period 2 lending is zero: $h = 0$.

Let us consider a more general case with $h > 0$. We can solve for the interbank rate r_R as functions of $x_i + \eta_i$. If the difference in reserves is small enough, $|(x_1+\eta_1)-(x_2+\eta_2)| \leq h$, then h alone can equalize marginal values between the two banks, $x_1+\eta_1+h_1 = x_2+\eta_2+h_2 = X/2$ and $y_1 = y_2 = 0$. In this case $r_R = v'(X/2)$. If $|(x_1 + \eta_1) - (x_2 + \eta_2)| > h$ then there are two possibilities. First h may be enough to reduce the difference in marginal values to under c_B, in the sense that $|(x_1 + \eta_1) - (x_2 + \eta_2)| - h \leq s^{-1}c_B$. If this happens, the bank i with smaller $x_i + \eta_i$ borrows all of h so that $r_R = v'(x_i + \eta_i + h)$, and there is no interbank trade $y_1 = y_2 = 0$. Otherwise $|(x_1+\eta_1)-(x_2+\eta_2)|-h > s^{-1}c_B$ and there is a nontrivial interbank

trade in the equilibrium. In this case the bank i with smaller $x_i + \eta_i$ borrows all of h, and at the same time also borrows in the interbank market, $y_i > 0$, so that $x_i + \eta_i + h + y_i = X - b$.

Using this characterization we can show that $x_1 = x_2$ following arguments similar to those in used in proving lemma 3 for the baseline model. Suppose that $x_1 > x_2$ in equilibrium. We can compute $w' \equiv \partial w(x_1, \eta_1, r_R)/\partial x_1$ as a function of η_1.

Following the notation used in the proof of 3, we define $a = X - b$, $b' = inf\{z|v'(z) = v'(b)\}$ and $a' = sup\{z|v'(z) = v'(a)\}$. There are five distinct intervals of η_1 within which w' has distinct expressions:

1. If $b' - x_1 \leq \eta_1$ then $w' = r_0 - c_B$ where $r_0 \equiv v'(b) + c_B$.
2. If $X/2 - x_1 \leq \eta_1 \leq b' - x_1$ then $w' = v'(x_1 + \eta_1)$.
3. If $X/2 - x_1 - h \leq \eta_1 \leq X/2 - x_1$ then $w' = v'(X/2)$.
4. If $a' - x_1 - h \leq \eta_1 \leq X/2 - x_1 - h$ then $w' = v'(x_1 + \eta_1 + h)$.
5. If $\eta_1 \leq a' - x_1 - h$ then $w' = r_0$.

Note that w' is decreasing in x_1. Also note that $\partial w(x_2, \eta_2, r_R)/\partial x_2$ has the same form as $\partial w(x_1, \eta_1, r_R)/\partial x_1$ if we replace x_1 and η_1 by x_2 and η_2 in the earlier characterization of the latter. The equilibrium requires

$$\int_{-\bar{\eta}}^{\bar{\eta}} \frac{\partial}{\partial x_1} w(x_1, \eta, r_R(\eta)) g(\eta) d\eta = \int_{-\bar{\eta}}^{\bar{\eta}} \frac{\partial}{\partial x_2} w(x_2, \eta, r_R(-\eta)) g(\eta) d\eta, \tag{26}$$

where $g(\eta)$ is the probability density function for η_i. Given the form of w' and the assumption $x_1 > x_2$, this equation holds if and only if x_1, x_2, h and X are such that η_1 and η_2 are both always in intervals 1, 3 or 5 so that $\partial w/\partial x_1$ and $\partial w/\partial x_2$ are always the same constant; w' is strictly decreasing in intervals 2 and 4 so η_1 and η_2 must never touch these intervals to make equation (26) hold.

Suppose that both η_1 and η_2 always end up in interval 1. Then the marginal values of both banks after period 2 market are always $r_0 - c_B$, making $r_R = r_D = r_0 - c_B$. Therefore given $c_h > 0$ it must be $h = 0$ which is a contradiction. The same can be said of cases where both η_1 and η_2 always end up in intervals 3 or 5.

Therefore the necessary condition for equilibrium, equation (26), cannot hold if $x_1 > x_2$. Similarly $x_2 > x_1$ is not allowed in equilibrium and thus $x_1 = x_2$.

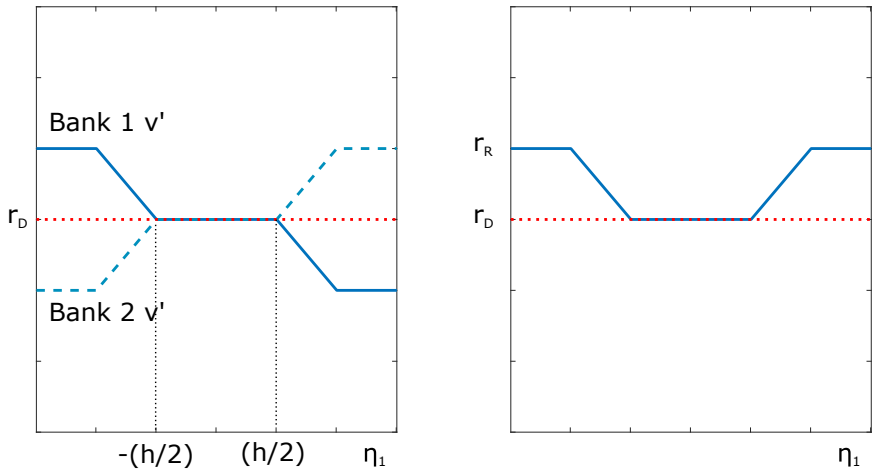

Figure 13: marginal values and interbank rate with $h > 0$.

As discusssed in the main text even with $h > 0$ the expected value of $v'(x_i + \eta_i + h_i + y_i)$ is still $v'(X/2)$ so r_D does not change with h. However, $E[r_R]$ decreases as h increases because h introduces a flat part of zero of length h around $\eta_1 = 0$. Figure 13 illustrates this. The figure also makes it clear that for $h \geq 2\bar{\eta}$, the difference $E[r_R] - r_D$ is zero. Intuitively $h \geq 2\bar{\eta}$ is enough to remove the difference in excess reserves between the two banks in period 2 and completely offset η shocks.

7.4.1 When excess reserves are not scarce

For a given level of excess reserves, allowing period 2 lending by the investor weakly decreases interbank volume because it weakly reduces the difference in marginal values pre-interbank trading between the two banks under any realization of η_i. This can be seen from figure 13 where the flat sections of v' near the lower and upper end of η_1 become smaller due to $h > 0$.

A less obvious fact is that even with some period 2 lending by the investor, interbank trade volume increases monotonically as reserves are drained. This can be seen formally by solving the model for different values of X as discussed in section 5.5.

As in the baseline model, if excess reserves are abundant, $X/2 \geq \bar{\nu} + \bar{\eta}$, there is no interbank trade or period 2 lending by the investor and $r_D = r_E - c_B$.

To describe how the equilibrium changes as $X/2$ decreases within the intermediate region, $\bar{\nu} - \bar{\eta} \leq X/2 \leq \bar{\nu} + \bar{\eta}$, we first define $f(X) \equiv E[r_R] - r_D$ with $c_h = \infty$ as a function of total excess reserves. In other words it is the expected per-unit revenue from lending an

infinitesimal amount of fund in period 2 for the investor. Generally $r_R(\eta_1)$ is indeterminate if η_1 is such that there is no interbank trade. In such a case r_R is defined as the interest rate that would prevail with an infinitely small but nonzero h or equivalently the marginal value v' of the bank with fewer excess reserves.

Using our prior analysis of the benchmark model we can show that $f(X)$ is a decreasing function. We skip the computational steps here but the result is intuitive: As excess reserves decrease the demand for borrowing in period 2 increases thus increasing the willingness to pay for a period 2 loan.

For the purpose of illustration we assume that $f(\hat{X}) = c_h$ for some positive constant $2\bar{\nu} < \hat{X} < 2\bar{\nu} + 2\bar{\eta} - c_B/s$. Then $f(X) < c_h$ for $X = 2(\bar{\nu} + \bar{\eta} - c_B/s) > \hat{X}$. This means that some level of interbank activity will come back before the investor decides to lend in period 2.

As discussed previously, for $\bar{\nu} + \bar{\eta} - c_B/s \leq X/2 \leq \bar{\nu} + \bar{\eta}$ there is no interbank trade but r_D strictly increases from $r_E - c_B$ as $X/2$ decreases from the upper bound.

For $\hat{X}/2 \leq X/2 \leq \bar{\nu} + \bar{\eta} - c_B/s$ there is no period 2 lending by the investor and the equilibrium is the same as in the model with $c_h = \infty$: As total excess reserves decrease, interbank lending volume increases, $r_R = r_E$ and r_D increases.

For $\bar{\nu} \leq X/2 \leq \hat{X}/2$ the equilibrium outcomes are constant except that period 2 late lending by the investor increases as excess reserves decrease. To maintain $E[r_R] - r_D = c_h$, the investor sets the supply of period 2 lending h so that $h = \hat{X} - X$. Computing the shape of w' shows that this makes the shape of w' as a function of η_i invariant for any given X and corresponding $h = \hat{X} - X$.

Finally over $\bar{\nu} - c_B/(2s) \leq X/2 \leq \bar{\nu}$, as excess reserves decline equlibrium outcomes no longer stay constant. The reason is that if $X/2 \leq \bar{\nu}$ then v' has a negative slope on both sides of $X/2$: On the contrary if $X/2 \geq \bar{\nu}$ then $v'(x)$ has a negative slope only for some $x < X/2$. As $X/2$ declines over the region $\bar{\nu} - c_B/(2s) \leq X/2 \leq \bar{\nu}$ the volumes of both interbank trade and period 2 non-bank lending increase, the lower limit of r_R increases from $r_E - c_B$ to $r_E - c_B/2$, and r_D increases to $r_E - c_B/2$. The upper limit of r_R stays at r_E.

For $-\bar{\nu} + c_B/(2s) \leq X/2 \leq \bar{\nu} - c_B/(2s)$, excess reserves can be considered scarce: r_D and the lower and upper limits of r_R all increase at the same rate as total excess reserves decline. The volume of interbank trades and the amount of period 2 lending by the investor stay constant.

7.5 Model with fixed cost for late deposits

We compute the total private surplus of allowing the investor to lend in the interbank market at no cost. The total surplus is defined as the sum of the surplus of the investor and the two banks. There are two sources of surplus. First it displaces interbank lending completely, $y_1 = y_2 = 0$, and saves balance sheet costs. Second free period 2 lending by the investor always equalizes the marginal values of the two banks, $v'(x_1 + \eta_1 + y_1 + h_1) = v'(x_2 + \eta_2 + y_2 + h_2)$. Equalizing marginal values increases surplus by reducing the system-wide cost of discount window borrowing, caused by the rate penalty $r_W - r_E$ and the balance sheet cost c_B.

Let $S(\eta_1)$ be the extra surplus from free period 2 lending by the investor, conditional on η_1. If $X/2 \geq \bar{\nu} + \bar{\eta}$ then both banks always end up with $v' = r_E - c_B$ even without any nonbank lending in period 2, and $S(\eta_1)$ is always zero. Therefore, we only consider $-\bar{\nu} - \bar{\eta} < X/2 < \bar{\nu} + \bar{\eta}$.

From lemma 3 we know that $x_1 = x_2 = X/2$ in equilibrium if the public good is not constructed. Also there is no interbank trade if $|\eta_1|$ is small enough so that $v'(x_1 - |\eta_1|) \leq v'(x_1 + |\eta_1|) + c_B$. Let $\eta_0 > 0$ be the maximum value of such $|\eta_1|$ which is a function of X. For $|\eta_1| \leq \eta_0$,

$$S(\eta_1) = \int_0^{|\eta_1|} [v'(x_1 - \eta) - v'(x_1 + \eta)] d\eta.$$

For $|\eta_1| > \eta_0$, the surplus from just equalizing marginal values is $S(\eta_0)$. In addition there is extra surplus from saving balance sheet costs from interbank trades. For $|\eta_1| > \eta_0$

$$S(\eta_1) = S(\eta_0) + (|\eta_1| - \eta_0) c_B.$$

Given the shape of v' it is clear that the expected total surplus $E[S(\eta_1)]$ is a decreasing function of X. Therefore as reserves are drained it is possible that after a certain amount of draining, $X \leq X^*$ for some X^*, the surplus exceeds the cost C_h and late trading becomes free for the investor eliminating all interbank trades. If C_h is such that there is some interbank trading at $X = X^*$ then there is non-monotonicity in how the volume of interbank trading responds to a drain of reserves. Similarly and as illustrated by figure 11 there can be non-monotonicity in the path of the deposit rate r_D.

8 References

Afonso, G. and R. Lagos (2015) "Trade Dynamics in the Market for Federal Funds" Econometrica 83(1), 263 - 313.

Armenter R. and B. Lester (2017) "Excess Reserves and Monetary Policy Normalization," Review of Economic Dynamics 23, 212 - 235.

Bindseil, U. (2016) "Evaluating Monetary Policy Operational Frameworks." Federal Reserve Bank of Kansas City's Jackson Hole Symposium.

Chen, H., J. A. Clouse, J. Ihrig, and E. Klee (2016) "The Federal Reserve's Tools for Policy Normalization in a Preferred Habitat Model of Financial Markets," Journal of Money, Credit and Banking, vol. 48, no. 5, pp. 921 - 955.

Ennis, H. M. and T. Keister (2008) "Understanding Monetary Policy Implementation." Federal Reserve Bank of Richmond Economic Quarterly.

Martin, A., J. J. McAndrews, A. Palida, and D. R. Skeie (2013) "Federal Reserve Tools for Managing Rates and Reserves," Federal Reserve Bank of New York Staff Report, No 642.

Norges Bank (2011) "Background for the System for Managing Bank Reserves in Norway." Website information (http://www.norges-bank.no/en/Liquidity-and-markets/The-liquidity-management-system/The-management-of-bank-reserves-The-system-in-Norway/Background-system-managing-bank-reserves/).

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

Potter, S. (2016) Discussion of "Evaluating Monetary Policy Operational Frameworks" by Ulrich Bindseil. https://www.newyorkfed.org/newsevents/speeches/2016/pot160826.

Potter, S. (2017) "Money Markets at a Crossroads: Policy Implementation at a Time of Structural Change." https://www.newyorkfed.org/newsevents/speeches/2017/pot170405. Figures at: https://www.newyorkfed.org/medialibrary/media/newsevents/speeches/2017/pot140405/pot170405.pdf.