Liquidity saving mechanisms

Antoine Martin and James McAndrews *

Federal Reserve Bank of New York

September 2006

Abstract

We study the incentives of participants in a real-time gross settlement with and without the addition of a liquidity saving mechanism. Participants in our model face a liquidity shock and different cost of delaying payments. They trade-off the cost of delaying a payment with the cost of borrowing liquidity from the central bank. The heterogeneity of participants in our model gives rise to a rich set of strategic interactions. The main contribution of our paper is to show that the design of a liquidity saving mechanism has important implications for welfare. In particular, we find that adding one type of liquidity saving mechanism can either increase or decrease welfare depending on parameters.

JEL classification: E42, E58, G21

Keywords: Liquidity saving mechanisms; Real-time gross settlement; Large-value payment systems

*We thank James Chapman for useful comments. We also thank Enghin Atalay for excellent research assistance. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of New York or the Federal Reserve System.

1 Introduction

Large-value payment systems are an increasingly important part of monetary economies that rely on fractional reserve banking systems to settle the claims of the economy's agents. In recent decades, for example, the turnover on U.S. dollar large-value payment systems has increased from 68 times annual GDP per year in 2000, to 102 times annual GDP in 2004. Correspondingly, a great deal of attention has been paid in recent years to features of their design and operation. Most notably, designs known as "real-time gross settlement" (RTGS) systems have proliferated widely in the last 15 years. More recently, various liquidity savings mechanisms (LSMs) have been designed or put in place to operate in conjunction with several RTGS systems.

The efficiency of the operation of these monetary transfer systems depends on the behavior of their participants. We examine a model of an RTGS system in which this behavior is influenced by two key features. First, carrying negative balances is costly for the system participants, while having positive reserves has no direct benefits. Second, participants are heterogeneous.

Starting the day with zero balances on account at a central bank, the system participants face an intraday borrowing costs to acquire the balances required to make a payment. In our model, system participants are subject to two types of shocks. First, some payments may turn out to be time-sensitive; if they are, the participant would suffer a cost if the payment is delayed. Second, the participant may be required to make a pay-in to or to receive a pay-out from a settlement system early in the day. Various combinations of these two shocks create quite a rich array of situations that a participant could face when it decides to make a payment or to delay it. Participants seek to settle their payments to minimize the sum of their delay costs and their intraday borrowing costs.

By modeling heterogeneity of payment types and participants' liquidity positions, our model adds significant new elements to the existing literature. In addition, we focus attention on the behavior of participants in an RTGS system supplemented by one of two types of LSMs. While LSMs may differ in many ways, we identify two broad types of LSMs, balance-reactive mechanisms, and receipt-reactive mechanisms, which release payments according to two different criteria. The balance-reactive LSM can be thought of as a state-contingent mechanism; a participant can specify a level of reserves above which payments committed to the LSM will settle. In contrast, the receipt-reactive LSM can be thought of as a mechanism that is independent of the liquidity-shock received by the participant; it releases payments dedicated to the LSM in accordance purely with a participant's momentary receipt of payments, and not its previous balance on account. In both types of the LSMs we consider, we allow offsetting payment messages in various participants' queues of LSM payments to settle.

Our model maintains some assumptions that influence our analysis. We assume that participants can obtain funds during the day at a proportional cost. This assumption is meant to approximate the policies of many central banks which provide intraday credit to participants either for a small fee or against collateral pledged to the central bank. We assume that there is no probability of default. In exploring this model, we are concerned with the efficiency of the operation of the system to settle the participants' obligations where the efficiency criterion is to minimize the sum of the delay costs suffered by participants and the costs of borrowing funds intraday. These concerns are fundamental to more general models that would include risks of default by participants. In any case, adding default in our model is straightforward and is left for future research. Finally, we assume that participants cannot have access to an intraday market within which they could borrow and lend funds on account.

We arrive at a number of conclusions. First, and perhaps most surprising, the ordering of welfare between the levels achieved in the three systems is as follows: a receipt-reactive LSM always generates welfare at least as high, and sometimes higher, than the levels achieved in an RTGS; at the same time a balance-reactive LSM can generate levels of welfare below RTGS or above the level achieved in a receipt-reactive LSM. These conclusion may be surprising in that a common intuition has been that system of an RTGS supplemented by a balance-reactive LSM can duplicate any outcome of an RTGS alone, and can therefore provide higher welfare. The intuition provided by our model is different. In a pure RTGS system, participants must either send a payment or delay that payment before they know whether they will receive an offsetting payment. Hence, while it might be ex-ante desirable to send a payment, a participant may regret, ex-post, having sent that payment if they do not receive an offsetting payment. Nevertheless, the high degree of coordination that can be achieved when many participants make their payments early benefits all system participants. In contrast, a balance-reactive LSM gives agents the option to condition the release of a payment on the receipt of an offsetting payment. While this is perceived to be beneficial by individual participants, it can reduce the incentives participants have to make payments early and therefore hurt all participants as more payments are delayed. Another novel conclusion, but one that is in accord with industry practice and intuition (and in contrast to the existing literature) is that only payments that have sufficiently low cost of delay are put in LSM queues. Our model also shows that a receipt reactive LSM makes agents decisions independent of each other. With such a design, there is a unique equilibrium while a pure RTGS system displays multiple equilibria. Whether there are multiple equilibria with a balance-reactive LSM depends on the underlying graph of payments. Finally, we find that if liquidity shocks are small, a balance-reactive LSM yields higher welfare than a receipt-reactive LSM or RTGS. When liquidity shocks are sufficiently large, either the receipt-reactive, or the balance reactive can provide the highest welfare, depending on parameters.

While our results suggest that LSMs can always outperform RTGS systems when taking into account participants' incentives in an environment of liquidity shocks and of differing types of payments, they suggest that the design of the LSM matters. An inappropriately designed LSM can yield a worse outcome than RTGS itself.

1.1 Real Time Gross Settlement Systems and Liquidity Savings Mechanisms

Modern banking systems use large-value transfer systems to settle payment obligations of commercial banks. The payment obligations can represent obligations of bank customers or obligations of the commercial banks themselves. In an RTGS system, now the common design used by most central banks (Bank for International Settlements, 1997), payment orders submitted by an individual participant (typically a bank) to the system are processed individually and released against funds in the bank's account, or against an extension of credit, up to some limit, by the central bank. Because the individual payments are processed and released in a "gross" fashion, that is, the complete value of the payment is transferred from the sender to the receiver when released, the RTGS system is widely recognized to require large amounts of liquidity, in the form of available balances or central bank credit. Alternative systems, such as a netting system in which payments are deferred and released on a "net" basis and only non-offsetting values are transferred between the accounts of banks, require much less liquidity, but impose delays relative to alternatives, such as an RTGS system.

Banks use RTGS system for both customer payments and their own payments. Among the bank's own payments we would note three types. First, a bank often uses an RTGS for the return and delivery of money market loans. The return of money market borrowings are fully known at the time the RTGS opens for business on a particular day. A second type of payment is a payment to a special-purpose settlement system, such as a securities settlement system or a foreign-exchange settlement system. In the U.S., for example, the Continuous Linked Settlement Bank (CLS Bank) is a special-purpose bank that settles foreign-exchange trades on its books. Banks use Fedwire, the Federal Reserve System's RTGS system, to make payments into and to receive payouts from CLS Bank early in the morning hours. The amounts of the payments into CLS Bank may not be known precisely at the start of the Fedwire business day. Finally, another type of payment made by banks on an RTGS are settlement or progress payments under a derivatives contract with another party. The interest rate on a particular day may trigger one of the parties to make a payment to the other; the amount of the payment may not be known in advance.

Both customer-initiated payments and a bank's own payments may or may not be time-sensitive. Consider a payment to settle a real estate transaction of a customer, in which many people are gathered in a closing or settlement meeting. The customer's demand for the payment is highly time-sensitive. Alternatively, a customer may be funding an account of their own at a brokerage firm; as long as the transfer is made on the particular day the customer's demand has been met.

The considerations just outlined suggest that banks are subject to liquidity shocks on any day. They may be required to use the RTGS to pay-in (at least on a net basis) more or less on a particular settlement system that day. In addition, a bank may find itself with many or few time-sensitive payments to make on a particular day.

Liquidity savings mechanisms to be used in conjuction with RTGS systems are a fairly recent phenomenon.[**?**, See McAndrews and Trundle (2001) and Bank for International Settlements (2005) for reviews of LSMs.] At least in part, LSMs are one way to attempt to reduce the demands for liquidity in the RTGS system, while maintaining the flexibility to make timely payments. There are many possible design alternatives for a LSM, but some features are common among all such LSMs. An LSM offers to the bank participating in the payment system two alternatives by which to submit payment orders. The first alternative (sometimes called the "express" route) is to submit the payment order for immediate settlement as though the system were a plain RTGS system. The second alternative is to submit payment orders to the LSM–a queue in which the payment order remains pending some event that will release the payment (this route is sometimes called the "limit order" route). The types of events the could trigger the release of payment orders from the limit queue would be the arrival into the bank's account of sufficient funds so that the bank's balance rises above some threshold, or the appearence in another bank's queue of an offsetting payment, or the receipt by the bank of a payment equal in size to the pending payment order. In all these cases, the release of the payment order in the limit queue is contingent on some state of the world. An LSM offers a new alternative, not available in RTGS, to make the settlement of payments state contingent in a particular way.

1.2 Relevant literature

Several papers examine the theoretical behavior in RTGS sytems. Angelini (1998, 2000) considers the behavior of banks in an RTGS systems in which banks face delay costs for payments as well as costly borrowing costs for funds. His results show that the equilibria of RTGS systems involve excessive delay of payments, as banks don't properly internalize the benefits to banks from the receipt of funds. Bech and Garratt (2003) carefully specify a game-theoretic environment in which they find that RTGS systems can be characterized by multiple equilibria, some of which can involve excessive delay. Roberds (1999) compares gross and net payment systems with systems offering an LSM. He examines the incentives participants have to engage in more risk-taking behavior in the different systems, and finds that under certain circumstances the risk profiles of LSM and net systems are indentical. McAndrews and Trundle (2001) and BIS (2005) provide extensive descriptive material on balance-reactive LSM. Johnson, McAndrews, and Soramaki (2004) introduce the concept of receipt-reactive LSMs and carry out simulation exercises in which they compare the liquidity economizing aspects of receipt-reactive LSMs and netting systems.

Willison (2005) examines the behavior of participants in an LSM and is most similar to our paper. Willison models agents as having an ordering of priority of their payments, from most time-sensitive to least so. That is similar in spirit to our assumption that some participants payments are time-sensitive and others' payments are not. The design of the LSM modeled by Willison admits only the offsetting feature–payments settle out of the LSM queue only when offsetting payments are entered into other participants' LSM queues; as a result we consider a wider array of LSMs in this paper. Willison also models the extension of credit from the central bank as an ex ante amount to be borrowed by participants, while in our paper the credit is tapped ex post, depending on a participant's per period balance. In a crucial difference with our paper, there are no liquidity shocks in Willison's paper.

2 The environment

The economy is populated by a continuum of mass 1 of risk neutral agents. We call these agents the core payment system participants or simply the participants, when there is no risk of confusion. There is also a nonstrategic agent which we identify with a settlement institutions such as CLS Bank. One can think of the nonstrategic agent as aggregating several distinct institutions. Each core participant makes two payments and receives two payments each day. One payment is send to another core participant while the other payment is sent to the nonstrategic agent. Similarly, one payment is received from another core participant and one is received from the nonstrategic agent. Both the payment sent to and received from core participants have size μ. The payments sent to and received from the nonstrategic agent have size $1-\mu$. Consistent with our interpretation of the nonstrategic agent as settlement institutions, we assume that $\mu \geq 1/2$. Our results extend to the case where $\mu < 1/2$.

The economy lasts two periods, morning and afternoon. At the beginning of the morning, core participants learn whether they receive a payment from the nonstrategic agent in the morning or in the afternoon. The probability of the payment being received in the morning is $\bar{\pi}$. We assume that $\bar{\pi}$ also denotes the fraction of core participants who receive a payment from the nonstrategic agent in the morning. More generally, throughout the paper we assume that if x represents the probability of an event occurring for a participant, then the fraction of participants for whom this event occurs is x as well. Core participants also learn whether they must make a payment to the nonstrategic agent in the morning or in the afternoon. The probability of having to make the payment in the morning is $\bar{\pi}$ and is independent of receiving a payment from the nonstrategic agent. We assume that payments to the nonstrategic agent cannot be delayed. Let $\sigma \equiv \bar{\pi}(1-\bar{\pi})$. A fraction σ of agents receive a payment from the nonstrategic agent in the morning and do not need to make a payment until the afternoon. We say that these agents receive a positive liquidity shock. A fraction σ of agents must make a payment from the nonstrategic agent in the morning and do not receive an offsetting payment until the afternoon. We say that these agents receive a negative liquidity shock. The remaining agents, a fraction $1-2\sigma$ make and receive a payment from the strategic agent in the same period, either in the morning or in the afternoon. We say that these agents do not receive a liquidity shock.

Core participants also learn whether the payment they must make to another core participant is time-critical. The probability that a payment is time-critical is denoted by θ. If an agent fails to make a time-critical payment in the morning a cost γ is incurred. Delaying non-time-critical payments until the afternoon has no cost. Core participants must choose whether to make the payment in the morning or in the afternoon before they know if they will receive a payment from another core participant in the morning, but after they know their liquidity shock. Participants form rational expectations about the probability of receiving a payment from some other core participant in the morning. We denote this expectation π.

Each core participant starts the day with zero reserves. Reserves can be borrowed from the CB at an interest cost of R. Participants who receive more payments that they make in the morning have excess reserves. We assume that these reserves cannot be lent to other core participants so that participants receive no benefit from excess reserves.[1] Payments received and sent in the same period offset each other. Hence, a core participant only needs to borrow from the CB if the payments it makes in the morning exceed the payments it receives in the morning.

3 A real-time gross settlement system

In this section, we study a real-time gross settlement system. We start by considering a system is which participants are not affected by liquidity shocks. This assumption is similar to what has been done in previous studies. Next we introduce liquidity shocks and show that many aspects of

[1]We could allow lending between core participants without changing our results as long as the return to lending is strictly less than R. This corresponds to an assumption that there is some cost associated with lending.

the model change.

3.1 RTGS without liquidity shocks

Assume that there are no liquidity shock, $\bar{\pi} = 0$, and normalize the size of the payment to other core participant to one, $\mu = 1$. We let π denote the probability with which participants expects to receive a payment from an other core participant in the morning.

Participants choose to delay their payments if the cost of delay is smaller than the expected cost borrowing if the payment is made early. Since participants who must make a non-time-critical payment face no cost of delay, they never strictly prefer to pay early. Participants who must make a time-critical payment strictly prefer to delay if $\gamma < (1 - \pi)R$.

Throughout the paper, we focus on pure strategy equilibria. Absent liquidity shocks, there are three candidate equilibria. Either all participants pay in the morning ($\pi = 1$), or only participants who must make time-critical payments pay in the morning ($\pi = \theta$), or all participants delay ($\pi = 0$).

If $\gamma < (1-\theta)R$, then core participants with time-critical payments choose to delay their payments if non-time-critical payments are delayed. In this case, either all payments are made in the morning, and $\pi = 1$, or no payments are made in the morning, and $\pi = 0$. If $\gamma > (1 - \theta)R$, then core participants prefer to make time-critical payments in the morning, even if non-time-critical payments are delayed. Hence, $\pi = 0$ is not an equilibrium if γ is large enough. In this case, participants with non-time-critical payments are indifferent between making their payments in the morning or delaying if all other participants pay in the morning. Since we have assumed that payments are made early in such cases, $\pi = 1$ is an equilibrium. Participants with non-time-critical payments prefer to delay if other participants with non-time-critical payments also delay. Hence, $\pi = \theta$ is also an equilibrium.

We calculate welfare in this economy using a ex-ante criterion. Consider all participants prior to their knowing whether their payment is time-critical. Welfare is defined at the expected utility of these agents. When all payments are made in the morning, no costs are incurred and welfare is maximized.

To summarize, when there is no liquidity shock, multiple equilibria can occur. The equilibrium such that all participants make their payment early maximizes welfare since no payment is delayed and no participant has to borrow, as all payments offset.

3.2 Participants behavior under RTGS

In this section, we assume that agents are affected by a liquidity shock, so that $\bar{\pi} > 0$. The size of the payment to other core participant is $\mu \in [1/2, 1)$. The expected cost of making a payment in the morning now depends on the pattern of payments made and received from the nonstrategic agent.

We first consider participants who receive a positive liquidity shock. To derive the cost of sending a payment early, note that with probability π, a payment from another core participant is received in the morning. In this case, participants with a positive liquidity shock have excess reserves between the morning and the afternoon and incur no borrowing cost. With probability $1 - \pi$ no payment is received from another core participant in the morning.

In that case, an amount $\mu - (1 - \mu) = (2\mu - 1)$ must be borrowed from the CB. This amount is the difference between the amount sent to another core participant and the payment received from the nonstrategic agent. Hence, the expected cost of making the payment to another core participant in the morning is $(1 - \pi)(2\mu - 1)R$ for participants who receive a positive liquidity shock.

The cost of delaying a time-critical payment is γ for participants with a positive liquidity shock, since they do not need to borrow. We assume that a participant sends her payment in the morning if she is indifferent between sending it in the morning or the afternoon. Consequently, participants who receive a positive liquidity shock choose to send a time critical payment early if $\gamma \geq (1 - \pi)(2\mu - 1)R$. Since delaying a non-time-critical payment has no cost, such a payment is paid early if $\pi = 1$.

Now consider participants who do not receive a liquidity shock. To derive the cost of sending a payment early, note that with probability π, an offsetting payment is received in the morning and no money needs to be borrowed while with probability $1 - \pi$, no offsetting payment is received and μ must be borrowed. Hence, the expected cost of making the payment to another core participant in the morning is $(1 - \pi)\mu R$ for participants who do not receive a liquidity shock.

The cost of delaying a time-critical payment is γ for these participants, since they do not need to borrow. In summary, participants who receive no liquidity shock send a time-critical payment early if $\gamma \geq (1 - \pi)\mu R$. Non-time-critical payments are paid early if $\pi = 1$.

Finally, consider participants who receive a negative liquidity shock. To derive the cost of sending a payment early, note that with probability π, a payment is received in the morning from another core participant and only $1-\mu$ must be borrowed while, with probability $1-\pi$, no payment is received in the morning and 1 must be borrowed. Hence ,the expected cost of making the payment to another core participant in the morning is $[\pi(1-\mu)+(1-\pi)]\,R = (1-\pi\mu)R$ for participants who have received a negative liquidity shock.

To derive the cost of delay, note that with probability π a payment is received in the morning and, since $\mu > (1-\mu)$, the participant does not need to borrow from the CB. With probability $1-\pi$ no payment is received in the morning and $1-\mu$ must be borrowed. Hence, the expected cost of delaying a time-critical payment is $\gamma + (1-\pi)(1-\mu)R$ for participants who have received a negative liquidity shock.

Hence, participants who receive a negative liquidity shock send their time-critical payment early if $\gamma \geq [\mu - \pi\,(2\mu-1)]\,R$. Non-time critical payments are delayed since $\mu < 1$. It can be verified that $\mu - \pi\,(2\mu-1) \geq (1-\pi)\mu \geq (1-\pi)(2\mu-1)$. We can summarize the results of this section in a proposition.

Proposition 1 *Core participants delay all non-time-critical payments. They make time-critical payment according to the following rules:*

1. *If $\gamma \geq [\mu - \pi(2\mu-1)]\,R$, then all core participants make time-critical payments in the morning.*

2. *If $[\mu - \pi(2\mu-1)]\,R > \gamma \geq (1-\pi)\mu R$, then core participants who re-*

ceive a negative liquidity shock choose to delay time-critical payments. Other core participants do not.

3. *If $(1-\pi)\mu R > \gamma \geq (1-\pi)(2\mu-1)R$, then only core participants who have received a positive liquidity shock choose to make time-critical payments in the morning. All others delay their time-critical payments.*

4. *Finally, if $(1-\pi)(2\mu-1)R > \gamma$, then all core participants delay time-critical payments.*

3.3 Equilibria under RTGS

The probability of receiving a payment in the morning depends on the behavior of the participants in the economy. Hence, π must be determined in equilibrium. We focus on symmetric subgame perfect Nash equilibria in pure strategies. We use the decision rule derived in the previous section to determine equilibrium strategies.

First, we note that if $\bar{\pi}$ and $(1-\mu)$ are strictly positive, then non-time-critical payments are always delayed. Indeed, for such parameters core participants that receive a negative liquidity shock must borrow from the CB if they make a payment early, regardless of what other participants do.

Proposition 2 *Four different equilibria can exist:*

1. *If $\gamma \geq [\mu - \theta(2\mu-1)]\, R$, then it is an equilibrium for all time critical payments to be made in the morning.*

2. *If*

$$\{\mu - \theta\,(1-\sigma)\,(2\mu-1)\}\,R > \gamma \geq [1-\theta\,(1-\sigma)]\,\mu R,$$

then it is an equilibrium for core participants who received a negative liquidity shock to delay time-critical payments while other participants pay time-critical payments in the morning.

3. *If*

$$(1-\sigma\theta)\mu R > \gamma \geq (1-\sigma\theta)(2\mu-1)R,$$

then it is an equilibrium for only core participants who received a positive liquidity shock to make time-critical payments in the morning.

4. *If* $(2\mu-1)R > \gamma$, *then it is an equilibrium for all core participants to delay time-critical payments.*

Proof. If $\gamma \geq [\mu - \pi(2\mu-1)]\,R$, then from proposition 1 we know that all core participants make time-critical payments in the morning. Since the fraction of time-critical payments in the economy is θ, then we have $\pi = \theta$.

If $[\mu - \pi(2\mu-1)]\,R > \gamma \geq (1-\pi)\mu R$, then from proposition 1 we know that core participants who received a negative liquidity shock choose to delay time-critical payments. There is a fraction σ of such participants in the economy. If all other participants sent time-critical payments early, then we have $\pi = \theta\,(1-\sigma)$.

If $(1-\pi)\mu R > \gamma \geq (1-\pi)(2\mu-1)R$, then from proposition 1 we know that only core participants who received a positive liquidity shock choose to make time-critical payments in the morning. This implies that the fraction of delayed time-critical payments is $1-\sigma$ and that $\pi = \sigma\theta$.

If $(1-\pi)(2\mu-1)R>\gamma$, then from proposition 1 we know that all core participants delay time-critical payments and $\pi=0$. ■

An interesting aspect of introducing a liquidity shocks is that it eliminates some equilibria that can arise when such shocks are ignored. With liquidity shocks it is not an equilibrium for all participants to pay early, while this was an equilibrium without liquidity shocks.

The equilibria of proposition 2 can co-exist. For example, equilibria 1 and 2 co-exist if $\theta(2\mu-1)>\gamma\geq\theta\left(1-\sigma\right)\left(2\mu-1\right)$. In fact, for some parameters all four equilibria co-exist, as shown in the following lemma.

Lemma 1 *The four equilibria described in proposition 2 can co-exist if*

$$\frac{1}{1+\sigma}>\mu>\frac{1+\theta}{1+2\theta}.$$

Proof. Equilibrium 4 exists whenever $(2\mu-1)R\geq\gamma$. Equilibrium 3 exists whenever equilibrium 4 exists since $1-\sigma\theta<1$. Equilibrium 1 exists whenever $\gamma>\left[\mu-\theta(2\mu-1)\right]R$. Hence, equilibria 1, 4, and 3 coexist if

$$(2\mu-1)R\geq\gamma>\left[\mu-\theta(2\mu-1)\right]R. \tag{1}$$

A little algebra shows that there exists a γ satisfying condition (1) whenever $\mu>\frac{1+\theta}{1+2\theta}$. Recall that equilibrium 2 exists if.

$$\left\{\mu-\theta\left(1-\sigma\right)\left(2\mu-1\right)\right\}R>\gamma\geq\left[1-\theta\left(1-\sigma\right)\right]\mu R.$$

It can be checked that $\left\{\mu-\theta\left(1-\sigma\right)\left(2\mu-1\right)\right\}R>(2\mu-1)R$ for all parameters. However, for $\left[\mu-\theta(2\mu-1)\right]R>\left[1-\theta\left(1-\sigma\right)\right]\mu R$ to hold, it must be the case that $\frac{1}{1+\sigma}>\mu$. With this restriction, the four equilibria coexist.
■

The condition $\mu > \frac{1+\theta}{1+2\theta}$ holds if μ and θ are sufficiently large. In particular, for any $\theta > 0$, there is a μ large enough that the condition hold. Note, however, that the condition cannot hold if $\mu < 2/3$.

3.4 Welfare under RTGS

Welfare is defined as the expected utility of a participant before the liquidity shock and time-criticality of the participant's payment is know. Equivalently, it is a weighted average of the welfare of all participants in the economy, where the weights are given by the population sizes.

First, we calculate the welfare of participants under equilibrium 1 of proposition 2, denoted by W_1, if such an equilibrium exists. With probability $1 - \theta$, a participant has to make a non-time-critical payment. In this case, the participant delays the payment at no cost. However, conditional on having to make a non-time-critical payment, the participant receives a negative liquidity shock with probability σ and incur an expected borrowing cost of $(1 - \theta)(1 - \mu)R$.

With probability θ, the participant has to make a time-critical payment. Under equilibrium 1, such payments are paid in the morning. Conditional on having to make a time-critical payment, a participant will receive a positive liquidity shock with probability σ and incur an expected borrowing cost of $(1 - \theta)(2\mu - 1)R$. A participant will receive a no liquidity shock with probability $1 - 2\sigma$ and incur an expected borrowing cost of $(1 - \theta)\mu R$. A participant will receive a negative liquidity shock with probability σ and incur a cost an expected borrowing cost of $(1 - \theta\mu)R$.

Recall that under equilibrium 1 of proposition 2, $\pi = \theta$. Putting these costs together, we obtain

$$\begin{aligned} W_1 &= -(1-\theta)\sigma(1-\theta)(1-\mu)R \\ &\quad -\theta\sigma(1-\theta)(2\mu-1)R \\ &\quad -\theta\left[1-2\sigma\right](1-\theta)\mu R \\ &\quad -\theta\sigma(1-\theta\mu)R. \end{aligned} \tag{2}$$

With similar steps, and a little algebra, we obtain the welfare of participants under equilibrium 2 of proposition 2, denoted by W_2, if such an equilibrium exists.[2] Under this equilibrium, $\pi = \theta(1-\sigma)$.

$$\begin{aligned} W_2 &= -(1-\theta)\sigma\left[1-\theta(1-\sigma)\right](1-\mu)R \\ &\quad -\theta\sigma\left[1-\theta(1-\sigma)\right](2\mu-1)R \\ &\quad -\theta\left[1-2\sigma\right]\left[1-\theta(1-\sigma)\right]\mu R \\ &\quad -\theta\sigma\left\{\gamma+\left[1-\theta(1-\sigma)\right](1-\mu)R\right\}. \end{aligned} \tag{3}$$

Under equilibrium 3 of proposition 2, $\pi = \sigma\theta$. The welfare of participants under this equilibrium, if it exists, is denoted by W_3 and given by the following expression.

$$\begin{aligned} W_3 &= -\theta(1-\sigma)\gamma-(1-\theta\sigma)\sigma(1-\mu)R \\ &\quad -\theta(1-\theta\sigma)\sigma(2\mu-1)R. \end{aligned} \tag{4}$$

Under equilibrium 4 of proposition 2, $\pi = 0$. The welfare of participants under this equilibrium, if it exists, is denoted by W_4 and given by the

[2] Details of the calculations for W_2, W_3, and W_4, are provided in the appendix.

following expression.

$$W_4 = -\theta\gamma - \sigma(1-\mu)R. \tag{5}$$

Proposition 3 $W_1 \geq W_2 \geq W_3 \geq W_4$ *whenever the corresponding equilibria exist.*

Proof. We show that, when comparing two equilibria, the equilibrium associated with the higher value of π yields higher welfare. Consider two equilibria, denoted by A and B, with π_A and π_B. Assume that $\pi_A > \pi_B$. Focus on a particular agent and let S_A and S_B denote the equilibrium strategies of this agent corresponding to each equilibrium and let $W(S_A, \pi_A)$ and $W(S_B, \pi_B)$ denote the welfare of this agent associated with each equilibrium. Now note that $W(S_B, \pi_A) \geq W(S_B, \pi_B)$, since all the actions that system participants can take have a cost that is (weakly) decreasing in π. Further, by definition of an equilibrium, $W(S_A, \pi_A) \geq W(S_B, \pi_A)$. If follows that $W(S_A, \pi_A) \geq W(S_B, \pi_B)$. ■

Another way to think about the result is in terms of the two sources of costs in this model: payment delay and borrowing from the CB. Bunching of payments, either in the morning or in the afternoon, reduces the cost of borrowing because payments can offset. Making payment in the morning, however, reduces the delay cost. If a participant decides to make her payment in the morning rather than in the afternoon, the effect is to reduce the offsets in the afternoon and increase the offsets in the morning. These two effects cancel each other out, at least partly. The benefit from reduced delay means that welfare is higher in an equilibrium in which more payments are paid in the morning.

4 A liquidity saving mechanism

In this section, we consider an arrangement that shares important features with liquidity saving mechanisms. This arrangement lets payments be made only if they are offset by an incoming payment. At the beginning of the morning period, after they observe their liquidity shock and the time criticality of their payment, core participants have the choice to put the payment they must make to another participant into a queue. The payment will be released if an offsetting payment is received by the participant or if an offsetting payment resides in the queue of another participant. Payments in the queue can offset multilaterally. We assume that the non-strategic agent does not use the queue.

The benefit from this arrangement is that it allows some participants to make sure that they will not incur a borrowing cost. The drawback, however, is that a payment put in the queue may not be released in the morning.

4.1 The case with no liquidity shock

The impact of a liquidity saving mechanism is easy to see in an economy with no liquidity shock. As we have seen above, without an LSM three equilibria can occur in this case: $\pi = 0$, $\pi = \theta$, and $\pi = 1$.

We contrast this result with what happens when a liquidity saving mechanism is available. First, note that participants have the choice between three actions: make a payment early, put a payment in the queue, or make a payment late. We use λ_e to denote the fraction of participants who send their payment early, λ_q to denote the fraction of agents who put their payments in the queue, and λ_d to denote the fraction of agents who delay their payments. Clearly, $\lambda_e + \lambda_q + \lambda_d = 1$.

When there are no liquidity shocks, putting a payment in the queue weakly dominates delaying a payment. Indeed, if the payment in the queue is not released in the morning, then both strategies yield the same outcome. If the payment in the queue is released in the morning, then the delay cost is avoided. Moreover, since participants do not need to borrow from the CB when the payment is released early, there is no borrowing cost with either strategy.

If delaying is a (weakly) dominated strategy, then either payments will be put in the queue or they will be paid in the morning outright. In either case, all payments are released in the morning and $\pi = 1$. Hence, the liquidity saving mechanism eliminates all equilibria with $\pi < 1$. With $\pi = 1$, all participants achieve the highest possible payoff of zero. This result is similar to Willison's (2005) result that an offsetting mechanism in the first period with no liquidity shocks results in the first-best outcome of all payments settling.

4.2 The case with liquidity shocks

With liquidity shocks ($\bar{\pi} > 0$ and $\mu < 1$), delaying payments is no longer dominated by the strategy of putting payments in the queue. Core participants who receive a negative liquidity shock may prefer to delay their payments since this can reduce the amount such participants need to borrow from the CB. With liquidity shock, understanding how the queue works is important.

4.2.1 The queue

In this section, we describe the way the queue works and derive the expressions for the probability that a participant receives a payment conditionally on being in the queue or not. The first thing to note is that the set off all payments must offset multilaterally. There may be one or more groups of payments that offset. We call any such group a cycle. At one extreme, the set of all payments could constitute the only cycle, as illustrated in Figure 1, so that any two participants are connected through a sequence of payments. At the other extreme, all cycles could be of length 2, as illustrated in Figure 2, so that all payments form pairs.

[Figures 1 and 2]

Turning to the queue, a payment in the queue may belong to a cycle having the property that all other payments in the cycle are also in the queue, as illustrated in Figure 3. In this case the payments are released by the queue since they offset multilaterally (or bilaterally if the cycle is of length 2). A payment in the queue may also be part of a cycle having the property that at least one payment in the cycle is not in the queue, as illustrated in Figure 4.[3] In this case, the payment belongs to a path (within the queue). Payments in a path cannot offset multilaterally. However, it is possible that

[3]Of course, a queue could contain both payments in a cycle and payments in path.

the participant who must make the "first" payment in the path receives a payment from outside the queue. In that case, the first payment in the path is released, creating a cascade of payments until eventually a payment is made to someone outside the queue. We denote by χ the probability that a payment in the queue is part of a cycle and $1-\chi$ the probability that it is part of a path.

[Figures 3 and 4]

We consider the value of χ for the two extreme cases described above. If all payments form only one cycle, then the probability that a payment in the queue is in a cycle is zero unless all participants put their payment in the queue. Formally, $\chi = 0$ if $\lambda_q < 1$ and $\chi = 1$ if $\lambda_q = 1$. Under this assumption, the queue releases the fewest payments. This case is also interesting because the role of the queue is only to allow agents to send their payment conditionally on receiving another payment. The queue no longer plays the role of settling multilaterally offsetting payments.

At the other extreme, if all payments are in cycles of length 2, then the probability that a payment in the queue is in a cycle is λ_q^2. Note that participants cannot take advantage of the fact that they know who they receive a payment from because they do not know whether that agent has received a liquidity shock or must make a time-sensitive payment. Moreover, since the probability that the same participants are paired again is zero, it is not possible to sustain dynamic incentives.

Next, we can derive the expressions for π^o, the probability of receiving a payment conditionally on not putting the payment in the queue, and π^q, the probability of receiving a payment conditionally on putting a payment in the queue. The latter probability is equivalent to the probability that a payment in the queue is released.

Suppose that there are no payments in the queue. Then, the probability of receiving a payment is given by the mass of participants who send a payment outright divided by the total mass of participants. Formally, $\pi^o = \lambda_e/(\lambda_e + \lambda_d)$. It turns out that the expression for π^o does not change when there are payments in the queue. Indeed, note that every payment made early by some participant outside the queue to a participant inside the queue must ultimately trigger a payment from a participant is inside the queue to a participant whose payment is outside the queue. From the perspective of participants outside the queue, this is the same as if the payment had been made directly from a participant outside the queue to another participant outside the queue. For this reason, we can ignore the queue. In summary, the expression for π^o is

$$\pi^o \equiv \frac{\lambda_e}{\lambda_e + \lambda_d} = \frac{\lambda_e}{1 - \lambda_q}. \tag{6}$$

If a participant puts a payment in the queue, the payment will be in a cycle with probability χ, in which case it is released for sure. With probability $1-\chi$, the payment is in a path. The probability that a payment in a path is released is equal to the probability of receiving a payment from outside the queue. This probability is equal to π^o. So the expression for π^q is given by

$$\pi^q \equiv \chi + (1-\chi)\frac{\lambda_e}{\lambda_e + \lambda_d} = \chi + (1-\chi)\pi^o. \tag{7}$$

Under our "long-cycle" assumption, $\chi = 0$ if $\lambda_q < 1$ so that $\pi^o = \pi^q = \lambda_e/(\lambda_e + \lambda_d)$. If $\lambda_q = 1$, then $\pi^o = 0$ and $\pi^q = 1$, since all the payment are put in the queue. Under our "short-cycles" assumption, $\chi = \lambda_q^2$ so that

$$\pi^q = \lambda_q^2 + (1 - \lambda_q^2)\frac{\lambda_e}{\lambda_e + \lambda_d} = \lambda_q^2 + (1 - \lambda_q^2)\pi^o.$$

4.2.2 Participants' behavior

Now we turn to describing the behavior of the participants. There are six types to consider. Participants who must send a time-critical payment may have a negative, a positive, or no liquidity shock. Similarly for participants who must send a non-time-critical payment. We first consider participants who must send time-critical payments.

For participants who must make a time-sensitive payment and have received no liquidity shock, the cost of delay is γ, since they do not need to borrow from the CB, whether or not they receive a payment early. The expected cost of putting a payment in the queue is $(1 - \pi^q)\gamma$. Since $\pi^q \geq 0$, the cost of delay is always at least a large as the cost of putting the payment in the queue. The expected cost of sending the payment early is $(1 - \pi^o)\mu R$, since with probability $1 - \pi^o$ no offsetting payment is received and μ must be borrowed at the CB. Hence, participants who must make a time-sensitive payment and have received no liquidity shock make their payment early if

$$(1 - \pi^q)\gamma \geq (1 - \pi^o)\mu R, \tag{8}$$

and put their payment in the queue otherwise.

Next, we consider the decision of participants who must make a time-sensitive payment and have received a positive liquidity shock. For these participants, the expected cost of delaying a payment or putting the payment in the queue is the same as for participants who did not receive a liquidity shock. Hence, for these participants also, it is always better to put a payment in the queue rather than to delay. The expected cost of sending the payment early is different, however, as the liquidity shock reduces the amount the participant needs to borrow whenever no offsetting payment is received. The cost of sending the payment early is given by $(1 - \pi^o)(2\mu - 1)R$. Hence, participants who must make a time-sensitive payment and have received a positive liquidity shock make their payment early if

$$(1 - \pi^q)\gamma \geq (1 - \pi^o)(2\mu - 1)R, \tag{9}$$

and put their payment in the queue otherwise.

Finally, we consider the decision of participants who must make a time-critical payment and have received a negative liquidity shock. To find the expected cost of delay for such participants, note that with probability $1-\pi^o$, no payment is received and these participants must borrow $1-\mu$ to cover the liquidity shock. With probability π^o, a payment is received and since $\mu > 1 - \mu$ the participants do not need to borrow. These participants also suffer the delay cost γ, so the expected cost of delay is given by $\gamma + (1-\pi^o)(1-\mu)R$.

If the payment is put in the queue, the participants must borrow $1-\mu$ from the CB whether the payment is released or not, since receiving a payment triggers the release of the queued payment. If the payment is not released, which happens with probability $1-\pi^q$, then the delay cost must be added. So the expected cost of putting the payment in the queue is $(1-\pi^q)\gamma+(1-\mu)R$.

If the payment is sent outright, the delay cost is always avoided. With probability π^o, a payment is received early and only $1-\mu$ must be borrowed from the CB. Otherwise, 1 must be borrowed. So the expected cost of sending the payment outright is $\left[(1-\pi^o)+\pi^o(1-\mu)\right]R$.

Using a little algebra, the behavior of participants who must make a time-critical payment and have received a negative liquidity shock can be characterized as follow: If

$$(1-\pi^q)\gamma \geq (1-\pi^o)\mu R, \tag{10}$$

then these participants choose to send the payment early. If

$$(1-\pi^o)\mu R > (1-\pi^q)\gamma \quad \text{and} \quad (1-\pi^q)\gamma \geq (1-\pi^o)(1-\mu)R \tag{11}$$

then these participants choose to put the payment in the queue. Finally, if

$$\pi^o(1-\mu)R > \pi^q\gamma, \tag{12}$$

then these participants choose to delay their payment.

Next we consider participants who must make a non-time-critical payment. By setting $\gamma = 0$ in the analysis conducted above, we can see that participants who face no cost of delay always prefer to put a payment in the queue rather than send it early. We have also seen that participants who receive a positive or no liquidity shock do not need to borrow from the CB if they queue or delay their payment. Hence, participants who must make a non-time-critical payment and who receive either a positive or no liquidity shock are indifferent between delaying or queuing their payment since neither action imposes a cost on them. In this case, we assume that the participants queue their payments

This reasoning does not apply to participants who have received a negative liquidity shock. For these agents, the cost of delay is $(1 - \pi^o)(1 - \mu)R$ since they must borrow from the CB unless they delay their payment and they receive a payment early. The cost of putting a payment in the queue is $(1 - \mu)R$, since in that case they must borrow from the CB whether or not they receive a payment. Hence, if $\lambda_e > 0$, which implies $\pi^o > 0$, these agents prefer to delay. If $\lambda_e = 0$, then the probability that these participants receive a payment from outside the queue is zero, or $\pi^o = 0$. In that case, they are indifferent between delaying and and queuing their payment. Again, we assume that these participants queue their payment if they are indifferent.

4.2.3 Equilibrium under the long-cycle assumption

Under the long-cycle assumption, the value of χ can be either 0 or 1, depending on whether λ_q is strictly smaller or is equal to 1. We consider each case separately.

If $\lambda_q = 1$, then all payments are put in the queue and are released in the morning. Participants who receive a liquidity shock must borrow $1 - \mu$ from the central bank but, since $\lambda_e = 0$, these agents cannot avoid that cost. Since no delay cost is incurred it is an equilibrium for all participants to put their payment in the queue regardless of that cost. However, we will see below that if γ is high enough, then this equilibrium does not survive deletion of weakly dominated strategies.

Since $\lambda_q < 1$ is an equilibrium only if $\lambda_e > 0$, we need to find out the parameter values for which this condition can hold. The conditions for participants to prefer to pay early rather than queue are given by equations (8), (9), and (10). Since $\mu \geq 2\mu - 1$, we only need to consider equation (9). The condition $\lambda_q < 1$, implies $\chi = 0$ and $\pi^o = \pi^q$. It follows that if

$$\gamma < (2\mu - 1)R, \tag{13}$$

then $\lambda_e > 0$ cannot be an equilibrium and the unique equilibrium is such that $\lambda_q = 1$.

We now assume that $\gamma \geq (2\mu - 1)R$ and consider equilibria such that $\lambda_e > 0$. If $\mu \geq 2/3$, then $2\mu - 1 \geq 1 - \mu$. This guarantees that participants who have a negative liquidity shock and must make a time critical payment do not want to delay their payment. If $\mu R > \gamma \geq (2\mu - 1)R$, then only participants with a positive liquidity shock make time-critical payment early. In this case, $\lambda_e = \theta\sigma$, $\lambda_q = 1 - \sigma$, $\lambda_d = (1 - \theta)\sigma$, and $\pi^o = \pi^q = \theta$.[4] If $\gamma \geq \mu R$, then all time-critical payments are made early. In this case, $\lambda_e = \theta$, $\lambda_q = (1 - \theta)(1 - \sigma)$, $\lambda_d = (1 - \theta)\sigma$, and $\pi^o = \pi^q = \theta / (\theta + (1 - \theta)\sigma)$.

If $\mu < 2/3$, then $1 - \mu > 2\mu - 1$. For such parameters, we need to consider three cases: Either $(1 - \mu)R > \gamma \geq (2\mu - 1)R$, or $\mu R > \gamma \geq (1 - \mu)R$, or $\gamma \geq \mu R$. If $(1 - \mu)R > \gamma \geq (2\mu - 1)R$, then participants who receive a negative liquidity shock delay time-sensitive payments. In this case, $\lambda_e = \theta\sigma$, $\lambda_q = 1 - \sigma(1 + \theta)$, $\lambda_d = \sigma$, and $\pi^o = \pi^q = \theta/(1 + \theta)$. The other two cases are identical to the two cases studies above when $\mu \geq 2/3$.

[4] Recall that whenever $\lambda_e > 0$ participants who have a negative liquidity shock delay non-time-critical payments while other non-time-critical payments are put in the queue.

We have shown that there can be multiple equilibria in some regions of the parameter space. However, with the appropriate refinement there is a unique equilibrium as is shown in the following lemma.

Lemma 2 *If both the equilibrium with $\lambda_q < 1$ and the equilibrium with $\lambda_q = 1$ exist, then the equilibrium with $\lambda_q = 1$ does not survive the deletion of weakly dominated strategies.*

Proof. We need to show that for some participants, putting their payment in the queue is a weakly dominated strategy when both equilibria exist. Consider participants who receive a negative liquidity shock and must make a non-time-critical payment. These participants are indifferent between delaying or putting their payment in the queue if $\lambda_q = 1$. However, they strictly prefer to delay their payment if $\lambda_q < 1$. Hence, the strategy consisting of putting the payment in the queue is weakly dominated for these agents whenever both equilibria exist. ■

This result is in contrast to section 3.3, where we found multiple equilibria that are robust. In the remainder of this paper, we focus on the equilibrium with $\lambda_q < 1$ when it exists. The results of this section can be summarized in the following proposition

Proposition 4 *Under the long-cycle assumption, we have the following equilibria:*

1. *If $\gamma < (2\mu - 1)R$, then all participants put their payment in the queue*

2. *If $\gamma \geq (2\mu - 1)R$ and $\mu \geq 2/3$, then*

(a) *If* $\mu R > \gamma \geq (2\mu - 1)R$*, then only participants with a positive liquidity shock make time-critical payment early. Participants with a negative liquidity shock delay non-time-critical payments and all others put their payment in the queue.*

(b) *If* $\gamma \geq \mu R$*, then all time-critical payments are made early. Participants with a negative liquidity shock delay non-time-critical payments and other non-time-critical payments are put in the queue.*

3. *If* $\gamma \geq (2\mu - 1)R$ *and* $\mu < 2/3$*, then*

(a) *If* $(1 - \mu)R > \gamma \geq (2\mu - 1)R$*, then participants who receive a negative liquidity shock delay their payment. Participants who receive a positive liquidity shock send their time-critical payment early. All other payments are put in the queue.*

(b) *If* $\mu R > \gamma \geq (1 - \mu)R$ *, then only participants with a positive liquidity shock make time-critical payments early. Participants with a negative liquidity shock delay non-time-critical payments and all others put their payment in the queue.*

(c) *If* $\gamma \geq \mu R$*, then all time-critical payments are made early. Participants with a negative liquidity shock delay non-time-critical payments and other non-time-critical payments are put in the queue.*

4.2.4 Equilibrium under the short-cycles assumption

Under the short-cycles assumption, $\chi = \lambda_q^2$. The analysis in this section is similar to the analysis in the previous section. For the same reason, it is an equilibrium for all participants to put their payments in the queue. This equilibrium is unique if γ is sufficiently small. If γ is not so small, there exists an equilibrium where some agents send their payments early while some agents delay.

We want to characterize the equilibria having the property that some payments are not put in the queue. First, consider the case where $\mu \geq 2/3$. Since $(1-\pi^o)/(1-\pi^q) \geq \pi^o/\pi^q$, it must be the case that

$$\frac{1-\pi^o}{1-\pi^q}(2\mu - 1)R \geq \frac{\pi^o}{\pi^q}(1-\mu)R.$$

This implies that equation (12) holds whenever equation (9) holds. Hence, if some participants with a positive liquidity shock choose to send time-critical payments outright, participants with a negative liquidity shock prefer to put time-critical payments in the queue rather than delay. In this case, only participants who have a negative liquidity shock and must make a non-time-critical payment delay. Formally, $\lambda_d = (1-\theta)\sigma$. We also know that other participants will put non-time-critical payments in the queue.

We need to determine what participants do with time-critical payments. We have seen that participants with a positive liquidity shock send payments early if $\gamma \geq \frac{1-\pi^o}{1-\pi^q}(2\mu-1)R$ and other participants send payments early if $\gamma \geq \frac{1-\pi^o}{1-\pi^q}\mu R$. Payments are put in the queue otherwise. So we need to focus on the ratio $(1-\pi^o)/(1-\pi^q)$.

Recall, from equation (7), that $\pi^q = \chi + (1-\chi)\pi^o$, which implies

$$1 - \pi^q = 1 - \chi - (1-\chi)\pi^o = (1-\chi)(1-\pi^o),$$

Thus, we can write

$$\frac{1-\pi^o}{1-\pi^q} = \frac{1}{1-\chi} = \frac{1}{1-(1-\lambda_e-\lambda_d)^2}. \tag{14}$$

Since λ_d is known, we need to find the value of λ_e. We focus on pure strategies and show that λ_e can take three values: Either no participants make time-critical payments early, or only participants with a positive liquidity shock make time-critical payments early, or all participants make time-critical payments early.[5]

If no participants make time-critical payments early, $\lambda_e = 0$, then

$$\frac{1-\pi^o}{1-\pi^q} = \frac{1}{1-[1-(1-\theta)\sigma]^2}. \tag{15}$$

If only participants with a positive liquidity shock make such payments early, $\lambda_e = \sigma\theta$, then

$$\frac{1-\pi^o}{1-\pi^q} = \frac{1}{1-(1-\sigma)^2}. \tag{16}$$

If all participants make time-critical payments early, $\lambda_e = \theta$, then

$$\frac{1-\pi^o}{1-\pi^q} = \frac{1}{1-[(1-\theta)(1-\sigma)]^2}. \tag{17}$$

It can be verified that

$$\frac{1}{1-[1-(1-\theta)\sigma]^2} > \frac{1}{1-(1-\sigma)^2} > \frac{1}{1-[(1-\theta)(1-\sigma)]^2}. \tag{18}$$

We can now characterize the equilibria of this economy.

Proposition 5 *The following cases can arise:*

[5] As is standard when there are multiple equilibria in pure strategies, there also exists equilibria in mixed strategies. We ignore such equilibria here.

1. *If* $\gamma \leq \min\left\{\left[1-[(1-\theta)(1-\sigma)]^2\right]^{-1}\mu R; \left[1-(1-\sigma)^2\right]^{-1}(2\mu-1)R\right\}$, *then the condition for participants to send time-critical payments early cannot be satisfied. Hence, the only equilibrium is for all participants to put their payments in the queue.*

2. *If* $\left[1-[(1-\theta)(1-\sigma)]^2\right]^{-1}\mu R > \gamma > \left[1-(1-\sigma)^2\right]^{-1}(2\mu-1)R$, *then it is an equilibrium for only participants with a positive liquidity shock to make time-critical payments early but it is not an equilibrium for all time-critical payments to be made early.*

3. *If* $\left[1-(1-\sigma)^2\right]^{-1}(2\mu-1)R > \gamma > \left[1-[(1-\theta)(1-\sigma)]^2\right]^{-1}\mu R$, *then it is an equilibrium for all participants to make time-critical payments early, but it is not an equilibrium for only participants with a positive liquidity shock to make time-critical payments early.*

4. *If*

$$\frac{\mu R}{1-(1-\sigma)^2} > \gamma \geq \max\left\{\frac{\mu R}{1-[(1-\theta)(1-\sigma)]^2}; \frac{(2\mu-1)R}{1-[1-\sigma]^2}\right\},$$

then it is an equilibrium for all participants to make time-critical payments early and it is also an equilibrium for only participant with a positive liquidity shock to make time-critical payments early.

5. *If* $\gamma \geq \frac{\mu R}{1-(1-\sigma)^2}$ *then it is an equilibrium for all participants to make time-critical payments early.*

6. *If* $\gamma \leq \left\{1-[1-(1-\theta)\sigma]^2\right\}^{-1}(2\mu-1)R$, *then it is an equilibrium for no time-critical payments to be made early.*

These expressions are obtained by substituting the values of $(1-\pi^o)/(1-\pi^q)$ given by equations (15), (16), and (17) into equations (8), (9), (10), and (11).

In case 2, the decrease in $(1-\pi^o)/(1-\pi^q)$ when more participants pay early is not large enough to provide incentives for participants who have not received a positive liquidity shock to pay early. Case 2 can arise if $\mu < 1$ and θ is sufficiently small, for example.

In case 3, $(1-\pi^o)/(1-\pi^q)$ is large enough that the cost of delay does not provide incentives for only participants with a positive liquidity shock to pay early. However, if all participants make time-critical payments early, $(1-\pi^o)/(1-\pi^q)$ decreases so much that this strategy is consistent with an equilibrium. Case 3 can arise if μ is sufficiently close to 1, for example.

In case 5, participants who have a positive liquidity shock would have incentives to make time-critical payments early even if other participants did not. However, the cost of delay is so high that other participants also find it beneficial to make such payments early.

It should be noted that if the equilibrium such that all participants put their payment in the queue is one of multiple equilibria, this equilibrium can be robust. Proposition 2 applies whenever all participants who must make time-sensitive payments would choose to put such payments in the queue only if no participants delay their payment. In this section, we have seen that for some parameter values all participants who must make time-critical payments would choose to put their payments in the queue even if some payments were delayed. In such a case, the equilibrium in which all participants put their payment in the queue is robust.

We now turn to the case where $\mu < 2/3$. For such values of μ, it is possible that

$$\frac{\pi^o}{\pi^q}(1-\mu)R > \gamma > \frac{1-\pi^o}{1-\pi^q}(2\mu-1)R,$$

in which case all participants who receive a negative liquidity shock have an incentive to delay. If these inequalities do not hold, the analysis is the same as above. Note also that

$$\frac{1-\pi^o}{1-\pi^q}\mu R \geq \frac{\pi^o}{\pi^q}(1-\mu)R.$$

Hence, if an equilibrium exists such that participants with a negative liquidity shock delay time-critical payments and participants with a positive liquidity shock make time-critical payments early, then it must be the case that participant with no liquidity shock queue their payment. In that case we have $\lambda_e = \sigma\theta$.

Restricting our attention to equilibria in pure strategies, λ_d can take two values. If only participants who must make a non-time critical payment and receive a negative productivity shock delay their payment, then $\lambda_d = \sigma(1-\theta)$ and $\pi^o/\pi^q = \theta/\left[\theta + (1-\theta)(1-\sigma)^2\right]$, where the expressions for π^o and π^q are given by equations (6) and (7). This case was studied in the previous section. If all participants who receive a negative liquidity shock delay their payment, then $\lambda_d = \sigma$ and $\pi^o/\pi^q = \theta/\left[\theta + (1-(1+\theta)\sigma)^2\right]$. This case is summarized in the following proposition.

Proposition 6 *If*

$$\gamma \leq \frac{1}{1-\left[1-\sigma(1+\theta)\right]^2}(2\mu-1)R,$$

then the unique equilibrium is for all agents to put their payment in the queue. If

$$\frac{\theta}{\theta + (1-(1+\theta)\sigma)^2}(1-\mu)R > \gamma > \frac{1}{1-[1-\sigma(1+\theta)]^2}(2\mu-1)R,$$

then it is an equilibrium for participants who receive a positive liquidity shock to make their time-critical payment early, for participants who receive a negative liquidity shock to delay all payments, and for other participants to put their payment in the queue.

4.2.5 Welfare under the long-cycle assumption

There is no systematic relationship between the equilibria described in proposition 4. To illustrate this, we show that depending on parameter values equilibrium $2a$ can provide more or less welfare than equilibrium $2b$. We use $W_{4,2a}$ and $W_{4,2b}$ to denote welfare under the corresponding equilibria in proposition 4. Note that

$$\pi_{4,2b} - \pi_{4,2a} = \frac{\sigma\theta(1-\theta)}{1-\sigma(1-\theta)} \geq 0.$$

Also, equilibrium $2b$ of proposition 4 exists if $\gamma = \mu R - \varepsilon$, where $\varepsilon \in (0, (1-\mu)R)$. Hence, we can write

$$\begin{aligned} W_{4,2b} - W_{4,2a} &= (1-\theta)\sigma\left(\pi_{4,2b} - \pi_{4,2a}\right)(1-\mu)R \\ &\quad + \theta\sigma\left(\pi_{4,2b} - \pi_{4,2a}\right)(2\mu-1)R \\ &\quad + \theta\left[1-2\sigma\right]\left[\left(\pi_{4,2b} - \pi_{4,2a}\right)\mu R - (1-\pi_{4,2a})\varepsilon\right] \\ &\quad + \theta\sigma\left[\left(\pi_{4,2b} - \pi_{4,2a}\right)\mu R - (1-\pi_{4,2a})\varepsilon\right]. \end{aligned}$$

Given $\pi_{4,2b} - \pi_{4,2a} > 0$, we can choose $\varepsilon > 0$ so small that $W_{4,2a} - W_{4,2b} > 0$. Now set $\bar{\pi} = 1/2$, so that $\sigma = 1/4$, $\mu = 3/4$, and $\varepsilon = (1 - \mu)R = R/4$. With these parameters, we have

$$W_{4,2b} - W_{4,2a} = \frac{1-\theta}{3-\theta}\frac{\theta}{16}\left[13\theta - 8\right],$$

so that that $W_{4,2b} - W_{4,2a} < 0$ if $\theta < 8/13$.

Next, we would like to compare equilibria with and without a LSM. To prove that an equilibrium with LSM yields higher welfare than an equilibrium without LSM, it is enough to show that the value of $\pi^o = \pi^q$ corresponding to the equilibrium with LSM is higher that the value of π corresponding to the equilibrium without an LSM. The argument is the same as in the proof of proposition 3.

If the size of the liquidity shock, $1 - \mu$, is small enough, then a LSM achieves higher welfare.

Proposition 7 *If $\mu \geq 2/3$, then welfare is higher with a LSM that without.*

Proof. As we have seen in section 3.3, the largest value of π in an equilibrium without LSM is θ. In contrast, we have seen in section 4.2.3 that the lowest value of $\pi^o = \pi^q$ is θ, if $\mu \geq 2/3$. Hence, welfare is always at least as high with a LSM that without under our long-cycle assumption. ■

If the size of the liquidity shock is small, then participants who receive a negative liquidity shock always prefer to put their payment in the queue rather than delay. This increases the probability of receiving a payment and makes it always greater the the probability of receiving a payment without a LSM.

Now we show that if the delay-cost of time-critical payments is sufficiently high, or sufficiently low, a LSM helps achieve higher welfare.

Proposition 8 *If $\gamma > [\mu - \theta(2\mu - 1)]\, R$ or if $\gamma < (2\mu - 1)R$ then the level of welfare is at least as high with a LSM than without.*

Proof. If $\gamma > [\mu - \theta(2(\mu - 1)]\, R$, then equilibrium 1 of proposition 2 exists when there is no LSM. We already know, from proposition 3, that this equilibrium provides the highest welfare of all equilibria without a LSM. So we need to prove that LSM equilibria consistent with $\gamma > [\mu - \theta(2\mu - 1)]\, R$ achieve higher welfare than equilibrium 1 of proposition 2. To do that, it is enough to show that the value of $\pi^o = \pi^q$ corresponding to these LSM equilibria are greater or equal than θ.

Note that $[\mu - \theta(2(\mu - 1)] > 1 - \mu$. So we need to check the value of π for equilibria $2a$, $2b$, $3b$, and $3c$ of proposition 4. Simple algebra reveals that the values of $\pi^o = \pi^q$ corresponding to these equilibria are greater or equal than θ.

If $\gamma < (2\mu - 1)R$, all participants put their payments in the queue with a LSM, so that $\pi^q = 1 \geq \theta$. ■

If the cost of delay is sufficiently high, participants who must make time-critical payments have strong incentives not to delay their payments, whether or nor a LSM is available. Adding a LSM increases welfare because it allows participants who must make non-time-critical payments to make payment conditional on receiving a payment. This protects them from the risk of having to borrow from the CB. If the cost of delay is sufficiently low, no participant has an incentive to make a payment early. When this occurs, no participant has an incentive to delay and all payments are put in the queue. This allows all payments to be released from the queue. Below we provide an example showing that if the cost of delay is intermediate, it is possible that welfare is higher without a LSM.

A LSM will also yield higher welfare if the fraction of time critical payments is not too high.

Proposition 9 *If $\sigma/(1-\sigma) > \theta$, then higher welfare is achieved with a LSM than without.*

Proof. It is enough to show that under this condition the highest π that can be achieved without a LSM is lower than the lowest π^o that can be achieved with an LSM. The highest π that can be achieved without a LSM is θ, under equilibrium 1 of proposition 2. The lowest π^o that can be achieved with an LSM is $\theta/(1+\theta)$ under items $3a$ in proposition 4. $\pi^o > \pi \Leftrightarrow \sigma/(1-\sigma) > \theta$. ■

If the fraction of time critical payments is small, the benefit from putting a large fraction of non-time-critical payment in the queue is large and an LSM provides higher welfare.

We can show by an example that for some parameter values the welfare under an equilibrium without an LSM can be higher the the welfare with an LSM.

Example 1 *Let $\mu = 1/2 + \delta$, $\gamma = \mu R - \varepsilon$, where $\delta, \varepsilon > 0$. For these parameters, welfare with a LSM is smaller than without a LSM, if δ and ε*

are small enough.

Details are provided in the appendix.

If γ is not too large participants with a negative liquidity shock who must make time-critical payment may choose to delay them. At the same time, γ must not be too small, so that participants with a positive liquidity shock prefer to send their payment rather than put it in the queue. In this case, the queue participants who have not received a liquidity shock and must make a time-critical payment have the option not to make the payment early, but to put it in the queue instead. This is beneficial for these agents, because it protects them from having to borrow from the CB. However, from the perspective of society, the decrease in the fraction of participants who make their payment early decreases welfare. In other words, the queue can eliminate some beneficial coordination between agents by giving them the option to send a payment conditionally on receiving once. This reduction in coordination has a cost for society.

4.2.6 Welfare under the short-cycles assumption

Propositions 7 and 9 carry over to the case with short-cycles without modification. Proposition 8 is slightly modified. We restate it without proof.

Proposition 10 *If* $\gamma > [\mu - \theta(2\mu - 1)]\, R$ *or if*

$$\gamma < \left\{1 - \left[1 - \sigma(1 + theta)^2\right]\right\}^{-1} (2\mu - 1)R$$

then the level of welfare is at least as high with an LSM than without.

Now we provide one example showing that for some parameters, welfare can be lower with an LSM.

Example 2 *Assume $\theta = 1$, $\sigma = 1/4$, and $\mu = 1/2$. If $\gamma/R \in (3/8, 2/5)$, the unique robust equilibrium with an LSM offers lower welfare than the unique equilibrium without an LSM.*

Details are provided in the appendix.

5 Balance vs. receipt reactive mechanisms

In this section we compare two different and common types of LSMs. A balance-reactive LSM (BRLSM) is a mechanism in which a participant can specify a level of its balances below which no payment is sent from the queue. A receipt-reactive LSM (RRLSM) is a mechanism in which payments received during a given interval of time are used to offset and release payments in the queue, regardless of the level of the participant's balance. In order to compare a BRLSM and a RRLSM, we change the timing at which participants must make the decision to send, delay, or put a payment in the queue. In this section, we assume that participants make this decision before they learn their liquidity shock. All other aspects of the environment are unchanged. In particular, participants know whether their payment is time critical or not before they decide to delay, queue, or send it.

Under a BRLSM, participants can condition the release of a queued payment on their level of balance. This is equivalent, in our environment, to conditioning the release of a payment on the participant's liquidity shock.

As a consequence, BRLSM equilibria in this section are equivalent to the equilibria described in section 4. The remainder of the section focuses on the RRLSM.

Under a RRLSM, offsetting payments are used to trigger the release of payments in the queue regardless of the participant's balance. In our model, this is equivalent to participants making their choice to queue a payment before knowing their liquidity shock. At that time, the expected welfare from any decision is the same for all participants whose must make payments with the same time-criticality. The cost of any decision corresponds to the weighted average of the costs described in section 4 with the weights given by the probability of receiving a particular liquidity shock.

5.1 Participants' behavior under a receipt reactive LSM

If a payment is delayed, all participants must pay the delay-cost for time-critical payments, regardless of their liquidity shock. With probability σ, a participant receives a negative liquidity shock and must also borrow an amount $1 - \mu$ if an offsetting payment is not received. The expected cost of delay is thus given by

$$\gamma + \sigma(1 - \pi^o)(1 - \mu)R. \tag{19}$$

If a payment is queued, it is not released with probability $1 - \pi^q$, in which case the delay cost is incurred for time-critical payments. In addition, participants who receive a negative liquidity shock must borrow $1-\mu$. Hence, the expected cost of putting a payment in the queue is

$$\sigma\left[(1-\pi^q)\gamma + (1-\mu)R\right] + (1-\sigma)\,(1-\pi^q)\gamma. \tag{20}$$

If a payment is sent early, then participants who receive a negative liquidity shock must borrow 1 if they do not receive a payment in the morning. This occurs with probability $1-\pi^o$. Otherwise they must borrow $1-\mu$. Participants who receive no liquidity shock must borrow μ if they do not receive a payment in the morning. Finally, participants who receive a positive liquidity shock must borrow $(2\mu-1)$ if they do not receive a payment in the morning. It follows that the expected cost of sending payments in the morning is

$$\sigma\left[(1-\pi^o) + \pi^o(1-\mu)\right]R + (1-2\sigma)\,(1-\pi^o)\mu R + \sigma(1-\pi^o)(2\mu-1)R. \tag{21}$$

The behavior of participants under a RRLSM is described in the next proposition

Proposition 11 *Assume $\bar{\pi} \in (0,1)$ and $\mu \in [0.5, 1)$, Under a receipt reactive LSM, participants who must make a time-critical payment*

1. *delay the payment if $\pi^o\sigma(1-\mu)R > \pi^q\gamma$,*
2. *queue the payment if $(1-\pi^o)\left[\mu R - \sigma(1-\mu)R\right] > (1-\pi^q)\gamma$ and $\pi^q\gamma \geq \pi^o\sigma(1-\mu)R$,*
3. *make the payment early if $(1-\pi^q)\gamma > (1-\pi^o)\left[\mu R - \sigma(1-\mu)R\right]$.*

Participants who must make a non-time-critical payment delay unless time-critical payments are queued, in which case they are indifferent between delaying and putting the payment in the queue.

Proof. The boundaries for delaying, queuing, or sending payments in the morning come from comparing equations (19), (20), and (21). Since $(\pi^o/\pi^q)\sigma(1-\mu)R > 0$, participants who must send non-time-critical payments prefer to delay unless $\pi^o = 0$, which occurs if time-critical payments are put in the queue. ■

5.2 Equilibria under a receipt reactive LSM

We can now describe the equilibria. With a RRLSM, either all participants put their payment in the queue or time-critical payments are paid early and non-time-critical payments are delayed.

Proposition 12 *There are two equilibria. Either all participants put their payment in the queue, in which case $\pi^o = 0$ and $\pi^q = 1$, or time-critical payments are paid early and non-time-critical payments are delayed, in which case $\pi^o = \pi^q = \theta$. If $\gamma > [\mu - \sigma(1-\mu)]$, only the latter equilibrium survives deletion of weakly dominated strategies.*

Proof. First, we show that it cannot be an equilibrium strategy for time-critical payments to be delayed. Indeed, if such payments are delayed, then no payment is made early and $\pi^o = 0$. Participants who must make non-time-critical payments are indifferent between delaying and putting payments in the queue and we assume that they queue payments so that $\pi^q > 0$. In that case, the condition $\pi^o \sigma(1-\mu)R > \pi^q \gamma$ indicates that for any $\gamma > 0$, time-critical payments are not delayed.

If time-critical payments are put in the queue, participants who must make non-time-critical payments are again indifferent between delaying and queueing. In this case, $\pi^o = 0$ and $\pi^q = 1$. This implies that the conditions for all participants to put payments in the queue are satisfied for all γ.

Finally, if time-critical payments are made early, then participants who must make non-time-critical payments strictly prefer to delay. In this case, $\lambda_e = \theta$, $\lambda_d = 1-\theta$, and $\lambda_q = 0$, which implies $\pi^o = \pi^q = \theta$. This equilibrium exists only if $\gamma > \mu R - \sigma(1-\mu)R$.

Lemma 2 applies in this case so that if $\gamma > \mu R - \sigma(1-\mu)R$, the equilibrium such that all participants queue their payment is not robust. ■

5.3 Welfare under a receipt reactive LSM

If the cost of delay is sufficiently high, we can show that a BRLSM provides higher welfare than a RRLSM.

Proposition 13 *If* $\gamma \geq \max\left\{[\mu - \sigma(1-\mu)]\, R; \frac{\theta}{\theta+(1-\theta)(1-\sigma)^2}(1-\mu)R\right\}$, *then welfare under a balance reactive LSM is higher than welfare under a receipt reactive LSM.*

Proof. To prove this result, we need to show that the values of π^o and π^q are higher under the RRLSM than under the BRLSM. Since $\gamma > [\mu - \sigma(1-\mu)]\, R$ the unique robust equilibrium with a RRLSM has $\pi^o = \pi^q = \theta$. The value of π^o and π^q will be at least as large with a BRLSM if time-critical payments are not delayed. The corresponding condition is $\gamma \geq \frac{\theta}{\theta+(1-\theta)(1-\sigma)^2}(1-\mu)R$.

$\blacksquare$

The next examples show that when the cost of delay is not too high and θ not too small, welfare under a RRLSM can be higher than under a BRLSM.

Example 3 *If $\gamma \leq [\mu - \sigma(1-\mu)]$ and μ is sufficiently small, then welfare under a RRLSM is higher than under a BRLSM.*

Details are provided in the appendix.

Finally, we compare welfare between a RRLSM and a system without LSM.

Proposition 14 *A system with a RRLSM provides at least as much welfare as a system without LSM.*

Proof. We have seen that the probability of receiving a payment early can be no greater than θ in the absence of a LSM. With a RRLSM, the probability of receiving a payment early is no smaller than θ. $\blacksquare$

6 Conclusion

We have shown that LSMs are potentially welfare-improving relative to an RTGS system alone. This conclusion was reached using a model in which banks settle their daily payments while seeking to minimize the costs of payment delays and intraday borrowing. The novel feature of the model we employ is that the participants are subject to two types of shocks. First, banks are randomly assigned to have either time-critical payments, whose late-period settlement imposes a cost on the bank, or non-time-critical payments. Second, banks are subject to liquidity shocks at the start of the day because of the operation of a settlement system, the timing of whose payments the banks don't control. These two shocks yield a rich array of strategic situations. The important parameters in our model are the costs of delay, the borrowing cost of intraday funds from the central bank, the relative size of the payments made to the settlement system versus other payments, and proportion of payments that are time critical.

We first show the importance of liqudity shocks in our model by comparing the equilibria of an RTGS system with and without the shocks. The liquidity shocks eliminate the "all pay early" equilibrium that exists without the shocks. The reason for this result is that, with liquidity shocks, the participants with a negative shock and non-time-sensitive payments will not choose to pay early (to avoid the borrowing costs they would otherwise incur). This simple intuition is important becuase it reminds us that coordination of parties that have different types of payments and who are subject to liquidity shocks is more difficult than if we assume all parties are identical. The equilibria under RTGS depend on the parameters in complicated ways, and, in certain parts of the parameter space, multiple equilibria exist.

Our model of LSMs considers two alternative designs of LSM. The first, a balance-reactive LSM, allows participants to queue payments to be settled from the queue conditional on the level of their balance being above some threshold level. The second, a receipt-reactive LSM, does not allow a participant to condition the settlement of its queued payment on the level of its balance. Surprisingly, these alternative designs lead to significant differences in behavior and performance.

To model the working of a balance-reactive LSM, we modeled the likelihood that payment messages assigned to the LSM would offset one another. We modeled two extreme cases. The first case, which we call the "long-cycle" case, is one in which no strict subset of payments is offsetting. In the alternative short-cycles case, every payment perfectly offsets with the payment from some other particular participant. In other words, payments offset bilaterally. In the short-cycles case, if both parties assign their payments to the LSM, the two payments would automatically settle. These two models provide different motives for using the LSM. In the long-cycle case, participants don't assign payments to the LSM queue in the hopes of offsetting them within the queue; instead, they assign them to the queue to have them settle only conditional on receiving another payment (and conditional on their threshold balance). In the short-cycles case, participants can also anticipate that some of the time their queued LSM payments will be offset and will settle directly as the result of their bilateral partner having queued their payment in the LSM as well.

The equilibria under balance-reactive LSMs depend on parameters, as in RTGS, and there can be multiple equilibria. Introducing an LSM has two effects on that multiplicity. On the one hand, being able to condition the release of a payment on the receipt of an offsetting payment eliminates strategic interactions and thus reduced the multiplicity of equilibria. This is particularly clear under the long-cycle assumption, where the equilibrium is unique. On the other hand, when payments settle bilaterally or multilaterally in the queue, as in the short-cycles case, then a new set of strategic interactions emerge and multiplicity of equilibria reappears.

In most parts of the parameter space, the presence of an LSM results in more payments being settled in the early period. For balance-reactive LSMs we find that the equilibrium welfare can exceed that of best equilibrium outcome in an RTGS system alone, but, perhaps surprisingly, it can also result in reduced welfare. Our examination of the conditions for higher welfare with a balance-reactive LSM show that when the costs of delay tend to be high, and the size of the outside settlement system and the proportion of time-critical payments are relatively low, the LSM results in higher equilibrium welfare. When the costs of delay are small and the payments to the outside settlement system are large, the RTGS can achieve higher welfare. The intuition for this result is that RTGS itself creates some beneficial incentives for coordination of payments. Allowing participants to condition their payments on their balances can, in some cases at least, lead some of the participants to queue time-critical payments rather than sending them outright in the early period, reducing beneficial coordination.

Under receipt-reactive LSM the welfare achieved is always at least as high as the level achieved in RTGS. Here the intuition is simpler. As participants cannot condition on their balance, they either submit all their payments to the queue, or simply pay all the time-critical payments in the early period.

In comparing balance-reactive and receipt-reactive LSMs we find that when delay costs are high, and the settlement systems are not too large, the balance-reactive system yields a better outcome than the receipt-reactive LSM. As a result, while our results point to LSMs being at least weakly preferred to RTGS for all parameter configurations, the practical choice can present more of a dilemma to the operator of the large-value payment system. The dilemma is that our results show that the LSM design matters. If the wrong LSM is implemented it can yield either lower welfare than RTGS, or lower welfare than a competing LSM design. The difficulty for an operator is knowing the sizes of the four parameters of interest. In addition, we considered basic design elements in choosing the LSMs to model; more complex designs would introduce other behavioral considerations that are beyond the scope of this paper.

Another parameter that our research shows is important to the incentive to place payments in an LSM queue is the probability of offsetting payments within any particular subset of payments submitted to the queue. Some evidence on this subject is available in simulation studies of LSMs and in the practical experience of large-value payment systems that employ LSMs.

Future research in this area can usefully focus on the question of the empirical magnitudes of the parameters of interest. The cost of delaying payments and the proportion of payments that are time-critical are especially important to measure, and difficult to observe. Further research employing alternative distributions of these parameters is a subject for the future, as is extending the current model to include many periods.

7 Appendix

Derivation of the expressions for W_2, W_3, and W_4

First, note that the expression for W_2 is given by

$$\begin{aligned} W_2 &= -(1-\theta)\sigma(1-\pi)(1-\mu)R \\ &\quad -\theta\sigma(1-\pi)(2\mu-1)R \\ &\quad -\theta(1-2\sigma)(1-\pi)\mu R \\ &\quad -\theta\sigma\left[\gamma+(1-\pi)(1-\mu)R\right]. \end{aligned}$$

Recall that under this equilibrium, $\pi = \theta\left[1-\sigma\right]$. Replacing π we get

$$\begin{aligned} W_2 &= -(1-\theta)\sigma(1-\theta\left[1-\sigma\right])(1-\mu)R \\ &\quad -\theta\sigma(1-\theta\left[1-\sigma\right])(2\mu-1)R \\ &\quad -\theta(1-2\sigma)(1-\theta\left[1-\sigma\right])\mu R \\ &\quad -\theta\sigma\left[\gamma+(1-\theta\left[1-\sigma\right])(1-\mu)R\right]. \end{aligned}$$

The expression for W_3 is given by

$$\begin{aligned} W_3 &= -(1-\theta)\sigma(1-\pi)(1-\mu)R \\ &\quad -\theta\sigma(1-\pi)(2\mu-1)R \\ &\quad -\theta(1-2\sigma)\gamma \\ &\quad -\theta\sigma\left[\gamma+(1-\pi)(1-\mu)R\right] \\ &= -\theta\left[1-\sigma\right]\gamma-(1-\pi)\sigma(1-\mu)R \\ &\quad -\theta(1-\pi)\sigma(2\mu-1)R. \end{aligned}$$

Under this equilibrium, $\pi = \theta\sigma$. Replacing π we get

$$\begin{aligned} W_3 &= -\theta\left[1-\sigma\right]\gamma - (1-\theta\sigma)\sigma(1-\mu)R \\ &\quad -\theta(1-\theta\sigma)\sigma(2\mu-1)R. \end{aligned}$$

The expression for W_4 is given by

$$\begin{aligned} W_4 &= -(1-\theta)\sigma(1-\pi)(1-\mu)R \\ &\quad -\theta\sigma\gamma \\ &\quad -\theta(1-2\sigma)\gamma \\ &\quad -\theta\sigma\left[\gamma + (1-\pi)(1-\mu)R\right] \\ &= -\theta\gamma - (1-\pi)\sigma(1-\mu)R. \end{aligned}$$

Under this equilibrium, $\pi = 0$. Replacing π we get

$$W_4 = -\theta\gamma - \sigma(1-\mu)R.$$

Details of Example 1

It can be checked that if δ and ε are small enough, there is a unique equilibrium without an LSM given by equilibrium 3 of proposition 2. Let W_3 denote the welfare associated with this equilibrium. The equilibrium with a LSM is given by equilibrium $3a$ in proposition 4. Let $W_{4,3a}$ denote the welfare associated with this equilibrium. The probability of receiving a payment early without an LSM is $\pi = \theta(1-\sigma)$, while the probability of receiving a payment early with the LSM is $\pi^o = \pi^q = \theta/(1+\theta)$. Clearly,

$\pi > \pi^o \Leftrightarrow \theta(1-\sigma) > \sigma$, which holds if σ is small enough. We can write

$$\begin{aligned} W_3 - W_{4,3a} &= \sigma(\frac{1}{2} - \delta)R(\pi - \pi^o) + \theta\sigma(2\delta)R(\pi - \pi^o) \\ &\quad + \theta(1-2\sigma)\left[\left(\frac{1}{2} + \delta\right)R(\pi - \pi^o) - \varepsilon(1-\pi^o)\right]. \end{aligned}$$

If $\pi > \pi^o$ and ε is small enough, this expression is positive.

Details of Example 2

First, we establish that the unique robust equilibrium with an LSM is such that all participants who receive a negative liquidity shock delay their payment while only participants with a positive liquidity shock send time-critical payments early.[6] Formally, $\lambda_e = \sigma\theta = 1/4$ and $\lambda_d = \sigma = 1/4$.

Since $\mu = 1/2$,

$$\gamma > \frac{1-\pi^o}{1-\pi^q}(2\mu - 1)R = 0$$

for all $\gamma > 0$. This implies that participants who receive a positive liquidity shock choose to send time-critical payments if γ is not zero. Other participants choose not to make time-critical payments early if

$$\frac{1-\pi^o}{1-\pi^q}\mu R > \gamma.$$

The smallest value that $(1-\pi^o)/(1-\pi^q)$ can take is 1, so the above condition becomes $R/2 > \gamma$

Participants who receive a negative liquidity shock choose to delay if $\frac{\pi^o}{\pi^q}(1-\mu)R > \gamma$. The smallest value that π^o/π^q can take is 4/5, so the above condition becomes $2R/5 > \gamma$.

[6] While it is an equilibrium for all participants to put their payment in the queue, lemma 2 applies here.

Next, we show that the unique equilibrium without an LSM is equilibrium 2 of proposition 2. That equilibrium exists if $R/2 \geq \gamma > R/8$. Equilibrium 1 of proposition 2 does not exist if $R/2 > \gamma$ and equilibrium 3 of proposition 2 does not exist if $\gamma > 3R/8$.

To summarize, the conditions for the equilibria we consider to exist and be unique are $2R/5 > \gamma > 3R/8$.

Now we can evaluate the welfare associated with each equilibrium. Equilibrium 2 of proposition 2 provides welfare $W_2 = -\gamma/4 - 3R/32$. The equilibrium with a LSM yields welfare $W = -7\gamma/16 - R/16$. If follows that $W_2 > W$ if $\gamma > R/6$, which contains the interval $2R/5 > \gamma > 3R/8$.

Details of Example 3

If $\gamma \leq [\mu - \sigma(1-\mu)]$, then the only equilibrium with a RRLSM is for all participants to queue payments. Any BRLSM such that not all payments are queued will yield lower welfare. The condition for participants with a positive liquidity shock prefer to send time-critical payments early is $(1-\pi^q) \geq (1-\pi^o)(2\mu-1)R)$. For any $\gamma > 0$, there exists a μ sufficiently close to 1/2 such that this condition is satisfied. In those cases, it is not a robust equilibrium for all payments to be queued when a BRLSM is available.

References

[1] Angelini, P., 1998. "An analysis of competitive externalities in gross settlement systems." *Journal of Banking and Finance* 22, 1-18.

[2] Angelini, P., 2000. "Are banks risk averse? Intraday timing of operations in the interbank market." *Journal of Money, Credit and Banking* 32, 54-73.

[3] Bech, M. L., and R. Garratt, 2003. "The Intraday Liqudity Management Game." *Journal of Economic Theory* 109, 198-219.

[4] BIS, 2005. "New developments in large-value payment systems" CPSS Publications No. 67.

[5] McAndrews, J. and J. Trundle, 2001. "New Payment System Designs: Causes and Consequences." *Financial Stability Review, Bank of England*, December, 127-36.

[6] Johnson, K., J. J. McAndrews, and K. Soramaki, 2004. "Economizing on Liquidity with Deferred Settlement Mechanisms." *Economic Policy Review*, Federal Reserve Bank of New York, 51-72.

[7] Roberds, W., 1999. "The Incentive Effect of Settlement Systems: A Comparison of Gross Settlement, Net Settlement, and Gross Settlement with Queuing." IMES Discussion Paper Series, 99-E-25, Bank of Japan.

[8] Willison, M., 2005. "Real-Time Gross Settlement and Hybrid Payment Systems: A Comparison." Bank of England working paper No 252.

Figure 1: A unique cycle

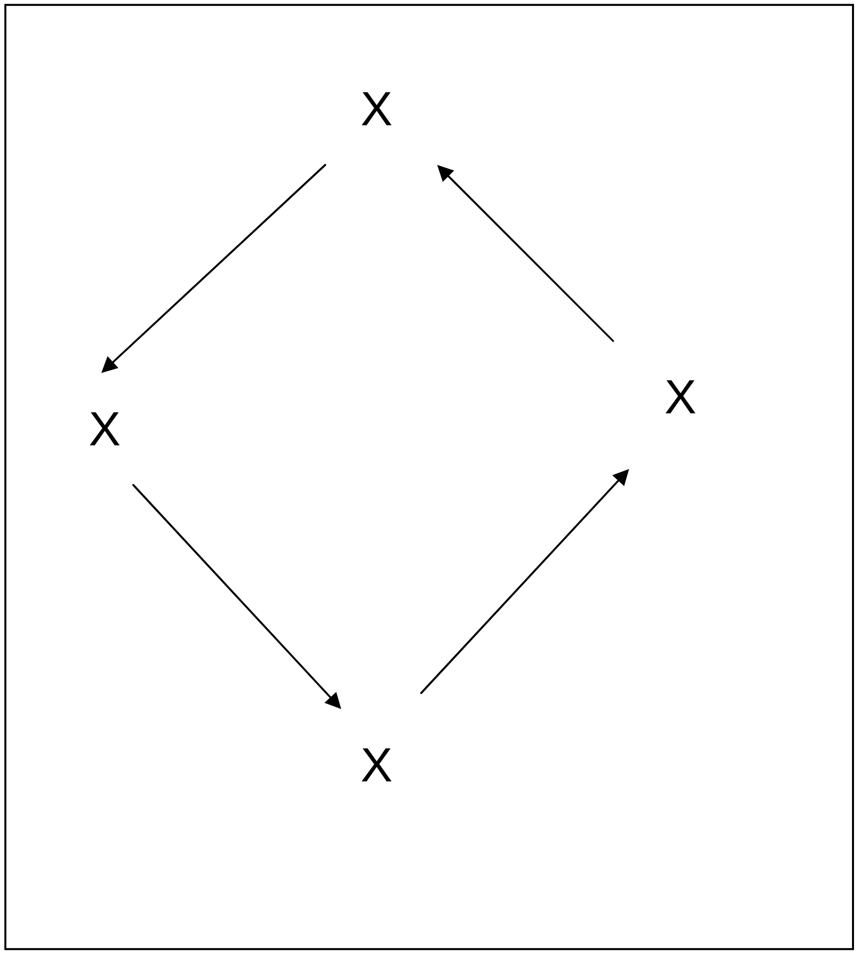

Figure 2: Several cycles

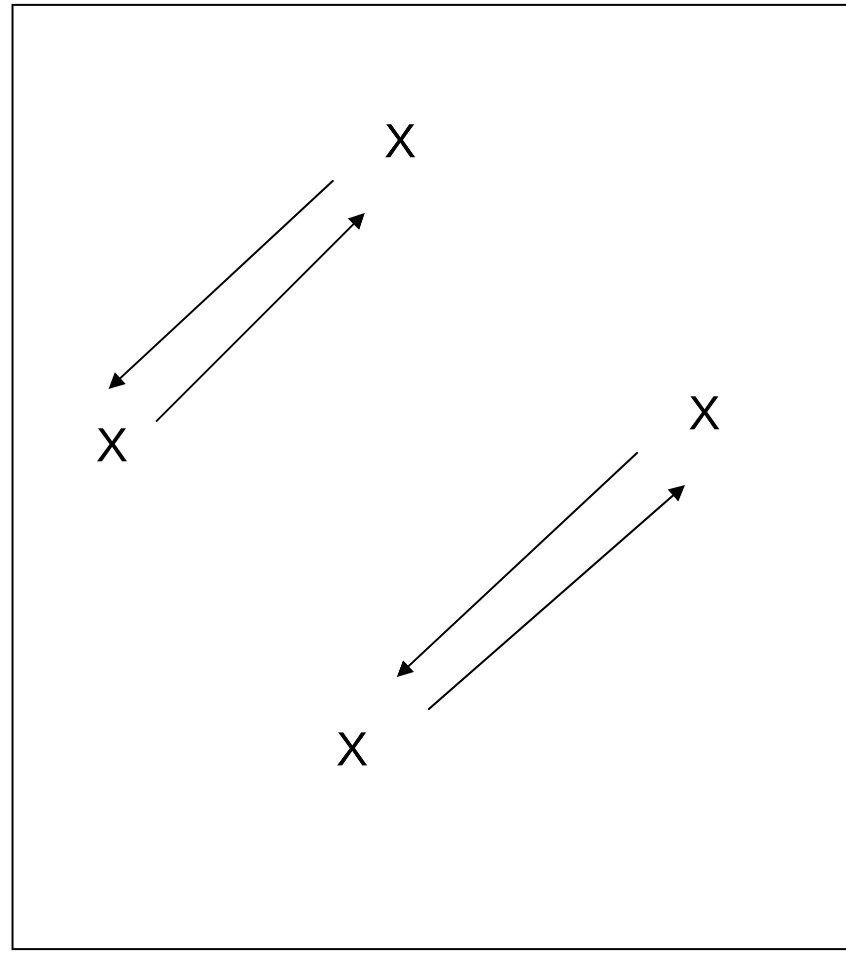

Figure 3: Queued payments in a cycle

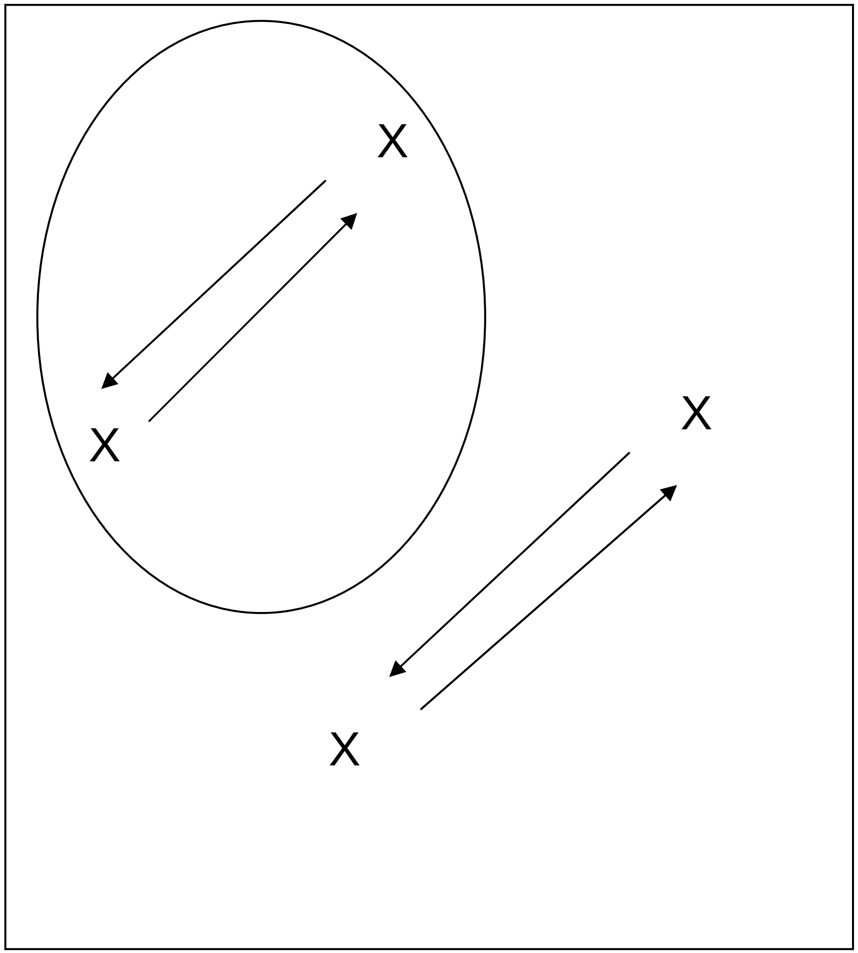

Figure 4: Queued payment in paths