Federal Reserve Bank of New York
Staff Reports

A Study of Competing Designs for a Liquidity-Saving Mechanism

Antoine Martin
James McAndrews

Staff Report no. 336
July 2008

A Study of Competing Designs for a Liquidity-Saving Mechanism
Antoine Martin and James McAndrews
Federal Reserve Bank of New York Staff Reports, no. 336
July 2008
JEL classification: E42, E58, G21

Abstract

We study two designs for a liquidity-saving mechanism (LSM), a queuing arrangement used with an interbank settlement system. We consider an environment where banks are subjected to liquidity shocks. Banks must make the decision to send, queue, or delay their payments after observing a noisy signal of the shock. With a balance-reactive LSM, banks can set a balance threshold below which payments are not released from the queue. Banks can choose their threshold such that the release of a payment from the queue is conditional on the liquidity shock. With a receipt-reactive LSM, a payment is released from the queue if an offsetting payment is received, regardless of the liquidity shock. We find that these two designs have opposite effects on different types of payments. Payments that are costly to delay will be settled at least as early, or earlier, with a receipt-reactive LSM. Payments that are not costly to delay will always be delayed with a receipt-reactive LSM, while some of them will be queued and settled early with a balance-reactive LSM. We also show that parameter values will determine which system provides higher welfare.

Key words: liquidity-saving mechanisms, real-time gross settlement, large-value payment systems

---

Martin: Federal Reserve Bank of New York (e-mail: antoine.martin@ny.frb.org). McAndrews: Federal Reserve Bank of New York (email: jamie.mcandrews@ny.frb.org). The views expressed in this paper are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System.

1 Introduction

This paper provides a comparative study of two types of designs for a liquidity-saving mechanism (LSM), using the framework of Martin and McAndrews (2008). LSMs are queuing arrangements that operate in conjunction with a interbank settlement system, or large-value payment system (LVPS). TARGET 2, the LVPS used in the Euro area already uses an LSM. The Bank of Japan plans to introduce an LSM in October 2008. The Federal Reserve is studying the benefits and costs of implementing an LSM for Fedwire, its LVPS. Hence, evaluating the performance of different LSM designs is an important policy issue.

Theoretical studies of interbank settlement systems typically depart from the standard theory of intertemporal trade because there are no organized markets for short-term (intraday) credit in interbank settlement systems. Hence, there is no market price at which banks can trade reserve balances. Instead, a bank that needs funds to make a payment usually has the choice between borrowing from the central bank at a fixed price or delaying the payment. The incentive to delay has led to a pattern of a highly concentrated surge of payments late in the day. These surges, together with concerns about systemic and operational risk, has made the management of interbank payment flows an important issue for central banks.

Until the mid 1990s, almost all LVPS consisted of netting arrangements between banks. However, concerns that these systems may be prone to cascade of failures lead to the gradual adoption, over the next 20 years, of real-time gross settlement systems (RTGS) (Bech and Hobijn, 2007). With RTGS, each payment is settled on an individual basis and in real time by a transfer of balances on the books of the central bank (CB). While eliminating the risk of cascades of defaults, RTGS systems come with their own problems. Because they require large amounts of liquidity, these systems are particularly prone to delay of payments.

LSMs are queuing arrangements that enhance an RTGS system and aim at reducing banks' need for liquidity. In such a system, a payment can be either sent through the RTGS stream, also known as express stream, or put into the LSM queue, also known as the limit stream. A payment is released from the queue, and settled, under pre-specified conditions. The standard design for an LSM allows banks to specify a balance threshold under which payments cannot be released from the queue. We call such a design a balance reactive LSM (BRLSM).

Johnson, McAndrews, and Soramaki (2004) propose a different kind of design. Under their system, banks cannot specify a balance threshold. However, payments can be released from the queue only if incoming balances in a predetermined interval of time is sufficient. For example, such a system may specify that only balances received by a bank in the last interval of 5 minutes can be used to settle the bank's payments in the queue. We call this kind of system a receipt reactive LSM (RRLSM). Using simulation techniques, Johnson, McAndrews, and Soramaki (2004) have shown that a RRLSM can outperform a BRLSM. RRLSM designs have also been used in policy analysis by Jackson and Ercevik (2008).

The goal of this paper is twofold. First, we want to compare the welfare benefits of a BRLSM and a RRLSM in a theoretical model. We show that depending on parameter values, either designs can provide higher welfare. This can be accounted for in part by the fact that these designs have different effects on different types of payments. We show that under a RRLSM, banks delay all payments that are costless to delay. In contrast, some banks queue such payments under a BRLSM. However, at least as many (or more) payments that are costly to delay will be settled early with a RRLSM than with a BRLSM. We also show that either design can provide higher or lower welfare than an RTGS system, depending on parameter values.

Second, our results suggest caution in evaluating the performance of LSMs using simulation techniques.[1] LSMs can affect the flow of payments for at least two reasons. First, an LSM allows for the conditional release of payments and some netting of payments can occur in the queue. Simulation techniques can be expected to capture these effects well. Second, LSMs affect the incentives of banks to send, queue, or delay payments. Simulations techniques cannot account for these effects. Moreover, our results show that different LSM designs can affect the incentives to delay different types of payments in opposite direction. We conclude that simulations are a poor substitute to good theory as a tool for evaluating settlement system design.

The remainder or the paper proceeds as follows: Section 2 presents some descriptive material concerning LSMs and discusses the related literature. Section 3 introduces the environment. Section 4 describes different designs for the settlement system. Section 5 compares the welfare provided by the designs we consider. Section 6 concludes.

2 RTGS 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 them-

[1] See, for example, the papers in Bank of Finland (2007), for papers that used simulation techniques to evaluate the performance of an LSM, and the references therein.

selves. Among the bank's own payments one can note three types. First, a bank often uses an RTGS for the return and delivery of money market loans. Second are payments to a special-purpose settlement system, such as a securities settlement system or a foreign-exchange settlement system.[2] For example, in the U.S. banks use the Federal Reserve System's RTGS system (Fedwire©) to make payments into and to receive payouts from CLS Bank, a special-purpose bank that settles foreign-exchange trades on its books. Finally, other types of payment made by banks on an RTGS include progress payments under a derivatives contract with another party and payments made on behalf of bank customers. The amounts of some of these payments may not be precisely known at the start of the business day.

Both customer-initiated payments and a bank's own payments may or may not be time-sensitive.[3] Consider a payment to settle a real estate transaction of a customer for which many people are gathered in a closing or settlement meeting. The customer's demand for the payment is highly time-sensitive. The considerations just outlined motivate the assumption that banks are subject to liquidity shocks and may have to make time-sensitive payments.

Liquidity-savings mechanisms to be used in conjunction with RTGS systems are a fairly recent phenomenon.[4] 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 an LSM, but some features

[2]In the U.S., many banks make payments into three private-sector special-purpose payment and settlement systems, CHIPS, CLS bank, and the Depository Trust Company (DTC).

[3]The terms time-critical and time-sensitive are used interchangeably.

[4]See McAndrews and Trundle (2001) and Bank for International Settlements (2005) for a review and extensive descriptive material on LSMs.

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 that 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 appearance 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 LSM 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.

2.1 Relevant literature

Several papers examine the theoretical behavior in RTGS systems. Angelini (1998, 2000) considers the behavior of banks in an RTGS systems in which they face delay costs for payments as well as costly borrowing of funds. He shows 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. Mills and Nesmith (2008) study an environment similar to the one in this paper. Their approach is complementary to our as they focus on the effect of risk, fees, and other factors on the incentives of banks to sent their payments early, or delay, in RTGS systems without LSMs.

Some recent work studying LSMs includes Roberds (1999), who compares gross and net payment systems with systems offering an LSM. He examines the incentives participants have to engage in risk-taking behavior in the different systems. Kahn and Roberds (2001) consider the benefits of coordination from an LSM in the case of CLS. Willison (2005) examines the behavior of participants in an LSM. Our paper extends the framework in Martin and McAndrews (2008) by introducing a noisy signal of a bank's liquidity shock. We also consider RRLSMs, which were not modeled previously. Atalay, Martin, and McAndrews (2008) study the efficient allocation in that environment.

3 The environment

The environment is similar Martin and McAndrews (2008). The economy is populated by a continuum of mass 1 of risk neutral agents. These agents are called payment system participants or banks. There is also a nonstrategic agent which is identified with settlement institutions.[5]

The economy lasts two periods, morning and afternoon. Each bank makes two payments and receives two payments each day. One payment is sent to another bank and is called the core payment. The other payment is sent to the nonstrategic agent and will be crucial in determining the bank's liquidity shock. Similarly, one payment is received from another bank and one is

[5]One can think of the nonstrategic agent as aggregating several distinct institutions such as the CLS bank, CHIPS, and DTC.

received from the nonstrategic agent. Core payments have size μ, while payments to and from the nonstrategic agent have size $1-\mu$. It is assumed that $\mu \geq 1/2$.

Three factors influence the decision of banks to send, queue, or delay their core payment. First, banks must pay a cost to borrow from the CB. Second, banks may need to send a time-sensitive payment. Third, banks may receive a liquidity shock. In contrast to Martin and McAndrews (2008), banks in this model may not know their liquidity shock.[6] Instead they receive a signal, which can be noisy.

Each bank starts the day with zero reserves. Reserves can be borrowed from the CB at an interest cost of R.[7] Banks that receive more payments than they send in the morning have excess reserves. It is assumed that these reserves cannot be lent to other banks so that banks receive no benefit from excess reserves. Payments received and sent in the same period offset each other. Hence, a bank only needs to borrow from the CB if the payments it makes in the morning exceed the payments it receives in the morning.

Banks learn in the morning whether the payment they must make to

[6] Banks may not know their liquidity shock with certainty at the time they plan their payment submission strategy because such a shock can arise from the activity of a securities settlement system. In the Fedwire Securities Settlement system in the U.S., for example, banks choose the time of the delivery of the securities which then triggers a payment of funds out of the account of the receiver of the securities. Consequently, the bank from whose account funds flow may not be aware of the timing at which its counterparty will deliver securities to it. As a result, a bank does not have perfect control, nor is it perfectly informed ex ante, about the state of liquidity in its account at the time it is making its decision to send a payment in the next instant.

[7] Evidence discussed in Mills and Nesmith (2008) suggests that the cost of intraday reserves can influence banks payment behavior. In our welfare analysis, we think of R as representing both the private and the social cost borrowing reserves.

another bank is time-critical. The probability that a payment is time-critical is denoted by θ.[8] 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. Banks choose whether to send the payment in the morning before they know if they will receive a payment from another bank in the morning. Participants form rational expectations about the probability of receiving a payment from some other bank in the morning. Let π denote this expectation.

Banks learn in the morning when they receive a payment from the nonstrategic agent and when they must send an offsetting payment. The probability of receiving the payment in the morning is $\bar{\pi}$ and so is the probability of having to send the payment in the morning. Both events are uncorrelated. 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. 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. 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. These agents do not receive a liquidity shock. The following table contains the definition of all parameters.

[8] Throughout the paper it is assumed that if x represents the probability of an event occurring for a bank, then the fraction of banks for whom this event occurs is x as well. Hence, a fraction θ of banks must make a time-critical payment.

Table 1: Parameters of the model

$\mu \in [0.5, 1]$	Size of payment to other banks
$R > 0$	Cost of borrowing
$\theta \in [0, 1]$	Probability of having to make a time-sensitive payment
$\gamma > 0$	Cost of delay
$\sigma \in [0, 0.25]$	Probability of a liquidity shock

The role played by the different frictions in the model can be summarized as follows. The cost of borrowing provides an incentive to bunch payments. Indeed, absent liquidity shocks banks could avoid borrowing if either all payments are sent in the morning or all payments are delayed. Time-sensitive payments provide an incentive for some banks to send their payment early in order to avoid the delay cost. In contrast, banks that receive a negative liquidity shock have an incentive to delay their payment to avoid having to borrow from the CB.

Before they observe the realization of the liquidity shock, banks observe a, possibly noisy, signal. The signal can take three values, s^+, s^0, and s^-. If a bank observes signal s^+, then the probability that it will experience a positive liquidity shock is $p \in [\sigma, 1]$. The bank will experience no liquidity shock with probability $(1-p)(1-2\sigma)/(1-\sigma)$ and a negative liquidity shock with probability $(1-p)\sigma/(1-\sigma)$. Similarly, a bank that observes signal s^- experiences a negative liquidity shock with probability p, no liquidity shock with probability $(1-p)(1-2\sigma)/(1-\sigma)$, and a positive liquidity shock with probability $(1-p)\sigma/(1-\sigma)$. Finally, a bank that observes the signal s^0 receives no liquidity shock with probability $p^0 \in [1-2\sigma, 1]$, a positive liquidity shock with probability $(1-p^0)/2$, and a negative liquidity shock with the same probability. We assume that p and p^0 satisfy the condition

$$\frac{1-p}{1-p^0} = \frac{1-\sigma}{2\sigma},$$

so the shares of banks that observe a positive, a negative, and no liquidity shock are σ, σ, and $1-2\sigma$, respectively, for all $p \in [\sigma, 1]$. These shares corresponds to the share of the population that will incur such shocks. The probability p can be thought of as the accuracy of the signal. If $p = p^0 = 1$, then the signal is perfectly accurate. If $p = \sigma$, which implies $p^0 = 1-2\sigma$, then the signal provides no information about the shock a bank will receive.

The timing of events in summarized in Figure 1. First, nature chooses the banks that receive a liquidity shock and the banks that must make time-critical payments. Next, banks receive the signal about their liquidity shock. After observing the signal, banks have the choice between sending their core payment, delaying the payment, or put it in an LSM queue, when such a queue is available. Next, the morning payments to the settlement systems are made, which result in the banks' liquidity shocks. At the end of the morning period, banks that must borrow from the CB incur a borrowing cost while banks that did not sent time-sensitive payments incur a delay cost. All remaining payments are made in the afternoon.

[Figure 1]

3.1 liquidity-saving mechanisms

This section briefly describes the way an LSM works. An LSM is a queue to which payments can be submitted. Payments are released from the queue according to pre-specified rules. In this paper we consider two types of LSM: A receipt-reactive LSM (RRLSM) and a balance reactive LSM (BRLSM). Under a BRLSM, banks can choose whether a queued payment should be released conditional on the level of their balance. In our model, this is equivalent to letting a bank choose whether to queue after it observes a perfect signal of its liquidity shock. Indeed, a bank that queues its core payment can forecast perfectly the amount of its reserves conditional on the liquidity shock it receives. By choosing its threshold appropriately, the banks can guarantee that its payment will be released from the queue, or not, depending on the liquidity shock and the receipt of an offsetting payment. The following table shows the balance of a bank if the bank's core payment is released, conditional on receiving an offsetting payment and on the liquidity shock.

Table 2: Bank balances

		Liquidity shock		
		negative	no shock	positive
Offsetting payment	not received	-1	$-\mu$	$-(2\mu - 1)$
	received	$-(1-\mu)$	0	$(1-\mu)$

If a bank chooses a threshold smaller than -1, then its queued payment will be released in all circumstances. Hence this is equivalent to sending the payment through the RTGS stream. If the threshold is greater than $(1-\mu)$, then the payment stays in the queue in all circumstance. This is equivalent to delaying the payment. If, for example, a bank chooses a threshold between -1 and $-\mu$, then the queued payment will be settled unless the bank has a negative liquidity shock and does not receive an offsetting payment. In an abuse of terminology, when talking about a BRLSM, we may say that a bank sends its payment if it receives a positive liquidity shock to mean that the bank queues its payment and sets a threshold such that the payment is released only if the bank's liquidity shock is positive.

Under a RRLSM, the release of a payment from the queue is independent of the bank's balance. In this case we assume that the bank must make its decision to queue after observing a noisy signal of its liquidity shock. If a payment is queued, it is released if and only if an offsetting payment is received, independently of shock the bank experiences. However, the bank can condition its behavior on the signal it receives.

The probability with which a payment is released from the queue depends in part on the underlying pattern of payments. Suppose, for example, that no strict subset of payments is multilaterally offsetting, as in Figure 2. In that case, no netting can occur in the queue unless all payments are queued. If all payments form pairs of bilaterally offsetting payments, as in Figure 3, the probability that payments in the queue will net is the highest. In this paper, we make the extreme assumption that there are no offsetting payments in the queue unless all payments are queued, as in Figure 2. This assumption allows us to consider the benefits of an LSM in the case where no netting occurs. Our results extend to other specifications of the underlying pattern of payments.

[Figure 2 and 3]

Let λ_e denote the fraction of banks that send their payment early, λ_q denote the fraction of banks that put their payment in the queue, and λ_d denote the fraction of banks that delay their payment. Clearly, $\lambda_e + \lambda_q + \lambda_d = 1$.

Martin and McAndrews (2007) 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 the payment in the queue. The latter probability is equivalent to the probability that a payment in the queue is released. Given our assumption on the underlying pattern of payments, we have

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

Define $\pi \equiv \pi^o = \pi^q$.

We call a settlement system in which a LSM is not available a real-time gross settlement system (RTGS). Under RTGS, a bank can choose either to delay or to send its core payment in the morning.

4 The settlement system designs

In this section, we characterize the equilibria with a RRLSM, a BRLSM, and a pure RTGS system. We focus mainly on RRLSMs since Martin and McAndrews (2008) already describe BRLSM and RTGS systems, albeit in a model without noisy signals.

4.1 Participants' behavior under a receipt-reactive LSM

First we consider the incentives to send a payment early, queue, or delay, for banks that receive a signal s^+. If a payment is delayed, the bank must pay the delay-cost for a time-critical payment, regardless of the liquidity shock. In addition, if the bank receives a positive liquidity shock, which occurs with probability p, or no liquidity shock, which occurs with probability $(1-p)(1-2\sigma)/(1-\sigma)$, no other cost is incurred. With probability $(1-p)\sigma/(1-\sigma)$, the bank 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

$$\gamma + \frac{(1-p)\sigma}{(1-\sigma)}(1-\pi)(1-\mu)R. \tag{2}$$

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

$$(1-\pi)\gamma + \frac{(1-p)\sigma}{(1-\sigma)}(1-\mu)R. \tag{3}$$

If a payment is sent early, then no delay cost is incurred. However, the bank must borrow 1 if it experiences a negative liquidity shock and fails to receive an offsetting payment in the morning. If the bank receives an offsetting payment, it must borrow $1-\mu$. If the bank experiences no liquidity shock and does not receive an offsetting payment in the morning, it must borrow μ. Finally, if the bank experiences a positive liquidity shock and does not receive an offsetting payment in the morning, it must borrow $(2\mu-1)$. Banks that experience a positive or no liquidity shock need not borrow from the central bank if they receive offsetting payments in the morning. It follows that the expected cost of sending payments in the morning is

$$\frac{(1-p)\sigma}{(1-\sigma)}(1-\pi\mu)R + \frac{(1-p)}{(1-\sigma)}(1-2\sigma)(1-\pi)\mu R + p(1-\pi)(2\mu-1)R. \tag{4}$$

Using similar steps, we can obtain the cost of delaying, queuing, and sending a payment early for banks with shocks s^0. These are

$$\gamma + \frac{1}{2}(1-p^0)(1-\pi)(1-\mu)R, \tag{5}$$

$$(1-\pi)\gamma + \frac{1}{2}(1-p^0)(1-\mu)R, \tag{6}$$

$$\frac{1}{2}(1-p^0)(1-\pi\mu)R + p^0(1-\pi)\mu R + \frac{1}{2}(1-p^0)(1-\pi)(2\mu-1)R, \tag{7}$$

respectively.

The cost of delay, queuing, and sending a payment early for banks with shocks s^- are

$$\gamma + p(1-\pi)(1-\mu)R, \tag{8}$$

$$(1-\pi)\gamma + p(1-\mu)R, \tag{9}$$

$$p(1-\pi\mu)R + \frac{1-p}{1-\sigma}(1-\pi)(1-2\sigma)\mu R + \frac{1-p}{1-\sigma}\sigma(1-\pi)(2\mu-1)R, \tag{10}$$

respectively.

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

Proposition 1 *Assume* $\bar{\pi} \in (0,1)$ *and* $\mu \in [0.5,1)$*, Under a receipt-reactive LSM, banks that receive a signal* s^+

- *delay the payment if* $\frac{(1-p)\sigma}{(1-\sigma)}(1-\mu)R > \gamma$*,*
- *queue the payment if* $[\mu - p(1-\mu)]\,R > \gamma \quad$ *and* $\quad \gamma \geq \frac{(1-p)\sigma}{(1-\sigma)}(1-\mu)R$*,*
- *make the payment early if* $\gamma > [\mu - p(1-\mu)]\,R$*.*

Banks that receive a signal s^0

- *delay the payment if* $\frac{1}{2}(1-p^0)(1-\mu)R > \gamma$*,*
- *queue the payment if* $\left[\mu - (1-p^0)\frac{1}{2}(1-\mu)\right] R > \gamma \quad$ *and* $\quad \gamma \geq \frac{1}{2}(1-p^0)(1-\mu)R$*,*
- *make the payment early if* $\gamma > \left[\mu - (1-p^0)\frac{1}{2}(1-\mu)\right] R$*.*

Banks that receive a signal s^-

- *delay the payment if* $p(1-\mu)R > \gamma$*,*
- *queue the payment if* $[\mu - \sigma + p(1-\mu)\sigma]\,\frac{R}{1-\sigma} > \gamma \quad$ *and* $\quad \gamma \geq p(1-\mu)R$*,*
- *make the payment early if* $\gamma > [\mu - \sigma + p(1-\mu)\sigma]\,\frac{R}{1-\sigma}$*.*

Proof. The boundaries for delaying, queuing, or sending payments in the morning come from comparing equations (2), (3), and (4), for banks with signal s^+, equations (5), (6), and (7), for banks with signal s^0 and equations (8), (9), and (10), for banks with signal s^-. ■

4.2 Equilibria under a receipt-reactive LSM

We can now describe the equilibria. First, it should be noted that it is always an equilibrium for all payments to be put in the queue. If all payments are in the queue, they all settle in the morning. This implies a cost only for banks that receive a negative liquidity shock. However, such banks cannot benefit from a deviation. Indeed, since no payment is sent outright in the morning, the only payments released from the queue are bilaterally or multilaterally offsetting payments. Hence banks that delay their payment cannot receive a payment in the morning from another bank, so delaying has no benefit over queueing.

When other equilibria exist, the equilibrium where all banks queue their payment can be refined away, as is shown in the following lemma.

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

The proof is provided in Martin and McAndrews (2007). In the remainder of this paper, we focus on the equilibrium with $\lambda_q < 1$ when it exists. Equilibria are characterized in the following proposition.

Proposition 2 *Under the long-cycle assumption, we have the following equilibria:*

1. *If $\gamma < [\mu - p(1-\mu)]\, R$, then all participants put their payment in the queue*
2. *If $\gamma \geq [\mu - p(1-\mu)]\, R$ and $\mu \geq \frac{2p}{1+2p}$, then*
   - (a) *If $\gamma \geq [\mu - \sigma + p(1-\mu)\sigma]\, \frac{R}{1-\sigma}$, then all time-critical payments are made early. All non-time-critical payments are delayed.*

(b) *If* $[\mu - \sigma + p(1-\mu)\sigma]\,\frac{1}{1-\sigma} R > \gamma \geq [\mu - p(1-\mu)]\,R$, *then participants with signal* s^+ *make time-critical payment early. Participants with signal* s^0 *or* s^- *put their time-critical payment in the queue. All non-time-critical payments are delayed.*

3. *If* $\gamma \geq [\mu - p(1-\mu)]\,R$ *and* $\mu < \frac{2p}{1+2p}$, *then*

   (a) *If* $\gamma \geq [\mu - \sigma + p(1-\mu)\sigma]\,\frac{R}{1-\sigma}$, *the equilibrium is the same as under 2a.*

   (b) *If* $[\mu - \sigma + p(1-\mu)\sigma]\,\frac{R}{1-\sigma} > \gamma \geq p(1-\mu)R$, *the equilibrium is the same as under 2b.*

   (c) *If* $p(1-\mu)R > \gamma \geq [\mu - p(1-\mu)]\,R$, *then participants with signal* s^+ *send their time-critical payment early. Participants with signal* s^0 *queue their time-critical payment. Participants with signal* s^- *delay their time-critical payment. All non-time-critical payments are delayed.*

Proof. First, we show that if $\gamma < [\mu - p(1-\mu)]\,R$, then all banks queue their payments. Under this condition, banks with a signal s^+ prefer to queue or delay time-critical payments, and so that no bank wants to send payments early. If no payments are sent early, banks that delay receive a payment early with probability zero. Hence, queuing is weakly preferred to delaying and all banks queue.

Next, note that if some payments are sent early, which happens if $\gamma \geq [\mu - p(1-\mu)]\,R$, then all non-time-critical payments are delayed. Indeed, notice that under proposition 1 banks always delay payments such that $\gamma = 0$, which corresponds to non-time-critical payments.

The other equilibria can be found by considering the thresholds for sending or queuing payments for banks with different signals. Note that the threshold between queuing and paying early is the same for banks with signals s^0 and s^-.

Lemma 1 applies here so that if $\gamma \geq [\mu - p(1-\mu)]\, R$, the equilibrium such that all participants queue their payment is not robust. ■

4.3 Equilibria under a balance reactive LSM and RTGS

An equilibrium under a balance reactive LSM corresponds to a situation of a perfect signal or $p = 1$. Notice that proposition 2 with $p = 1$ is identical to proposition 6 in Martin and McAndrews (2008). We briefly describe equilibria in the RTGS case.[9]

Under RTGS, banks do not have the option to send their payments into a queue. The only options are to delay a payment or to send it outright. The cost of delaying and sending a payment early are given by equations (2) and (4), respectively, for banks with signal s^+, equations (5) and (7), respectively, for banks with signal s^0, (8) and (10), respectively, for banks with signal s^-. Combining these expressions, yields the following proposition.

Proposition 3 *Banks delay all non-time-critical payments unless $\pi = 1$. Banks make time-critical payments according to the following rules:*

1. *If $\gamma \geq p\,[\mu - \pi(2\mu-1)]\, R + \frac{1-p}{1-\sigma}(1-\pi)(1-2\sigma)\mu R + \frac{1-p}{1-\sigma}\sigma(1-\pi)(2\mu-1)R$, then all banks make time-critical payments in the morning.*

2. *If $p\,[\mu - \pi(2\mu-1)]\, R + \frac{1-p}{1-\sigma}(1-\pi)(1-2\sigma)\mu R + \frac{1-p}{1-\sigma}\sigma(1-\pi)(2\mu-1)R > \gamma \geq \frac{1}{2}(1-p^0)\,[\mu - \pi(2\mu-1)]\, R + p^0(1-\pi)\mu R + \frac{1}{2}(1-p^0)(1-\pi)(2\mu-1)R$, then banks who receive a signal s^- choose to delay time-critical payments. Other banks do not.*

[9] For more details, see Martin and McAndrews (2008) we characterize RTGS equilibria in an environment without noisy signals.

3. *If* $\frac{1}{2}(1-p^0)\left[\mu-\pi(2\mu-1)\right]R+p^0(1-\pi)\mu R+\frac{1}{2}(1-p^0)(1-\pi)(2\mu-1)R >$ $\gamma \geq \frac{1-p}{1-\sigma}\sigma\left[\mu-\pi(2\mu-1)\right]R+\frac{1-p}{1-\sigma}(1-2\sigma)(1-\pi)\mu R+p(1-\pi)(2\mu-1)R,$ *then only banks that received a signal* s^+ *choose to make time-critical payments in the morning. All others delay time-critical payments.*

4. *If* $\frac{1-p}{1-\sigma}\sigma\left[\mu-\pi(2\mu-1)\right]R+\frac{1-p}{1-\sigma}(1-2\sigma)(1-\pi)\mu R+p(1-\pi)(2\mu-1)R > \gamma,$ *then all banks delay time-critical payments.*

Using these rules, we can characterize the different equilibria that can arise as in the next proposition:

Proposition 4 *Four different equilibria can exist:*

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

2. *If*
$$p\left[\mu-\theta(1-\sigma)(2\mu-1)\right]+\frac{1-p}{1-\sigma}\left[1-\theta(1-\sigma)\right]R\left[(1-2\sigma)\mu+\sigma(2\mu-1)\right] > \gamma$$
$$\geq \frac{1}{2}(1-p^0)\left[\mu-\theta(1-\sigma)(2\mu-1)\right]R+\left[1-\theta(1-\sigma)\right]R\left[p^0\mu+\frac{1}{2}(1-p^0)(2\mu-1)\right],$$
then it is an equilibrium for banks that received a signal s^- *to delay time-critical payments while other banks pay time-critical payments in the morning.*

3. *If*
$$\frac{1}{2}(1-p^0)\left[\mu-\sigma\theta(2\mu-1)\right]R+(1-\theta\sigma)R\left[p^0\mu+\frac{1}{2}(1-p^0)(2\mu-1)\right] > \gamma$$
$$\geq \frac{1-p}{1-\sigma}\sigma\left[\mu-\sigma\theta(2\mu-1)\right]R+\frac{1-p}{1-\sigma}(1-2\sigma)(1-\sigma\theta)\mu R+p(1-\sigma\theta)(2\mu-1)R,$$
then it is an equilibrium for only banks that received a signal s^+ *to make time-critical payments in the morning.*

4. *If* $\frac{1-p}{1-\sigma}\sigma\mu R + \frac{1-p}{1-\sigma}(1-2\sigma)\mu R + p(2\mu - 1)R > \gamma$*, then it is an equilibrium for all banks to delay time-critical payments.*

As in Martin and McAndrews (2008) multiple equilibria can arise for a given set of parameters.

5 Welfare comparison

In this section, we compare the welfare provided by a RRLSM, a BRLSM, and an RTGS system. We show that no system dominates the others in the sense of providing higher welfare for all parameter values. In fact, we can find parameter values such that any of the three systems provide the highest welfare. In addition, we can show that a RRLSM provides higher incentives to delay non-time-critical payments than a BRLSM, which tends to reduce welfare, but provides better incentives to send or queue time-sensitive payments than a BRLSM, which tends to increase welfare.

5.1 RRLSM vs. BRLSM

First we show that, in equilibrium, welfare increases with the share of payments that are received early.

Proposition 5 *Let* π_A *and* π_B *denote the fraction of payments received in the morning in two different equilibria denoted by* A *and* B*, respectively. Let* W_A *and* W_B *denote welfare under these equilibria. If* $\pi_A \geq \pi_B$*, then* $W_A \geq W_B$*.*

Proof. Focus on a particular bank and let S_A and S_B denote the equilibrium strategies of this bank corresponding to each equilibrium. Let $W(S_A, \pi_A)$ and

$W(S_B, \pi_B)$ denote the welfare of this bank associated with each equilibrium. Note that $W(S_B, \pi_A) \geq W(S_B, \pi_B)$, since all the actions that banks can take have a cost that is (weakly) decreasing in π, as shown in equations (2) to (10). Further, by definition of an equilibrium, $W(S_A, \pi_A) \geq W(S_B, \pi_A)$. It follows that $W(S_A, \pi_A) \geq W(S_B, \pi_B)$. ■

This is a powerful result as it allows us to compare the welfare provided by two different equilibria by looking only at the fraction of payments that are settled in the morning. The next two propositions compare the settlement of different types of payments with a RRLSM and a BRLSM.

Proposition 6 *If $\gamma \geq [\mu - \sigma(1-\mu)] R$, non-time-sensitive payments are delayed with a RRLSM. In contrast, banks queue their non-time-sensitive payment with a BRLSM.*

Proof. Proposition 1 shows that if $p = p^0 = 1$, which corresponds to a BRLSM, then banks who receive a positive or no liquidity shock choose to queue their non-time-sensitive payments. In contrast, since $\sigma(1-\mu)R > 0$, all non-time sensitive payments are delayed with a RRLSM if $p > \sigma$. ■

Proposition 7 *The share of time-critical payments that are released in the morning is at least as high with a RRLSM as with a BRLSM.*

Proof. The proof follow from inspection of the cutoff between different types of equilibria described in proposition 2. It can be verified that more payments are released from the queue in the morning under equilibrium 1 than under equilibria $2a$ or $3a$, under equilibria $2a$ or $3a$ than under equilibria $2b$ or $3b$, and under equilibria $2b$ or $3b$ than under equilibrium $3c$.

Recall that with a BRLSM, $p = 1$, so the value of p is smaller or equal with an RRLSM than with a BRLSM. With a smaller value of p, either the equilibrium is the same with RRLSM as with BRLSM, or the equilibrium has more time-critical payments released in the morning. ■

These two results show that a BRLSM and a RRLSM have very different effects on different types of payments. A RRLSM has bad incentives properties for non-time-sensitive payments, compared to a BRLSM, but it has at least as good or better incentive properties for time-sensitive payments.

With a BRLSM, banks can protect themselves against the risk of having to borrow at the central bank. This makes them willing to queue non-time-sensitive payments as this involves no cost. Balance thresholds will be set such that queued payments are not released from the queue if a bank receives a negative liquidity shock. If the bank receives a positive or no liquidity shock, queued payments will be released if an offsetting payment is received. In contrast, with a RRLSM a bank always runs the risk of having a negative liquidity shock and cannot protect itself against that risk. Since delaying a non-time-critical payment has no cost, banks prefer to delay such payments to protect themselves against the case of a negative shock.

The positive effect of an RRLSM on time-critical payments comes from two separate effects. The *uncertainty* associated with a RRLSM affects banks that have received different signals differently. The incentives of banks that receive a signal s^- to send a payment early or queue are increased by this uncertainty because more uncertainty is associated with a higher probability of receiving a positive liquidity shock or no shock. In contrast, the incentives of banks that receive a signal s^+ to send a payment early are reduced. However, this reduced willingness of banks with a signal s^+ to send payments early results, in equilibrium, in increased incentives to queue payments. Indeed, if no payments are sent early, there is no benefit of delaying payments, as noted previously. So reducing the incentive of banks with a signal s^+ to send their payment early can have the beneficial effect of inducing all banks to queue.

It is apparent from propositions 6 and 7 that welfare can be higher or lower with a RRLSM, compared to a BRLSM, depending on parameters. For example, if $(2\mu - 1)\, R < \gamma < [\mu - \sigma(1-\mu)]\, R$, then welfare with RRLSM is greater than with a BRLSM, since in that case a RRLSM achieves the equilibrium where all payments are settled in the morning but BRLSM does not. In contrast, if γ is sufficiently large and the fraction of time-sensitive payments, θ, is small, then a RRLSM provides lower welfare than a BRLSM. If γ is sufficiently large, all time-sensitive payments will be sent early with either a BRLSM and a RRLSM. However, all non-time-sensitive payments would be delayed with the RRLSM while banks with a positive and no liquidity shock would queue those payments with a BRLSM.

5.2 RRLSM vs RTGS

Now we compare welfare with a RRLSM or with an RTGS. Our first result shows that if the signal provides no information, then a RRLSM provides higher welfare than a RTGS system.

Proposition 8 *If $p = \sigma$, then a RRLSM provides higher welfare than a RTGS system.*

Proof. From proposition 2, if $p = \sigma$ then either all payments are queued, and released in the morning, or all time-sensitive payments are sent in the morning. The RTGS equilibrium with the highest welfare is such that all time-sensitive payments are made early. ■

However, the following example show that this result does not carry over to signals that contains some information, no matter how small.

Example 1 *Let $\mu = 0.5$, $\sigma = 0.25$, $\theta = 0.8$, $\gamma/R = 0.375$ and $p = \sigma + \varepsilon$, $\varepsilon > 0$, then welfare with RTGS system can be higher than with a RRLSM.*

For these parameter values, it is an equilibrium for banks with signals s^+ and s^0 to send their payments early under an RTGS system since

$$\gamma/R = 0.375 \geq \frac{1}{2}(1-p^0)\left[\mu - \pi(2\mu-1)\right] + p^0(1-\pi)\mu + \frac{1}{2}(1-p^0)(1-\pi)(2\mu-1) = 0.225 - \varepsilon/30.$$

With this equilibrium, $\pi^{RTGS} = \theta(1-\sigma) = 0.6$.[10] The equilibrium with a RRLSM is such that banks with a signal s^+ make time-sensitive payments early, while banks with signal s^0 and s^- queue time-sensitive payments since

$$\frac{[\mu - \sigma + p(1-\mu)\sigma]}{1-\sigma} = 0.375 + \varepsilon/6 > \gamma/R = 0.375 \geq [\mu - p(1-\mu)] = 0.375 - \varepsilon/2.$$

With this equilibrium,

$$\pi^{RRLSM} = \frac{\sigma\theta}{1-\theta(1-\sigma)} = 0.5.$$

Since $\pi^{RTGS} > \pi^{RRLSM}$, proposition 5 says that welfare is higher with an RTGS system.

Proposition 8 and example 1 show that, depending on parameter values, a RRLSM can provide more or less value than an RTGS system.

One should also note that if $\theta = 1$ and p is sufficiently close to 1, then the welfare provided by a RRLSM is arbitrarily close to the welfare provided by a BRLSM, except for a set of parameters that has measure zero in the parameter space. For this reason, the results in Martin and McAndrews (2008) comparing the welfare provided by a BRLSM and an RTGS system in a model where $p = 1$ can be extended to compare the welfare provided by a RRLSM and an RTGS system, at least when $\theta = 1$ and p is sufficiently close to 1.

[10] Details of the calculations are provided in the appendix

6 Conclusion

In this paper we compared two competing designs for an LSM in the model of Martin and McAndrews (2008), augmented with a noisy signal. Under a BRLSM, banks can choose a threshold balance below which payments are not sent. This implies that banks can disregard their signal and condition the release of their payment on their actual liquidity shock. With a RRLSM, and payment that is queued will be released upon receipt of an offsetting payment.

We have shown that these competing designs have very different effects on incentives to delay, queue, or send different types of payments. RRLSM provide incentives to delay non-time-sensitive payments compared to a BRLSM. In contrast, at least as many time-sensitive payments are settled early with RRLSM than with a BRLSM. Hence, depending on parameter values, either a RRLSM or a BRLSM can provide higher welfare. Similarly, either of these system designs can provide higher or lower welfare than a pure RTGS system.

Our results suggest that simulation techniques, which have been used to evaluate LSMs, may not be a reliable guide to policy. LSMs affect the flow of payments in two ways. First, they allow for the conditional release of payments and the offsetting of queued payments. These effects are likely to be captured well by simulation techniques. Second, LSMs change banks incentives to send, queue, or delay payments. Simulation techniques are unable to account for such effects. In addition, as our research shows, these effects can depend on design of an LSM and different designs can opposite effects on different types of payments.

7 Appendix

Details of example 1

Since $\mu = 0.5$,

$$\frac{1}{2}(1-p^0)\left[\mu - \pi(2\mu-1)\right] + p^0(1-\pi)\mu + \frac{1}{2}(1-p^0)(1-\pi)(2\mu-1) \tag{11}$$

$$= \frac{1}{2}(1-p^0)\mu + p^0(1-\pi)\mu = 0.25(1-p^0) + 0.2p^0. \tag{12}$$

Recall that

$$p^0 = 1 - (1-p)\frac{2\sigma}{1-\sigma} = 1 - \frac{2}{3}(0.75+\varepsilon),$$

so $0.25(1-p^0) + 0.2p^0 = 0.225 - \varepsilon/30$.

8 References

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

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

Atalay, E., Martin, A., McAndrews, J., 2008. The welfare effects of a liquidity-saving mechanism. Federal Reserve Bank of New York Staff Report 331.

Bank of Finland, 2007. Simulation studies of liquidity needs, risks and efficiency in payment networks. H. Leinonen (ed.). Scientific monographs E:39.

Bech, M. L., Garratt, R., 2003. The Intraday Liqudity Management Game. Journal of Economic Theory 109, 198-219.

Bech, M. L., Hobijn, B., 2007. Technology Diffusion within Central Banking: The Case of Real-Time Gross Settlement. International Journal of Central Banking 3, 147-181.

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

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

Kahn, C. M., Roberds W., 2001. The CLS bank: A solution to the risks of international payments settlement? Carnegie-Rochester Conference Series on Public Policy 54, 191-226.

Martin, A., McAndrews, J., 2007. liquidity-saving Mechanisms. Staff Report 282, Federal Reserve Bank of New York.

Martin, A., McAndrews, J., 2008. liquidity-saving Mechanisms. Journal of Monetary Economics 55, 554-567.

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

Mills, D. C. Jr., Nesmith, T. D., 2008. Risk and Concentration in Payment and Securities Settlement Systems. Journal of Monetary Economics 55, 542-553.

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.

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

Figure 1: Timeline

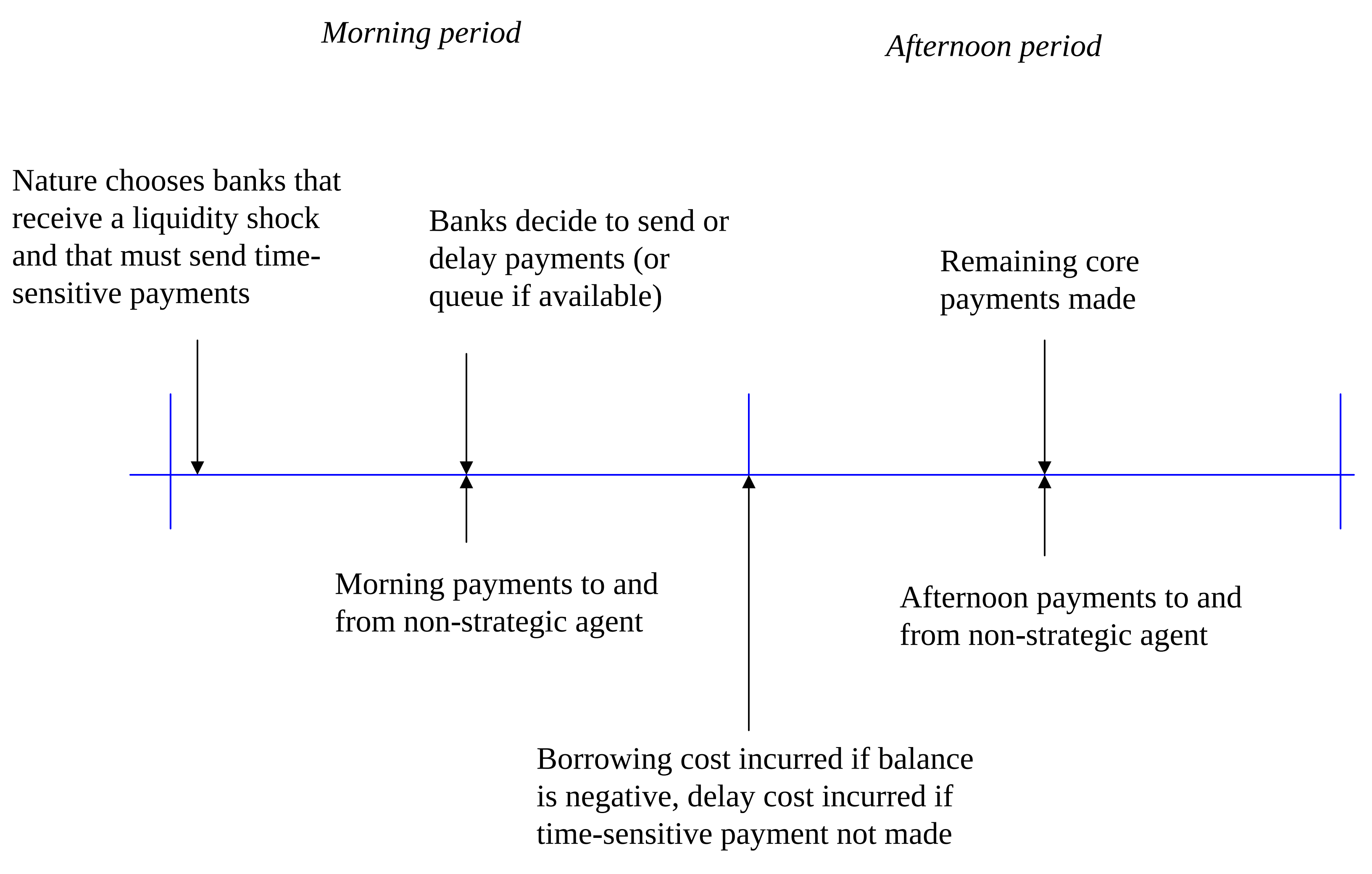

Figure 2: A unique cycle

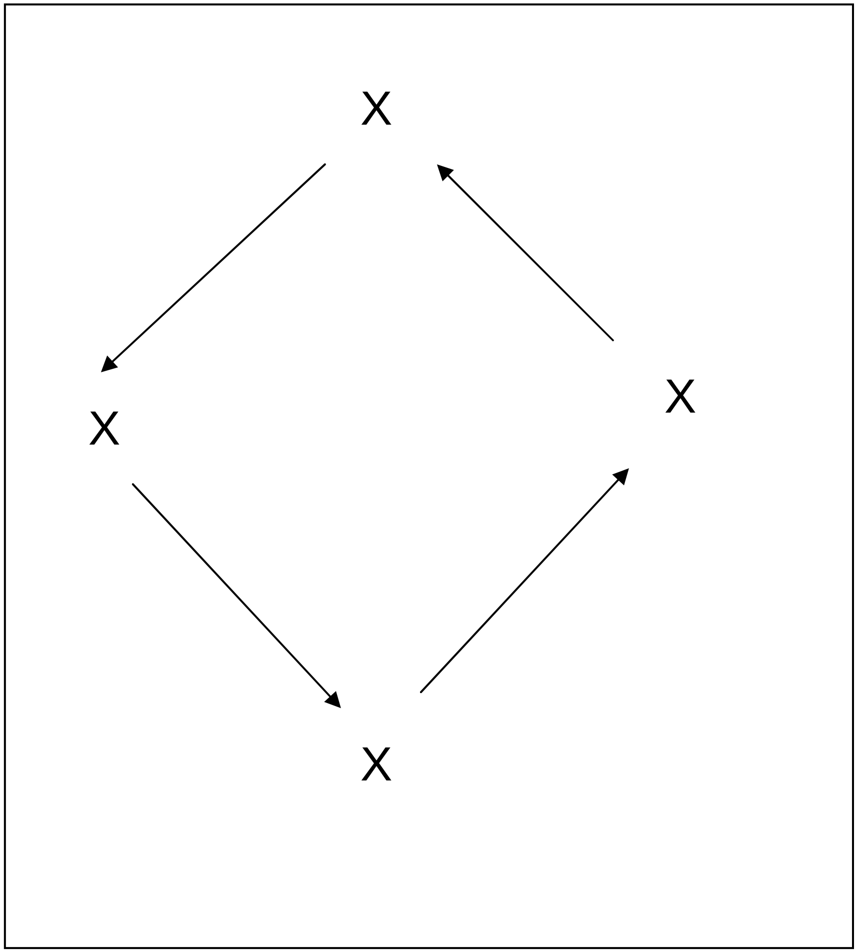

Figure 3: Several cycles