Understanding intraday credit in large-value payment systems

Ruilin Zhou

Introduction and summary

A large-value (or wholesale) payment system is a contractual and operational arrangement that banks and other financial institutions use to transfer large-value, time-critical funds to each other. It is operated either by a central bank or by a coalition of banks and financial institutions. In the U.S., the two largest and most commonly used large-value payment systems are Fedwire and CHIPS (Clearing House Interbank Payments System). Fedwire is a service provided by the Federal Reserve banks to about 10,000 depository institutions and government agencies for transferring funds and federal and agency securities. CHIPS is owned and operated by the Clearing House Interbank Payments Company L.L.C., whose members consist of 79 of the world's largest commercial banks.

The growth of the payment industry has been phenomenal. For example, the combined annual value of transfers through Fedwire and CHIPS was about 67 times U.S. gross domestic product (GDP) in 1988 and climbed to 89 times in 1999. The average *daily* value of transfers processed by the two systems reached $3,257.15 billion, about 35 percent of *annual* GDP, in 1999.

These staggering numbers give a glimpse of the increasing importance of large-value payment systems to the health and efficiency of the financial systems that our fast growing and ever more integrated national and international economies depend upon. The Federal Reserve, as well as other central banks around the world, has long recognized this importance, and continues to design and implement policies to ensure the proper and efficient operation of payment systems.

In this article, I examine the key issues in large-value payment systems and the optimal payment system design that addresses these issues. First, I discuss the main conflict faced by a large-value payment system—shortage of settlement liquidity versus potential credit risk—through the mechanics of the two main classes of payment systems. Customers of a real-time gross settlement (RTGS) system are constantly in need of liquidity to settle payments in real time, while those of a net settlement (NS) system face the uncertainty of potential settlement failure. The focus of this article is the remedy for liquidity shortage in a real-time gross settlement system, the provision of intraday liquidity by the central bank, and the policies designed to reduce the central bank's resulting exposure to credit risk. I describe three intraday-credit policies commonly used by central banks: a cap on an institution's net debit position during the day, an interest charge on the usage of intraday credit, and a collateral requirement to back the extension of intraday credit. In particular, I discuss the experience of Fedwire after the introduction of the first two policies.

The way a large-value payment system works and its key issues define the criteria for theoretical modeling, a common approach adopted by modern day economists to study optimal institutional design. I propose four main criteria that a payment system model should satisfy. First, it should directly model the underlying transactions of goods or financial assets for which payments have to be made so that the system design affects the allocation of real resources. Having met the first requirement, the model should treat consumption/investment debt, which is generated by the underlying real resources transaction, as distinct from payment debt, which is created only for payment needs. The final two criteria have to do with the two sides of the main conflict; the model should incorporate both settlement liquidity shortage and credit risk,

Ruilin Zhou is a senior economist at the Federal Reserve Bank of Chicago. *The author thanks Jeff Lacker (who suggested the topic), Ed Green, David Marshall, Jamie McAndrews, and Will Roberds for their generous help in clarifying issues and Liqian Ren for technical assistance.*

preferably generated endogenously by agents' choice or action.

The existing theoretical research on payment systems typically does not meet all four proposed criteria. It mostly focuses either on liquidity or risk and rarely models demand for settlement liquidity as a derived demand. Nevertheless, the literature provides significant insights into the merits of the three intraday-credit policies. I review this body of literature and find that articles focusing on credit risk as the main difficulty of a payment system tend to support the imposition of caps and collateral requirements, while those focusing on liquidity shortage support unconditional free provision of intraday credit by the central bank. There is one argument that supports the interest charge policy, but only as an outcome of another central bank policy, the non-interest bearing overnight reserve requirement.

I employ a model developed by Freeman (1996), which, to a large extent, satisfies all four measures[1] to explore some different insights into the payment problem. The original Freeman model is designed to study the use of open-market operations and the discount window in conducting monetary policy in environments with a liquidity shortage. To make the model better suited for settlement of payments, I assume that the underlying goods transaction is accomplished through a pairwise trade so that one debtor owes payment to one creditor, rather than each creditor holding a diversified portfolio of debt issued by different debtors, as in Freeman's model. To demonstrate explicitly the distinction between consumption debt and payment debt, I introduce money growth into the model, but maintain Freeman's assumption that there is no intertemporal real resources allocation opportunity within a day. Finally, I interpret the debt settlement market as a payment clearing market, resale of debt as private intraday borrowing/lending of reserves, and central bank injection of liquidity as intraday-credit lending rather than an open market operation or discount window lending.

My analysis leads to the following conclusions. First, when credit risk is not of concern, the interest rate (that is, the risk-free rate) on intraday credit should be zero. The overnight nominal interest rate, which serves to achieve efficient intertemporal allocation of real resources, is governed by the money growth rate.[2] Setting the intraday rate to zero is optimal because the sole purpose of intraday credit is to settle payments made for underlying transactions of goods and financial assets. Hence, the cost of intraday credit constitutes a transaction cost of the underlying goods/assets trade, and should be minimized to distort as little as possible the intertemporal allocation of real resources. For the same reason, any private provision of intraday liquidity through a market-like mechanism that generates a positive intraday risk-free rate is inefficient. Second, when a particular type of credit risk is under consideration, namely, aggregate default risk in the same model,[3] the free provision of intraday liquidity by the central bank remains the optimal policy (Freeman, 1999). Such a policy may inadvertently lead to price fluctuation and inflation because when some agents default on their payments, the central bank's temporary injection of settlement liquidity becomes permanent. Despite this potential side effect, free lending acts as an insurance mechanism that transforms the default risk, which would have been borne disproportionately by a subgroup of payment system users, to inflation risk that affects everyone with much less severity.

The conclusions of the model depend critically on the assumption that there is no need or opportunity to optimize the timing of consumption and production over the course of a single "day." A "day" can be interpreted as any length of time. The results will hold for any time interval for which the assumption remains valid. In the actual economy, despite the common practice of end-of-business-day settlement in the majority of payment systems, identifying the appropriate length of time for which one can reasonably claim that the assumption is valid is beyond the scope of this article. As communication technology linking financial markets in time zones around the world continues to advance, the length of time for which the assumption remains valid will presumably decline. Along the course of this development, policies to prevent credit intended for payment services from being used for short-term investments and speculations, including those that shorten the settlement period, may need to be considered.

How payment systems work

A fund transfer from bank A to bank B is usually accomplished in the following fashion. First, the sending bank A initiates the transfer by sending a payment message regarding the impending transfer to the payment system, which after processing is delivered to the receiving bank B. Then, either immediately or at some fixed time after the payment information is processed, the actual settlement occurs. The conventional means of settlement of large-value funds transfer systems is central bank funds (base money). So at the settlement stage, bank A's reserve or clearing account at the central bank is debited and bank B's account is credited. A settlement is final if the funds received by bank B are irrevocable (except in cases involving criminal fraud).

According to the way settlement takes place, a payment system can be classified into a gross settlement system (GS) or a net settlement system. With a gross settlement system, fund transfers at the settlement stage occur on a bilateral, gross (that is, transaction by transaction) basis. A common form of GS large-value payment system is the real-time gross settlement system, at which both the information processing and settlement take place continuously in real time. With a net settlement system, payment messages are processed continuously in real time, but settlement occurs only at the end of a clearing cycle, on a net debit/credit, bilateral, or multilateral basis.

RTGS versus NS systems

For a given payment system, the parties involved in a funds transfer, the sending bank and the receiving bank, face different problems. To understand this, let us consider the complete life cycle of a funds transfer. Suppose a manufacturing company I purchases $10 million worth of computer equipment and services from company II. The contract stipulates that on the day all purchased equipment and services are delivered, say, August 1, 2000, company I makes its $10 million payment to company II. On August 1, 2000, after verifying the delivery of its purchase, company I instructs its bank, bank A, to send $10 million to company II's account at bank B. Bank A may make the funds transfer right away upon request *if it is able to* or it can delay it if the contract permits. Under an RTGS system, bank B will receive the transfer with finality as soon as bank A's payment order is processed. Under an NS system with end-of-business-day settlement, bank B receives the message of $10 million incoming transfer once the system accepts bank A's payment order, but will know for sure whether the transfer is settled with finality only at the end of the day. This is just one outgoing funds request for bank A, and one incoming funds transfer for bank B during the day. Potentially, both banks may receive many such payment orders throughout the day. In general, a bank has little control over the arrival of its customers' outgoing payment requests, whether they are urgent (time-sensitive) requests, and the flow of its incoming funds transfers (which depend on other banks' timing decisions of payment initiation). For these reasons, contracting a precise-time funds transfer between companies I and II may be very costly. The end-of-business day settlement of a transaction is often the convention, possibly the best that can be accomplished.[4]

Under an RTGS system, bank B, the receiving bank, enjoys the real-time settlement finality; the $10 million, once received, can be used to cover outgoing payments the rest of the day without any uncertainty. Bank A, however, faces the problem of when to send the payment request. This decision depends on whether bank A has sufficient funds in its reserve or clearing account to cover the transfer, when it is expecting the arrival of incoming funds, and whether it should save the account balance for more urgent payment requests. The timing decision of every bank using the payment system may collectively slow down the speed of funds transfers or may even trigger gridlock of the whole system (in the case of two banks, this is a situation where bank A is waiting for bank B's payment and bank B is waiting for bank A's payment, so neither can pay the other). The concern for whether there will be a sufficient account balance to cover outgoing payments demand may raise the level of precautionary reserves that each bank holds (above reserve requirements), given the uncertain demand for payment. Therefore, the need for settlement liquidity in real time may be very costly not just for the funds-sending bank, but for the payment system as a whole.

Under an NS system with end-of-business-day settlement, the funds-sending bank A does not have the above concerns. The payment is settled only at the end of the day, at which time, it would have received all of the day's incoming funds, as well as having made all the outgoing transfers. It pays bank B and other banks the net amount it owes at the closing of the business day. In other words, for this particular transfer, bank A receives an implicit extension of $10 million interest-free intraday credit from bank B between the time the transfer is initiated and the time the net balance owed to bank B is paid. The end-of-day payment implies that sending banks, including bank A, have no incentive to delay sending the payment messages if there are no other payment-system-imposed constraints. Hence, there should be no costly delays or gridlock. Also, since each bank needs to pay only the net amount at the end of a day, which usually is a lot smaller than the value of total outgoing payments the bank has to make during the day, it needs to hold lower reserves or clearing balances as payment liquidity than under an RTGS system. On the other hand, bank B, the receiving bank, may face significant credit risk. If bank A fails during the day and can not make the payment at the end of the day, bank B may have to bear at least part of the loss, and, moreover, bank A's inability to settle may trigger the unwinding of the whole day's payments. The potential spillover of this settlement failure to other payment systems and financial markets, often termed systemic risk, is considered very costly.

From the above discussion, we understand that the main difficulty for an RTGS system is the provision of costly settlement liquidity in real time, while the difficulty for an NS system is the potential credit risk. For commercial banks and financial institutions, the everyday needs of settlement liquidity outweigh the risk of settlement failure, which is a possibility though it rarely happens in practice. Hence, private payment arrangements are often NS systems.[5] On the other hand, the increasingly integrated international economy, including financial markets and payment systems, intensifies the concern of many central banks over potential systemic risk. The recent technological advances in real-time monitoring and processing of financial transactions also make the implementation of RTGS systems easier. These factors facilitate the recent movement toward RTGS as the favored large-value payment system of central banks in many countries.[6]

To ease the shortage of settlement liquidity under an RTGS system, many central banks provide intraday liquidity with certain restrictions. That is, instead of waiting for the arrival of incoming funds to cover outgoing payments, a sending bank without a sufficient account balance can make a payment by borrowing from the central bank during the day and paying it back before the end of the day. This arrangement effectively turns an RTGS system into a netting-with-the-central-bank system. This is because the real-time settlement of a funds transfer using intraday credit is an initial settlement between the central bank and the receiving bank (after which the sending bank's debt to the receiving bank is owed to the central bank), followed by another settlement between the sending bank and the central bank at a later time, possibly with the sending bank's incoming funds or a net payment at the end of the day.

Intraday-credit policy

The extension of intraday credit by the central bank effectively transfers the credit risk from the receiving bank to the central bank. To reduce the potential credit risk posed to the central bank by the allowance of intraday credit, some form of explicit measure to control the use of intraday credit is often adopted, in addition to an intensified effort to monitor and control banks' financial situation and risk management. These policies include

1) the imposition of a quantitative limit (or "cap") on the amount of intraday credit that each bank can receive at any moment of a day,
2) charging an interest rate (though not necessarily the market rate) to discourage the improper usage of intraday credit, and
3) the requirement of collateral or intraday repos[7] to fully or partially back the amount of credit extended.

All three measures impose costs on the use of intraday credit. The potential punishment for violation of the cap (if violation is allowed) or the inability to borrow from the central bank above the cap (when violation is not allowed) are costly to banks. The interest charge is an explicit proportional cost for the usage of intraday credit. Collateral or repos carry an opportunity cost if the amount of qualifying safe assets required for adequate settlement liquidity is more than the amount a bank would hold without the requirement.

Different central banks adopt different intraday-credit policies for their RTGS systems. The TARGET system, which is a collection of inter-connected domestic RTGS systems of European Monetary Union (EMU) member countries that settle cross-border payments denominated in euros, mandates that each member central bank provides interest-free intraday credit on a fully collateralized basis. Switzerland, as a non-EMU member, had an extreme form of intraday-credit policy for its interbank funds transfer system, Swiss Interbank Clearing (SIC): no intraday provision of settlement liquidity by its central bank under any condition.[8] As a substitute, there is a very limited intraday money market for special time-critical payments in connection with securities transactions (BIS, 1997). Only in October 1999, the Swiss National Bank started to allow intraday repos-backed interest-free overdrafts in an effort to make its payment system more compatible with the EMU countries. In the U.S., Fedwire adopts the other two risk-management measures, the intraday overdraft cap and the interest charge.

The experience of Fedwire

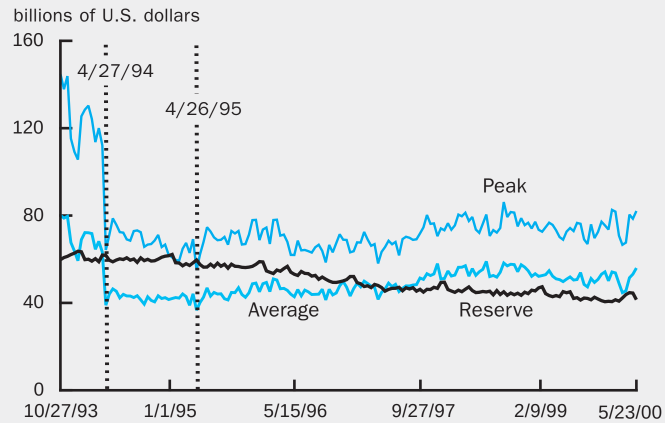
The Federal Reserve Banks used to have a quite liberal intraday-credit policy, with almost no restriction on the use of intraday credit by depository institutions. In 1986, the Federal Reserve moved toward a more cautious approach in its extension of intraday credit. It began by imposing a quantitative limit on the total amount of intraday credit each depository institution could incur for funds transfer over Fedwire and other private large-value payment systems (such as CHIPS). This cross-system limit was replaced by net debit caps on Fedwire alone in 1991 (CHIPS maintains its own net debit caps established by its participants, separate from those on Fedwire). Currently, each depository institution is subject to two capital-based net debit caps for overdrafts related to funds transfer and book-entry securities transfer:[9] a daily cap that limits the amount of intraday overdrafts the institution can incur at any moment in its Federal Reserve account and a two-week cap that the average overdraft by the institution over a two-week period should not exceed. Studies show that although the initial imposition of the net debit caps on funds-transfer-related overdraft may have restricted the growth of overdrafts (Richards, 1995) and forced a few heavy users of daylight overdrafts to improve their liquidity management, the overall effect of imposing caps on intraday credit has not been significant (Hancock and Wilcox, 1996).

In 1994, in an effort to intensify its control of intraday credits, the Federal Reserve imposed an explicit minute-by-minute interest charge of 24 basis points (annual rate) on the average daylight overdraft each institution incurred during a business day in addition to the net debit caps.[10] The rate was raised to 36 basis points in 1995. This interest charge is levied with deductibles. For ten hours each day, overdrafts valued at 10 percent of an institution's risk-based capital are exempted from the charge. In addition, any two-week total charge less than $25 is waived. Because of these deductibles, many institutions do not pay anything under the new policy. In fact, using data on aggregate fees collected and average overdrafts, the imputed effective average (not marginal) annual rate was only around 8 basis points before the raise in April 1995 and around 11 basis point after that. In 1999, the *average per minute* overdraft on Fedwire was on the magnitude of $50 billion, while the aggregate fee collected over a *two-week* period was only around $1 million.

Despite the low fee and the deductibles, the impact of the initial interest charge was significant, although the subsequent rate increase had no obvious effect. During the six months immediately following the imposition of the fee, both intraday peak and average overdrafts declined by about 40 percent, with security-related overdrafts decreasing more (45 percent) than the funds-related overdrafts (25 percent). According to Richards (1995), the reduction in intraday overdraft is driven by the reduction of large overdrafts: More than 90 percent of the drop in intraday overdraft comes from the top six overdrafting institutions. Figure 1 shows the Fedwire intraday peak overdraft and average overdraft for both funds transfer and book-entry security accounts from October 1993 onward.[11]

FIGURE 1

Total intraday peak overdrafts, average overdrafts, and reserves



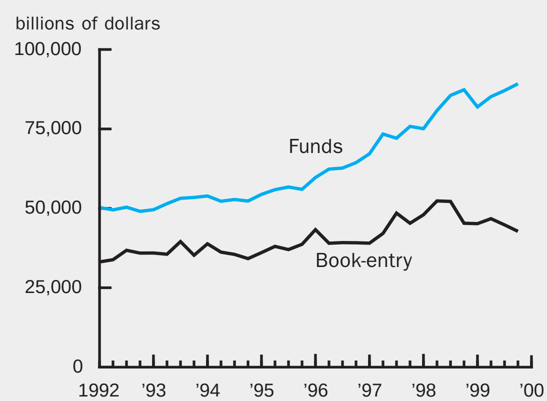
Notes: Intraday peak overdraft is the greatest value of the total intraday overdraft by all institutions at the end of each minute of the day for a given day. Average overdraft is the sum of the average per-minute intraday overdrafts of all institutions on a given day.

Sources: Data for Fedwire overdrafts are available at Federal Reserve Board, "Data: Daylight overdraft and related fees, April 27, 1994, to present," at www.bog.frb.fed.us/PaymentSystems/PSR/default.htm. For similar statistics before August 1993, see Richards (1995) and Hancock and Wilcox (1996). Data for reserves are from Federal Reserve Board, "Statistical release weekly historical aggregate reserves—Aggregate reserves of depository institutions (not adjusted for changes in reserve requirements, not seasonally adjusted)," table 5, at www.bog.frb.fed.us/releases.

Despite the significant impact of the interest charge on institutions' overdraft behavior, it has little effect on the amount of transactions processed over Fedwire. Figure 2 shows the total value of transactions made on both the funds transfer and book-entry security accounts. In other words, the imposition of the fee at its current level does not discourage transfer

FIGURE 2

Value of transfers originated



Sources: Data from Federal Reserve Board, "Data on transfer of funds and book-entry securities, quarterly," at www.bog.frb.fed.us/PaymentSystems/FedWire/default.htm.

activities over Fedwire. It affects only the timing of the transfers as institutions try to reduce the amount of overdrafts. This is evident from the apparent attempts by banks to finance a higher proportion of outgoing payments with incoming funds and utilize account balances more efficiently by delaying sending payment orders (Richards, 1995). McAndrews and Rajan (2000) find increased coordination among participating banks and conjecture that they synchronize payment activities through establishing regular times for funds transfers.

In summary, there are two basic models of intraday-credit policy for an RTGS system in use: the European model that allows interest-free intraday credit on a full collateral backed basis and the U.S. model with an intraday-overdraft cap and explicit intraday-credit pricing. Both models directly limit the central bank's exposure to credit risk due to the provision of intraday credit. Questions remain as to whether either model or any combination of the three measures outlined earlier or other measures serve the central bank's objective of promoting an efficient payment system while containing risk.

Modeling payment systems

Now that we understand how payment systems work, we can study the optimal design of a payment system by modeling the fundamental conflict of liquidity versus risk in an environment that incorporates most of the essential features of a modern payment problem. More specifically, a payment system model should satisfy the following criteria. First, it should model the underlying transaction of goods or financial assets for which payment has to be made in a different time (that is, a debt has to be issued and settled at different times). The choice of the payment system used to settle the payments matters, in the sense that it affects the underlying resources allocation. Second, there should be a distinction between consumption/investment debt and settlement debt if both are modeled. The former is created when the underlying trade of goods or assets is conducted, and the latter is generated when settlement liquidity is borrowed in order to settle the associated consumption/investment debt. This distinction will enable one to study the property of settlement debt independently of the underlying consumption/investment debt. Third, the model should have an endogenously generated settlement liquidity shortage, possibly derived from a payment structure in which settlements of different parties are interdependent and may even induce settlement gridlock. The shortage of liquidity makes borrowing and lending intraday settlement liquidity necessary. Last, the model should include the risk component: the possibility of settlement failure that could be triggered by genuine bank failure (for example, investment failure) or by moral hazard induced by the intraday-credit policy (for example, overuse of intraday credit or change in portfolio choice).

These four criteria are tall orders to fill. A substantial amount of theoretical research focuses either on costly settlement liquidity or on credit risk, rarely both. Most models ignore the reality that the demand for settlement liquidity is a derived demand for underlying trade of goods and financial assets (criteria 1 and 2). Despite their problems, these studies provide significant insights into the workings of different payment systems and intraday-credit policies. I survey this body of literature before discussing a model that satisfies the four criteria and the insights it provides.

Some theoretical arguments

It is generally argued that either collateral or a debit cap is required to limit the central bank's exposure to credit risk. The debate often centers on whether settlement liquidity should be allocated through a market-like mechanism, such as price. A generic argument for market allocation of settlement liquidity postulates that settlement liquidity is a resource, and by standard economic theory, efficient allocation of any resource can be achieved through a market mechanism (see Mengle et al., 1987, and Evanoff, 1988). The demand for intraday credit is assumed to derive from the fundamental difficulty of synchronizing payment flows; hence, having access to intraday credit would reduce settlement cost (such as excess holdings of reserve balances for settlement and the need for potentially costly precise-timing contracting). On the supply side, it is argued that the providers of settlement liquidity should be compensated for its cost, which includes the pure time cost of funds and the compensation for risk. The value of settlement liquidity to both sides of the market gives rise to the standard demand and supply and, hence, market clearing price. This argument is plausible heuristically. The challenge is to model explicitly the elements that determine the demand and supply for settlement liquidity and to evaluate the argument in a rigorous way. Some of the following arguments are derived from such explicit modeling.[12]

1. Charging a nominal overnight rate on intraday overdraft corrects the distortion created by the non-interest-bearing reserve requirement. This theory derives the value of intraday-credit pricing from the existence of another distortionary policy. Lacker (1997) argues that the central bank's standard policy requiring depository institutions to maintain a reserve balance with no interest paid is, in fact, an inflation tax on reserve balances. In a model where banks face a positive overnight interest rate but a zero intraday interest rate, the wedge between the two interest rates reinforces the distortionary reserve requirement. This is because there is no need for banks to hold overnight balances for the next day's payment needs given that they can borrow at zero interest rate (assuming there is no intraday-borrowing constraint). The requirement to hold a reserve balance overnight, and hence the foregone overnight interest on it, distorts banks' intertemporal resource allocation. If intraday liquidity is also priced at the overnight rate, then banks need to make provision for payment liquidity by either holding an overnight account balance with the opportunity cost of the interest rate or borrowing intraday at the same rate. For banks whose payment liquidity needs are at least as large as their reserve requirements, holding overnight balances equal to or above the requirements (which they are indifferent from borrowing intraday) renders the distortionary reserve requirement non-binding. For banks whose payment liquidity demand is smaller than the reserve requirement, the distortion created by the non-interest-bearing reserve requirement cannot be completely eliminated. However, it is not clear, due to the potential general equilibrium effect, that banks (even those with large payment liquidity needs) would prefer to pay the marginal cost of financing payment liquidity and not suffer the distortion brought about by the reserve requirements or vice versa.

2. Costly monitoring of borrowing banks is necessary, and requires compensation. Rochet and Tirole (1996) focus on the risk component of the cost to the supplier of intraday liquidity, and argue that the primary problem of a payment system is solvency, not liquidity. They speculate that in a world where banks were perfectly safe, a bank could get liquidity instantaneously since an intraday market would emerge if the cost of precise-time contracting was too high. Given that banks are not perfectly safe, the solvency of a borrowing bank requires monitoring by its lender. Hence, intraday lending should be costly, not free. Although Rochet and Tirole do not provide a framework for measuring the cost of monitoring, they do argue that a quantitative cap or a reasonable level of collateral requirement may be a better means to control the overuse of intraday credit than pricing. They argue that pricing may induce moral hazard, by increasing borrowers' failure rate, or adverse selection, by eliminating banks with low nonobservable risk and serving those with high nonobservable risk and, hence, a higher probability of failure.

3. Free intraday liquidity may encourage banks' risk-taking behavior. Kahn and Roberds (1999b) show that under an NS system in which intraday liquidity is free, banks may choose a more risky asset portfolio than they would under an RTGS system without the provision of intraday liquidity. This is because in an environment where each bank exists for only one period, say one day (hence, there is no need to consider the effect of its action in the long run), it is optimal to default (not settle) net payment at the end of the day. By doing so, losses from risky investment are shifted to other participants of the NS system (or to the central bank under an RTGS system with the central bank providing free intraday liquidity). Under an RTGS system without the provision of intraday liquidity, payment orders have to be settled with reserves or liquidation of safe assets as they are realized throughout the day. No strategic default at the end of a day is possible. In such an environment, the remedy for liquidity shortage is not charging interest on the net debit position, which would give more incentive for default, but imposing net debit caps or requiring collateral. The latter also dominates RTGS without liquidity provision given that the safe assets do not have to be liquidated as collateral.

4. The extension of free intraday credit eliminates inefficiency brought about by intraday liquidity constraints. It is well understood that the creation of inside money (debt) can sometimes improve intertemporal resource reallocation when agents face liquidity constraints with outside money alone. Kahn and Roberds (1999a) reinterpret consumption as funds transfer and payment with debt securities as payment with the central bank's intraday credit. In their model, the free extension of the right amount of intraday liquidity can eliminate the liquidity shortage and restore the first-best consumption allocation (that is, the allocation achievable when there is no liquidity constraint), while charging interest on intraday credit is distortionary. Furthermore, some combination of inflation and partial collateral requirements can also achieve the first-best consumption allocation.

5. The provision of free intraday liquidity reduces the possibility of holding a "sterile" reserve. This argument again relies on the central bank's usual practice of a zero-interest reserve requirement, making banks' above reserve-requirement balances "sterile." Kahn and Roberds (1999b) model the arrival of payment orders for a bank as completely stochastic; facing this flow, banks make beginning-of-the-day reserve and portfolio decisions. Under an RTGS system without the provision of intraday liquidity, a bank makes a payment either with its reserve balance or by liquidating asset holdings at a cost. The random nature of the payment flow implies that a bank may end up with a positive balance. Under an NS system (or an RTGS system with free intraday credit), the expected reserve is usually smaller than under an RTGS system, since only the end-of-day net credit/debit positions, which are often smaller than gross payments, need to be settled. Therefore, banks face a smaller chance of holding "sterile" reserves at the end of the day. Lacker (1997) makes a similar point.

6. Costly intraday liquidity may induce banks to delay sending payment orders, which generates a negative externality. Angelini (1998) introduces an exogenous cost structure for delaying payments in a model where payments among banks are interdependent. Given that intraday credit is costly (either because of pricing or collateral requirements), while making a payment sending/withholding decision, a bank faces the tradeoff between sending the payment order promptly by borrowing costly intraday credit or delaying the payment and suffering the delay cost. In such an environment, if banks cooperate to maximize joint profit, there will be no delay (no payment order is blocked by other banks' delayed payment). In a noncooperative equilibrium, however, a bank will delay a payment order to reduce its expected intraday-overdraft cost and wait for the incoming funds to arrive. By doing so, it transfers the intraday-credit cost to the payment-receiving bank. The negative externality generated by this delay is a dead-weight loss to the payment system, and it cannot be eliminated by the existence of the intraday money market because liquidity on the intraday market will also be costly. Kobayakawa (1997), in a similar setup, shows that while intraday-credit pricing induces delay, collateralized intraday credit does not, although it imposes other costs on banks. Humphrey (1989), on the other hand, argues that delays in sending less time-critical payments will improve reserve efficiency. As I mentioned in the discussion of Fedwire's intraday-credit policy, there is evidence both of banks delaying sending outgoing payments and of banks cooperating in making payments.

Among the above six arguments, only the first one supports market-based intraday liquidity pricing; and the argument holds only in conjunction with the existence of the distortionary non-interest-bearing reserve requirement. When credit risk is under consideration, argument 2 supports monitoring-cost based pricing, and 3 supports net debit caps and collateralization to control banks' risk-taking behavior. For a pure liquidity shortage concern, the last three theories support the provision of free intraday credit.

A model without settlement risk

Next, I discuss a version of a payment system model that fits the bill of the proposed theoretical framework, introduced by Freeman (1996). The Freeman model is intended to study the central bank's means of conducting monetary policy, via open market operations and the discount window, in an economy with a liquidity shortage. When applied to a large-value payment system that lacks intraday liquidity, the model offers different insights about the provision of intraday credit. To separate the problems of liquidity shortage and settlement risk, I first discuss a version of the model that only has a shortage of settlement liquidity; then I explore the effect of introducing credit risk. In the appendix, I solve a parametric version of the model without settlement risk.

The model is a standard overlapping generation model with added features to satisfy the first three criteria (that is, modeling the underlying transaction of real resources, distinguishing real debt and settlement debt, and incorporating a settlement liquidity shortage).[13] There are a large number of two-period-lived agents, one generation born in each period. Each generation has an equal number of creditors and debtors. They are so named to anticipate the roles they will play in their lifetime. There are two nonstorable goods, the *C*-good and the *D*-good, endowed to the young generation each period. At the beginning of a period, a young creditor receives one unit of the *C*-good, and a young debtor receives one unit of the *D*-good. In addition, the initial old creditors, who live only one period, are endowed with m_0 units of money per person. Debtors consume both goods only when young, while creditors prefer to consume the *C*-good when young and the *D*-good when old. All agents are risk averse.

This preference and endowment pattern leads to both intra-generation and inter-generation trades. More specifically, at any date, young debtors would want to purchase some *C*-good from young creditors, and old creditors would like to consume some of young debtors' endowment, the *D*-good. Suppose the former trade (intra-generation) occurs in the morning and the latter (inter-generation) takes place in the afternoon. When young debtors meet young creditors at the *C*-good market in the morning, they have no money, and have only their endowment, which the young creditors do not consume. The only way the two parties can trade is if young creditors accept young debtors' personal IOUs (a promise to pay a certain amount of fiat money tomorrow for the goods purchased today) as payment for goods.[14] To make the model more like a payment problem, given that there are an equal number of ex ante identical debtors and creditors, I assume that the debt for *C*-good transaction is bilateral, that is, each creditor holds one debtor's IOU after the trade.[15] Assume that all agents are able to issue nonfalsifiable, verifiable personal IOUs. Debtors pay back their creditors with money next morning in a central clearing market. By then, they should have obtained the fiat money necessary to settle their debts. Assume the central clearing market is operated by an infinitely lived central bank that has the authority to issue fiat money[16] and to enforce the settlement of debt contracts in the market. In the afternoon, old creditors use the money they received in the morning (debt payment) to purchase the *D*-good from young debtors at the *D*-good market.

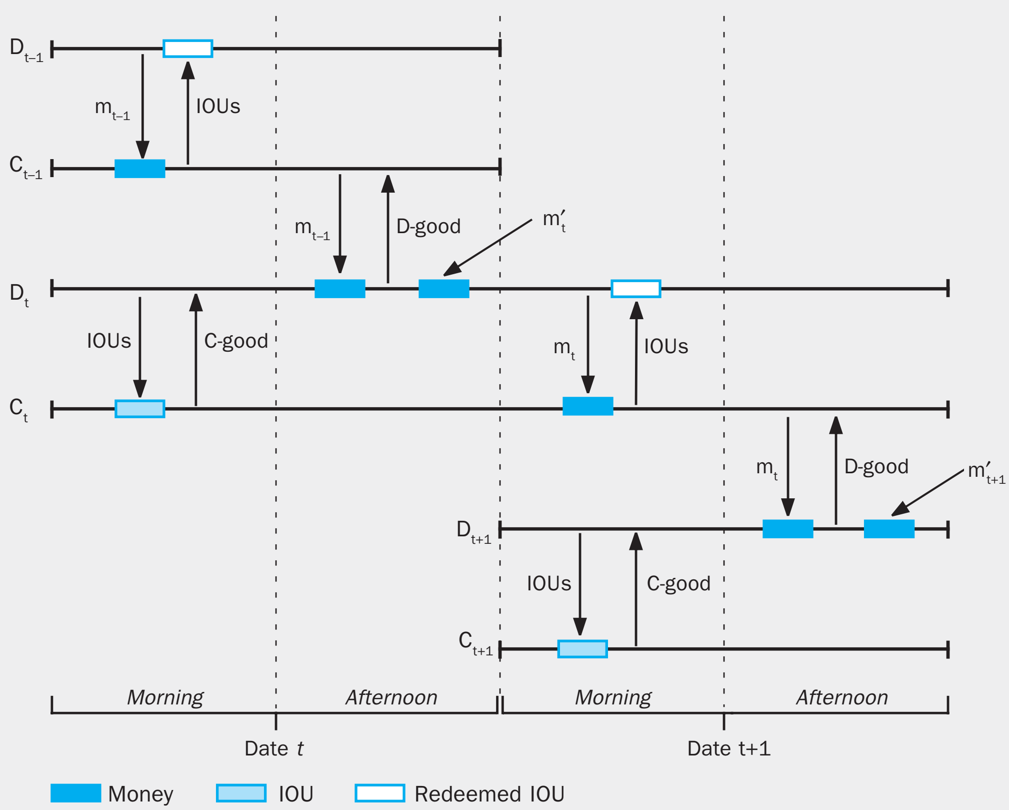
To generate a positive, overnight nominal interest rate, suppose that each young debtor receives a lump-sum transfer of fiat money at the end of a day. The money growth is exogenous, and the growth rate is $i > 0$. At the end of a day, all fiat money will be in the hands of young debtors, which they use to pay their debt next morning. The timing of different markets and the trading flows within and across generations are illustrated in figure 3.

Over the course of a lifetime, a creditor who wants to consume when old saves by selling a portion of her nonstorable endowments in exchange for debt when young and settles the debt for money, with which she purchases her old age consumption. A young debtor, on the other hand, first buys goods with personal IOUs, and then sells his endowment for money. He is alive in the second period of his life solely for the purpose of repaying his debt. Given that all debtors (creditors) are risk averse and ex ante identical, economic efficiency requires that all debtors (creditors) of each generation consume the same amount of their desired consumption goods.[17]

Without any settlement problem, with standard preferences (increasing and concave utility function), young and old agents of all generations will consume

FIGURE 3

Timing of markets



Notes: D_t is the generation-*t* debtor; C_t is the generation-*t* creditor; m_{t-1} indicates a debtor's money holding at beginning of *t*; m'_t indicates per debtor lump-sum money transfer at end of *t*; $m_t = m_{t-1} + m'_t$.

a constant portion of their desired goods, respectively. The prices of both goods grow at rate i, and the overnight nominal interest rate on the debt is also i. This outcome is efficient.[18] I call it *equilibrium* $(*)$.

To satisfy the third criterion (incorporating a shortage of settlement liquidity), I introduce a settlement problem at the central clearing market every morning. Suppose that the payment flows are not fully synchronized. When the clearing market opens, all creditors arrive, but only a fraction, say $\lambda \in 0[0,1]$, of debtors arrive. Before the remaining $1 - \lambda$ fraction of the debtors arrive, $1 - \alpha$ fraction of old creditors have to leave, $\alpha \in [0,1]$. For an individual agent, the timing of his or her arrival and departure (early or late) is completely random and is realized only before the settlement. An old creditor may fall into one of three categories:

X: the debt she holds is settled at par because

X_1: she leaves early and her debtor arrives early,

X_2: she leaves late and her debtor arrives late,

Y: she cannot settle directly with her debtor because she leaves early and her debtor arrives late, or

Z: she leaves late but her debt is settled directly with her early-arriving debtor.

The probabilities associated with categories X, Y, and Z are $\alpha + \lambda - 2\alpha\lambda$, $(1 - \lambda)(1 - \alpha)$, and $\alpha\lambda$, respectively.

The latter two groups of creditors can trade since group Y have unredeemed IOUs and have to leave early, while group Z receive their payment money and can wait for more debtors to arrive. Depending on whether they exchange money for debt, group Z can be divided into two subgroups: they either

Z_1: do not exchange repayment money for debt, or

Z_2: purchase debt with repayment money.

Figure 4 illustrates the timing and trading patterns among different groups of debtors and creditors on the clearing market.

Whether group Y creditors will be repaid in full for the debt they accepted in the previous period depends on the relative sizes of groups Y and Z. If there are more agents in group Z than in group Y, $\alpha\lambda \geq (1 - \lambda)(1 - \alpha)$ (or equivalently, the amount of debt held by early-leaving creditors is smaller than the amount of money brought in by early-arriving debtors, $1 - \alpha \leq \lambda$), then a portion of the creditors in group Z, group Z_2, purchase unredeemed debt from those in group Y at par (since the demand for debt is greater than the supply) and settle the purchased IOUs with late-arriving debtors. In this case, the asynchronization of payment flows does not create any problem. All creditors are repaid in full, and the consumption allocation is the same as in equilibrium $(*)$.

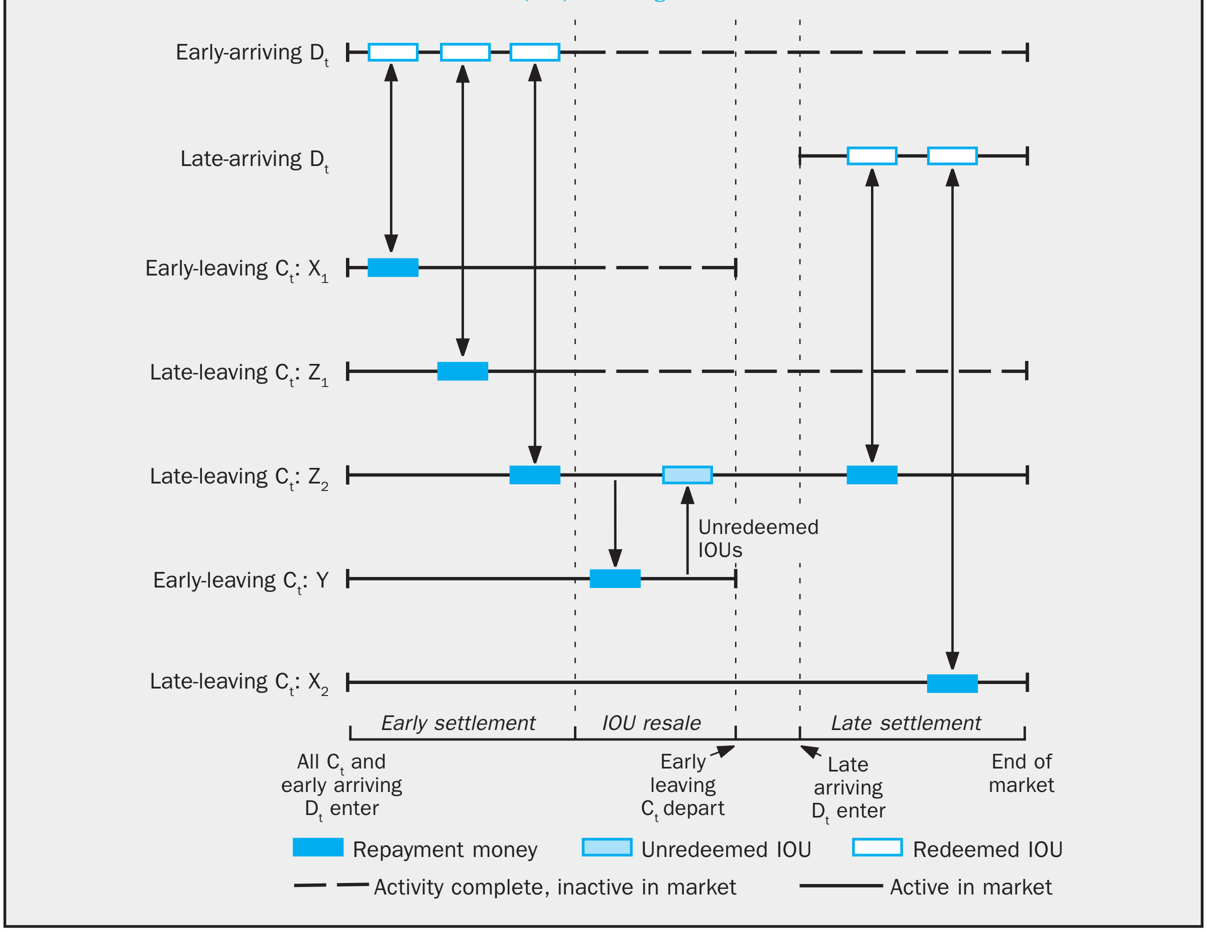
It is also possible that there are fewer agents in group Z than in group Y, $\alpha\lambda < (1 - \lambda)(1 - \alpha)$ (or equivalently, there is more debt than money available for settlement, $1 - \alpha > \lambda$). In this case, all creditors in group Z purchase unsettled debt from group Y creditors at a discount (that is, $Z_2 = Z$), each obtains more debt than her debt holdings before the settlement, and the repurchased debts are settled when the late-arriving debtors arrive. The amount discounted depends on how much smaller group Z is relative to group Y, in other words, the severity of the settlement liquidity shortage. Because the debt is discounted, group Y creditors receive less money and group Z creditors receive more money than they were originally promised. Group X creditors are unaffected since their debts are settled at par. The uneven distribution of fiat money among old creditors leads to an uneven allocation of consumption goods. Relative to old creditors in equilibrium $(*)$, group Y old creditors consume less, group Z consume more, and group X consume the same. Such an outcome is inefficient; group Z creditors benefit at the expense of group Y, despite being ex ante the same.

The inefficiency can be easily corrected if the central bank, which runs the clearing market, buys up the unredeemed debt at par with newly issued money directly from old creditors in group Y and then destroys the fiat money turned in by late-arriving debtors for repaying their debt. By doing so, the central bank temporarily increases the amount of fiat money in the economy intraday, but does not change the aggregate money supply overnight. Hence, the action does not alter the inflation path on both the C-good and the D-good market.[19] Since all creditors are able to redeem their debts at par, the consumption allocation is the same as in equilibrium $(*)$.

The trading of debt between early-leaving and late-leaving creditors at the central clearing market can be interpreted as the operation of a private interbank market trading reserves in a large-value payment system. Banks that receive payments early (group Z) extend credit to the banks that cannot pay early (late-arriving debtors), so that banks in urgent need of funds (group Y) can be paid in time. This lending transforms an overnight consumption debt (from group Y creditors to late-arriving debtors) into an intraday settlement debt (from group Z creditors to late-arriving debtors), which is settled later during the day. When debt is traded at par, the intraday lenders (group Z creditors) do not gain anything, hence the lending is free.

FIGURE 4

Date (*t*+1) clearing market



However, when debt is traded at a discount, the additional fiat money obtained by group *Z* creditors can be viewed as an interest payment for intraday lending.[20] The results above suggest that economic efficiency requires free intraday settlement lending, that is, an intraday interest rate of zero. The central bank's temporary injection of settlement liquidity at no cost to settlement parties accomplishes this goal.

Potentially, banks that receive incoming funds early can lend to banks in urgent need of settlement liquidity for a few hours or minutes during the day. The analysis suggests that the shortage of settlement liquidity under an RTGS system may not be completely resolved by the development of such a private intraday lending market. Given that there is very little intraday consumption or investment opportunity,[21] the settlement liquidity should be provided at zero interest rate. When private borrowing and lending of reserve balances is not able to achieve this objective, the central bank, which is the sole issuer of settlement money (base money), should step in and provide the needed liquidity free of interest. Fedwire may be such a case; figure 1 shows that Fedwire's peak intraday overdraft is above the total reserve balance and, for the last three years, even the average intraday overdraft has exceeded the reserve balance. That is, had a private intraday market substituted for the Federal Reserve's role of intraday liquidity provision, *ceteris paribus*, the outcome would be inefficient.

Introducing credit risk

In the model discussed above, there is no uncertainty regarding whether a debtor will repay his debt at full value; the only question is when he will arrive at the clearing market. Under such a setup, the optimal intraday-credit policy is to provide settlement liquidity free of charge. An obvious question is what if there is a possibility that a debtor does not repay his debt.

To answer this question, Freeman (1999) assumes that with some probability θ, a fixed fraction of debtors default on their debt and spend the repayment money on the *D*-good when old, $\theta \in [0,1]$.[22] The uncertainty of whether some debtors will default is resolved only after the late-arriving debtors show up at the clearing market, not before. Hence, the default risk is borne solely by the late-leaving creditor. In such an environment, Freeman shows that if the central bank is willing to tolerate price fluctuation, the free provision of intraday liquidity required to settle all debts at face value is still optimal. The reason, however, is for optimal risk sharing. For simplicity, I discuss the intuition of this result in the original Freeman model without any modifications (such as the exogenous money growth I imposed above).[23]

As I stated earlier, economic efficiency in this model environment requires that the allocation minimizes agents' ex post consumption difference, in particular, that of the creditors since they are the ones who suffer from the problem caused by asynchronized payment flows as well as the default risk. If the central bank does not help to settle late-arriving debtors' and defaulters' debt, creditors are divided into four groups (compared with three when there is no aggregate default risk) when default occurs:

a) creditors who are repaid at full value (these include both early-leaving and late-leaving creditors whose debtors do not default),

b) early-leaving creditors who have to sell their unredeemed debt to late-leaving creditors who have been repaid at a discount (to reflect the potential default risk),

c) late-leaving creditors who purchase debt from early-leaving creditors and are able to redeem the debt later, and

d) late-leaving creditors whose debt holdings are not redeemed because of default.

Among these four groups of agents, group c receive the most amount of money, hence the highest consumption; they are followed by groups a and b; and group d creditors consume nothing when old.

This allocation can easily be improved. One solution, though not the only one, is to have the central bank redeem all creditors' unpaid debt at par, including both the unsettled debt of early-leaving creditors and the defaulted debt of late-leaving creditors, and then take out an equal amount of fiat money repaid by debtors *whenever possible*. By doing so, all creditors receive exactly their promised payments in fiat money. When default actually occurs, the settlement liquidity injected can not be completely taken out, and hence, the goods price will inflate. But such inflation is felt by all creditors equally. When default does not take place, the central bank's temporary injection of liquidity is taken out completely by the end of a day, and goods market prices are not affected. The price volatility induced by default of payments acts as an insurance mechanism for risk-averse creditors; it transforms credit risk borne by late-leaving creditors alone to inflation risk borne by all creditors. Therefore, the resulting allocation is preferable to having a constant price but a bigger fluctuation in consumption.

The model assumes that the timing of agents' arrival and departure as well as whether debtors default are exogenous, rather than endogenously chosen by agents. Therefore, the model is not suited to studying the potential moral hazard problem induced by free or low-cost provision of intraday settlement liquidity and the possible policies to offset it. If such decisions are explicitly modeled, it is quite possible that the result may be different. For example, such a model may support monitoring-cost-based pricing that differentiates and punishes agents who use intraday liquidity imprudently, as argued by Rochet and Tirole (1996). However, the intuition provided by the simplified model presented here should survive. That is, the optimal design of intraday-credit policy for a payment system has to take into consideration its distributional effect on all members of the system, as well as its effectiveness in reducing risk.

Conclusion

The simple model presented here takes into account the basic elements of the four criteria I proposed earlier. It models the underlying goods transaction so that whether the central bank provides intraday liquidity affects the consumption allocation. The model yields the result that economic efficiency requires consumption debt to be priced at a positive interest rate while payment debt should be priced at zero interest rate. The assumed payment flows generate a shortage of settlement liquidity so that the central bank's provision of liquidity improves welfare. Finally, the model assumes default risk such that not all payments can be settled.

Through the analysis of the model, I argue that settlement debt is very different from consumption/investment debt; while the latter facilitates the allocation of real resources across time, the former exists only for settling the underlying intertemporal transaction. Hence, consumption/investment debt should be appropriately priced to give proper incentives for the efficient allocation of real resources, while the cost of settlement debt should be minimized so that it does not distort the underlying goods/assets transaction. The temporary injection of free intraday liquidity by the central bank helps to achieve this goal. The provision of intraday settlement liquidity through a private intraday money market in central bank funds may be too costly, in particular when total funds in reserve and clearing accounts are in short supply. Furthermore, even with potential aggregate default risk, the provision of free intraday liquidity by the central bank may be the best way to ensure banks do not bear the brunt of the risk disproportionately.

The model does not meet the proposed standards completely because the introduction of the liquidity problem and settlement risk is rather mechanical. A richer setup where the twin problems are induced by agents' action would allow us to study other major payment system problems, such as delaying payments and the associated gridlock, banks' endogenous risk-taking decisions, and potential moral hazard. Further research efforts are needed to enhance our understanding of the role of intraday liquidity and its connection to the conduct of monetary policy (for example, inflation) and other central bank policies (such as zero-interest reserve requirements).

APPENDIX

A parametric model without settlement risk

I show a parametric version of the model without the provision of settlement liquidity by the central bank (the model presented in the section, *A model without settlement risk*). Since all creditors are identical ex ante and all debtors are identical ex ante, I look for a symmetric (all creditors act the same and all debtors act the same), competitive (agents on both goods markets and debt-resale markets are price-takers) equilibrium.

Let P_{Ct} and P_{Dt} denote the date-t prices for C-good and D-good, respectively, and let R_t be the date-t nominal interest rate on consumption debt.

Consider a generation-t debtor first, $t \geq 1$. At date t, suppose that a young debtor purchases C_{dt} units of C-good, pays with personal debt valued at h_t dollars, and sells $1 - D_{dt}$ units of his endowment D-good to old creditors in exchange for m_{t-1} dollars. At the end of the day, the debtor consumes his purchase and the remaining D_{dt} units of his endowment, which yield utility $\log(C_{dt}) + \log(D_{dt})$, and receives a lump-sum transfer m_t' units of money. Next morning, the debtor visits the central clearing market, where he pays back the debt and interest $R_t h_t$ dollars in full with the money obtained at date $t, m_{t-1} + m_t'$ regardless of whether he arrives early or late. The debtor chooses his consumption bundle (C_{dt}, D_{dt}) subject to his budget constraints to maximize his expected utility. That is,

$$1)\quad \max \quad \log(C_{dt}) + \log(D_{dt})$$

$$2)\quad \text{such that:} \quad P_{Ct}C_{dt} = h_t$$

$$3)\quad P_{Dt} = m_{t-1} + P_{Dt}D_{dt}$$

$$4)\quad R_t h_t = m_{t-1} + m_t' \equiv m_t.$$

Next, consider a creditor born at date t, $t \geq 1$. Suppose that the young creditor sells $1 - C_{ct}$ units of her endowment C-good to young debtors, accepting l_t dollars personal debt in exchange in the morning. She consumes the remaining C_{ct} units of her endowment. At date $t + 1$, she goes to the central clearing market to settle the debt she holds in the morning, and then purchases $D_{c(t+1)}$ units of D-good for consumption with the repayment money from young debtors in the afternoon. Her lifetime utility is given by $\log(C_{ct}) + \beta\log(D_{c(t+1)})$, where β is the discount factor. The creditor's date-(t+1) consumption depends on her experience at the clearing market. If the creditor is able to settle her debt holdings directly with her debtor and does not purchase second-hand debt (group X or group Z_1 creditor), let her date-(t+1) consumption be $D^{XZ_1}_{c(t+1)}$. If she leaves the market before her debtor arrives (group Y creditor), assume that she trades the debt holdings for money at discount $\rho_{t+1} \leq 1$ with other late-leaving creditors who have been repaid, and the proceeds yield her $D^{Y}_{c(t+1)}$ units of D-good. If she leaves the market late but settles with her debtor early, and purchases group Y creditors' debt at discount ρ_{t+1} (group Z_2 creditor), she obtains unredeemed debt valued at $1/\rho_{t+1}$ times of her original debt payment. Suppose that the repayment of these debt purchases affords her $D^{Z_2}_{c(t+1)}$ units of D-good consumption. Mathematically, the creditor chooses her contingent consumption bundle $(C_{ct}, D^{XZ_1}_{c(t+1)}, D^{Y}_{c(t+1)}, D^{Z_2}_{c(t+1)})$ subject to her budget constraints to maximize her expected utility,

$$5)\quad \max \quad \log(C_{ct}) + \beta[(\alpha+\lambda-2\alpha\lambda)\log(D^{XZ_1}_{c(t+1)}) + (1-\lambda)(1-\alpha)\log(D^{Y}_{c(t+1)}) + \alpha\lambda\log(D^{Z_2}_{c(t+1)})]$$

$$6)\quad \text{such that:} \quad P_{Ct} = P_{Ct}C_{ct} + l_t$$

$$7)\quad l_t R_t = P_{D(t+1)} D^{XZ_1}_{c(t+1)}$$

$$8)\quad \rho_{t+1} l_t R_t = P_{D(t+1)} D^{Y}_{c(t+1)}$$

$$9)\quad (1/\rho_{t+1}) l_t R_t = P_{D(t+1)} D^{Z_2}_{c(t+1)}.$$

An initial generation debtor does not play any role. An initial old creditor spends her endowed m_0 dollar of money to purchase her consumption, a first generation young debtor's endowment.

Since there are an equal number of debtors and creditors each generation, and each agent is endowed with one unit of his/her type-specific goods, for any date $t \geq 1$,

$$10)\quad h_t = l_t$$

$$11)\quad C_{dt} + C_{ct} = 1$$

$$12)\quad D_{dt} + (\alpha + \lambda - 2\alpha\lambda) D^{XZ_1}_{c(t+1)} + (1-\lambda)(1-\alpha) D^{Y}_{c(t+1)} + \alpha\lambda D^{Z_2}_{c(t+1).} = 1.$$

Also since all young debtors obtain the same amount of money by selling their endowment, the lump-sum transfer of money m'_t satisfies, for any $t \geq 1$,

$$13)\quad m'_t = i * m_{t-1}.$$

The D-good market (goods for money) clears, that is, for any $t \geq 1$,

$$14)\quad P_{Dt}(1 - D_{dt}) = m_{t-1}.$$

Define the interest rate to be the relative nominal price of the D-good across periods,[1] for all $t \geq 1$,

$$15)\quad R_t = \frac{P_{D(t+1)}}{P_{Dt}}.$$

The debt discounting rate ρ_{t+1} is determined by the demand and the supply of the unredeemed debt,

$$16)\quad \rho_{t+1} = \min\left\{1, \frac{\alpha\lambda}{(1-\lambda)(1-\alpha)}\right\}.$$

In such a model, a stationary equilibrium with active trading is efficient. An equilibrium is stationary if all creditors (or all debtors) across generations consume the same amount, that is, for all $t \geq 1$,

$$17)\quad C_{dt} = C_d, \quad C_{ct} = C_c, \quad D_{dt} = D_d,$$

$$D^{XZ_1}_{ct} = D^{XZ_1}_{c}, \; D^{Y}_{ct} = D^{Y}_{c}, \quad D^{Z_2}_{ct} = D^{Z_2}_{c}.$$

Depending on the parameter values, there are two possible stationary trading equilibria. In equilibrium 1, $\alpha\lambda \geq (1-\lambda)(1-\alpha)$. Equilibrium ($*$) is in fact the same as this case; there is no liquidity shortage, and debt is not discounted, that is, $\rho_{t+1} = 1$. Hence, from equations 7 through 9, $D^{XZ_1}_c = D^Y_c = D^{Z_2}_c \equiv D_c$. The solution of the model is given by

$$18)\quad C_d = \frac{\beta}{\beta+1}, \quad C_c = \frac{1}{\beta+1},$$

$$D_d = \frac{i+1}{i+2}, \quad D_c = \frac{1}{i+2},$$

$$P_{D1} = m_0(i+2), \quad P_{C1} = m_0 \frac{\beta+1}{\beta},$$

$$\frac{P_{D(t+1)}}{P_{Dt}} = \frac{P_{C(t+1)}}{P_{Ct}} = R_t = i+1.$$

In equilibrium 2, $\alpha\lambda < (1-\lambda)(1-\alpha)$. In this case, debt is discounted. From equation 16, $\rho_{t+1} = \alpha\lambda/(1-\lambda)(1-\alpha)$. The solution to the model differs from that of equilibrium 1 only in a creditor's old age consumption, which is contingent on being a group X, Y, or Z creditor.

$$19)\quad D^{XZ_1}_c = \frac{1}{i+2}, \quad D^{Y}_C = \frac{\alpha\lambda}{(1-\lambda)(1-\alpha)} \frac{1}{i+2},$$

$$D^{Z_2}_c = \frac{(1-\lambda)(1-\alpha)}{\alpha\lambda} \frac{1}{i+2},$$

which satisfies $D^{Z_2}_c > D^{XZ_1}_c > D^{Y}_c$. This equilibrium is inefficient.

If the central bank purchases group Y creditors' unredeemed debt at par with newly issued money, and then takes the same amount of money out when late-arriving debtors repay their debt, we get back to the efficient equilibrium 1.

[1] In fact, the model determines jointly $R_t PC_t$, the gross nominal payment next period for the purchase of C good this period. It does not determine R_t separately. This definition of the nominal interest rate is by convention.

NOTES

[1]One unsatisfactory feature of the Freeman model is that the liquidity shortage is generated from the exogenously imposed payment flows, rather than as an outcome of agents' endogenous payment decision.

[2]Without money growth, the reinterpretation of Freeman's (1996) result suggests that both the intraday and the overnight interest rate should be zero.

[3]This is another unsatisfactory feature of the model—default on payments is exogenously imposed rather than the agents' choice.

[4]The conventional settlement lag for foreign exchange transactions is even longer—two days rather than one.

[5]Japan's BOJ-NET offers the choice of both designated-time NS and RTGS arrangements for each transaction, although the share of transactions settled by RTGS is very small. According to BIS (1997), the share of transactions settled by RTGS was 1.2 percent of total in terms of number and 0.1 percent in terms of value in 1995.

[6]Britain's large-value payment system, CHAPS, previously operated under an NS arrangement and converted to an RTGS system in April 1996. The European Monetary Union chooses RTGS for its large-value funds transfer system, TARGET (Trans-European Automated Real-Time Gross Settlement Express Transfer), which is currently in the process of implementation. The only major developed country that has chosen to adopt an NS system as its main large-value payment system is Canada. The LVTS, debuted in 1999, is a privately owned hybrid system that offers assurance of settlement in real time (guaranteed by the Bank of Canada), although the actual settlement occurs at the end of a business day. See BIS (1997).

[7]A sale of securities combined with a forward (same-day) repurchase.

[8]When sufficient funds are not available in sending parties' SIC accounts, payment orders are held in a central queue and processed in a first-in-first-out (FIFO) basis once the covering funds are received. See the Internet at www.snb.ch/e/snb/interbank/inter.html.

[9]However, a bank can increase its capacity of overdraft for security-transfer-related activities by pledging collateral.

[10]The annual rate assumes 360 days a year and 24 hours a day. The interest charge is in addition to the fixed transaction fee (independent of the size of the transaction) that has always been in place.

[11]Intraday peak and average overdrafts for funds transfer alone show a similar pattern.

[12]I have argued earlier that an RTGS system with central bank provision of intraday credit can be viewed as a netting-with-central-bank system, although with some intraday-credit measures, settlement liquidity may be more costly and proportional to the amount of usage relative to a standard NS system. (NS systems such as CHIPS usually also impose net debit caps and require collateral, although at a lower level. The collateral is intended for potential settlement failure as part of the loss-sharing agreement among participants.) Nevertheless, some arguments for an NS system can be used to argue for a low-cost, nonmarket provision of intraday liquidity with an RTGS system.

[13]The environment introduced here is similar to Green (1997), a variant of the Freeman model that preserves the spatial separation of markets, but without assuming an island economy structure as in Freeman (1996).

[14]This setup is analogous to the large-value payment problem that a goods transaction and its payment occur at different times.

[15]Freeman (1996) assumes that each creditor holds a diversified portfolio of debt issued by different debtors. This assumption affects how a creditor's budget constraint is written, as shown in the appendix.

[16]I assume that the only means to settle debt is fiat money, which a private clearinghouse cannot issue.

[17]This is the solution to a social planner's problem that maximizes a weighted sum of utilities of the debtors and creditors in each generation.

[18]The exogenously imposed inflation is distortionary. Equilibrium (*) is the second best given the existence of inflation.

[19]If the central bank were to purchase debt below par, it would thereby withdraw money from the economy. (The difference between par and the purchase price, times the quantity of debt purchased, would be withdrawn from the economy at each date.) I am assuming that in such a case, injection of new money into the economy would increase by this withdrawn amount, so that the net growth rate of aggregate money stock entering the goods market remains at i.

[20]Although the interest is paid by group Y creditors, rather than the presumed intraday-credit borrowers, the late-arriving debtor. This mismatch of the interpretation of the model and reality arises because, for simplicity, I assume the payment flows are exogenous, whereas in practice, funds transfers are initiated by funds-sending banks (debtors) in most large-value payment systems.

[21]The only industrialized countries with some form of intraday money markets are Japan and Switzerland (prior to 2000), and they exist solely to serve the liquidity needs of settling payments since these countries' central banks do not provide any form of intraday settlement lending.

[22]The debtors' preference has to be changed; in addition to consuming both goods when young, they now also consume the D-good when old. This assumption ensures that the defaulters' money is not withdrawn from the goods markets.

[23]For a detailed mathematical derivation of the result, see Freeman (1999).

REFERENCES

Angelini, Paolo, 1998, "An analysis of competitive externalities in gross settlement systems," *Journal of Banking and Finance*, Vol. 22, pp. 1–18.

Bank for International Settlements, 1997, "Real-time gross settlement systems," report prepared by the Committee on Payment and Settlement Systems of the central banks of the Group of Ten countries, available on the Internet at www.bis.org/publ/index.htm.

__________, 1993, *Payment Systems in the Group of Ten Countries*, Basel.

Evanoff, Douglas D., 1988, "Daylight overdrafts: Rationale and risks," *Economic Perspectives*, Federal Reserve Bank of Chicago, May/June, pp.18–29.

Freeman, Scott, 1999, "Rediscounting under aggregate risk," *Journal of Monetary Economics*, Vol. 43, pp. 197–216.

__________, 1996, "The payment system, liquidity, and rediscounting," *American Economic Review*, Vol. 86, pp. 1126–1138.

Green, Edward J., 1997, "Money and debt in the structure of payments," *Bank of Japan Monetary and Economic Studies*, Vol. 15, pp. 63–87, reprinted in 1999, *Quarterly Review*, Federal Reserve Bank of Minneapolis, Spring, pp. 13–29.

Hancock, Diana, and James A. Wilcox, 1996, "Intraday management of bank reserves: The effects of caps and fees on daylight overdrafts," *Journal of Money, Credit, and Banking*, Vol. 28, No. 4, pp. 870–908.

Humphrey, David B., 1989, "Market responses to pricing Fedwire daylight overdrafts," *Economic Review*, Federal Reserve Bank of Richmond, May/June, pp. 23–34.

Kahn, Charles M., and William Roberds, 1999a, "Real-time gross settlement and the costs of immediacy," Federal Reserve Bank of Atlanta, working paper, No. 98-21a.

__________, 1999b, "The design of wholesale payments networks: the importance of incentive," *Economic Review*, Federal Reserve Bank of Atlanta, Third Quarter, pp. 30–39.

Kobayakawa, Shuji, 1997, "The comparative analysis of settlement systems," Centre for Economic Policy Research, discussion paper, No. 1667.

Lacker, Jeffrey M., 1997, "Clearing, settlement, and monetary policy," *Journal of Monetary Economics*, Vol. 40, pp. 347–381.

McAndrews, James, and Samira Rajan, 2000, "The timing and funding of Fedwire funds transfers," *Economic Policy Review*, Federal Reserve Bank of New York, Vol. 6, No. 2, pp. 17–32.

Mengle, David L., David B. Humphrey, and Bruce J. Summers, 1987, "Intraday credit: Risk, value, and pricing," *Economic Review*, Federal Reserve Bank of Richmond, January/February, pp. 3–14.

Richards, Heidi Willmann, 1995, "Daylight overdraft fees and the Federal Reserve's payment system risk policy," *Federal Reserve Bulletin*, December, pp. 1065–1077.

Rochet, Jean-Charles, and Jean Tirole, 1996, "Controlling risk in payment system," *Journal of Money, Credit, and Banking*, Vol. 28, No. 4, pp. 832–862.

3 A direct participant in a large-value payment system (LVPS) is a participant that can settle transactions without using an intermediary. If not a direct participant, a participant will need to use the services of a direct participant (correspondent bank) to perform payments and settlements on its behalf.Â Local regulator needs to understand how correspondent banks mitigate intraday liquidity risk. Local regulator must identify behavioral changes in how banks manage intraday liquidity risk. Table 2 illustrates possible causes of intraday liquidity risk and corresponding mitigating actions.Â counterparties to defer payments and/or withdraw intraday credit lines. In such an. event, a direct participant must fund more of its payments from its own sources of Financial stress. This article explains how large-value payment systems work, using either gross or net settlement. The author discusses risk control in a real-time gross settlement system and analyzes the pricing of credit to provide intraday liquidity. She argues for distinguishing between consumption/investment debt and payment debt. A theoretical model suggests that, under the assumption that there are no opportunities for intraday optimization of consumption and production, the risk-free rate on intraday payment credit should be zero. This is because the cost of intraday liquidity is a transaction cost of