International Finance Disucssion Papers

Number 216

February 1983

PAYMENTS SYSTEMS: THEORY AND POLICY

by

Allen Frankel and Jeffrey Marquardt

January 1983

PAYMENTS SYSTEMS: THEORY AND POLICY

BY

Allen Frankel and Jeffrey Marquardt*

Introduction

The goal of this paper is to provide a simplified financial framework for examining a number of theoretical and policy issues raised by the existence and operation of interbank payments arrangements. To this end, payments arrangements are assumed to be embedded in an exogenous, real sector. The organization of payments activities, including their risks, is examined under increasingly general assumptions about the process of transferring payment from one agent in the nonbank sector to another. The most general framework presented in this paper consists of a multibank economy with trade and bank credit. The framework developed is intended for the analysis of payments arrangements in which transactions costs, information costs, and capital market imperfections significantly affect behavior.

Part I of this paper sets out the general framework. The economic rationale for the periodic settlement of payments is examined. Part II

*International Finance Division, Federal Reserve Board. This paper represents the views of the authors and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System or members of its staff.

examines the allocation of risk in an economy when supervisory authorities do not intervene to regulate payments networks. Part III examines one rationale for the regulation of payments networks -- moral hazard in the presence of a lender of last resort -- and some of the problems of attempting to regulate such networks.

The strategy of this paper is to combine a model of payments transactions processing with an Edgeworth-type portfolio model for bank providers of payments services. This combination is designed to capture the roles of real and financial costs in the determination of the settlement interval, the utilization of bank-provided payments services and the organization of the interbank payments mechanism.

Government intervention in our analysis is assumed to involve guarantees to bank providers of payments services. This view of government intervention in the payments mechanism is considerably different from the traditional U.S. model that has involved a direct Federal Reserve subsidy of the real costs of processing payments transactions through its check clearing services to member banks. On the other hand, we view government intervention as a guarantee against loss which consequently suppresses the value of credit analysis to providers of payment services. Moreover, this analysis is appropriate to private payments systems as well as the Federal Reserve System's own interbank funds transfer network, the Fedwire.[1/]

1/For a description of Fedwire see Association of Reserve at Bankers [2] pp. 15-16, 25-27. Coats and Frankel [3] contains a discussion of the operating rules of the Fedwire.

Several net settlement payments networks (such as the CHIPS network) have been, or are in the process of being, created. Interest in such mechanisms has been encouraged by the technological changes (computerization) that have reduced the costs to payments processors of more-timely updating of positions which facilitates the making of intraday credit judgements.1/ Interest has also been encouraged by the development of standardized formats for sending computerized funds transfer messages through organizations such as SWIFT2/

1/CHIPS is the acronym for the Clearing House Interbank Payments System. CHIPS, which is operated by the New York Clearinghouse Association, has 100 bank participants and processes more than $200 billion daily in gross payments.

2/SWIFT is the acronym for the Society for Worldwide Interbank Financial Telecommunication, a co-operative company created under Belgian law. It is wholly owned by member banks. SWIFT provides an international transaction processing network to which only banks (or bank-owned organizations) may connect.

I. Payments System Framework

This part presents several models of payments relationships. The strategy in this part is first to analyze the structure of payments in an economy with firms but no banks. The organization of payments activities is then examined as one bank and then multiple banks are introduced. The forms of credit in the economy are also allowed to become successively more complex. In the simplest economy, only interfirm trade credit is allowed. In the most complex economy, firms that trade with one another are allowed to extend each other credit and banks are allowed to extend credit to firms.

A. Economic Agents, Balance Sheets and Objectives

The agents in all regimes can be introduced by describing the agents in the most complex economy. Less complex regimes are constructed by placing restrictions on these agents and their balance sheets.

There are four agents in the most complex economy: firms, banks, a banker's bank, and a lender of last resort. Payments networks are a fifth conceptual entity in the economy. These networks, however, are composed of banks and are not independent decision-making entities.

The focus of our analysis is on the physical and financial relationships among the four types of economic agents arising out of the provision of payments services. Conceptually, there are two important dimensions to payments services. First, payment services are specialized forms of communication among banks. Second, payment services involve provision of a means for discharging debts among firms resulting from banks' abilities to reconcile financial positions vis-a-vis firms and other banks.

Firms. Firms are private agents that sell goods to one another. Trade between firms may be described as a set of contracts that specify quantities and types of goods, buyer and seller, prices, delivery dates and payment terms. It is assumed that only the payment terms (due date and means of payment) are endogenous variables.

The balance sheet of a representative firm is depicted in Figure 1. The firm may hold four types of assets. Inventories of goods are held for

Figure 1

Representative Firm's Balance Sheet

Assets	Liabilities
Gold	Accounts Payable
Bank Deposits	Bank Loans
Accounts Receivable	
Inventories	Net Worth

sale to other firms. Stocks of inventories and sales of goods reflect the underlying pattern and timing of trade in the real economy. Gold is a completely malleable store of wealth whose value appreciates at the economy's exogenous rate of return.1/ The exchange of gold is the ultimate means of monetary settlement. However, gold is very costly to transport and

1/This assumed behavior of the gold price, rising at the rate of interest, is consistent with a model in which, in part, the existing stock of gold (an exhaustible resource) is held by those who value its monetary services. Gold earns no rental income when its stock exceeds the quantity that agents enjoying its monetary services would want at any positive rental. A discussion of this point is found in Salant and Henderson [6].

protect.[1/] Bank deposits may also be held by firms, but earn an implicit return in the form of payment services. Firms may acquire deposits either by surrendering gold to a bank or by accepting the transfer of a deposit in payment for goods. Accounts receivable are acquired by selling goods to other firms and not receiving immediate payment in gold or bank deposits.

Firms have two liabilities: bank loans and accounts payable. Firms borrow from banks (receive bank loans) to finance payments. Firms obtain accounts receivable when they extend trade credit to buyers of their goods. Firms are assumed to minimize the expected costs of executing the payments required by underlying goods contracts.

Banks. The balance sheet of a representative bank is depicted in Figure 2. Banks have four types of assets. A bank may hold gold that it

Figure 2

Representative Bank's Balance Sheet

Assets	Liabilities
Gold	Deposits of Firms
Deposits at Banker's Bank	Loans from Banks
Loans to Banks	
Loans to Firms	Net Worth

acquires in exchange for the deposit of gold by firms, or possibly the settlement of deposit transfers from other banks. The bank may itself deposit gold at an institution called the banker's bank in exchange for deposit

1/From a modeling standpoint gold, rather than fiat currency, is useful as an ultimate monetary asset. Although exchanges of fiat currency involve opportunity and transactions costs, the physical qualities of gold make clear that real transactions costs exist in the making of payments without resorting to a full model of private sector behavior. There is also a long history of the use of gold as an ultimate monetary asset. A brief description of the role of gold shipments in the U.S. payments mechanism in the nineteenth century is found in Garbade and Silber [4].

liabilities of the banker's bank. The bank may also lend to firms or other banks acquiring claims on those entities.

Banks have two types of debts. They are debtors to firms that have deposited gold. Banks also become debtors if they receive interbank loans. Net worth is equal to the equity position of bank owners. Bank depositors are assumed to be the senior creditors of a bank.1/

It is assumed that each firm is initially assigned to one bank. A switch of banks by a firm is assumed to require a lump sum charge to be paid by the firm. The bank does not necessarily know when its firms will make or receive payments. Maximization of the value of bank equity involves the manipulation of both assets and liabilities so as to minimize the expected costs of making and receiving payment transfers.

Banker's Bank. A single banker's bank is assumed to exist. It performs the equivalent service for banks that they perform for firms. The balance sheet of the banker's bank is shown in Figure 3. The banker's bank accepts gold from banks or the lender of last resort in exchange for deposit

Figure 3

Balance Sheet of the Banker's Bank

Assets	Liabilities
Gold	Deposits from Banks
Loans to banks	Deposits from LLR
	Net Worth

liabilities that earn no explicit return. The banker's bank may also grant loans to banks. The loans are disbursed by creating and transferring deposit liabilities of the banker's bank.

1/There is no explicit deposit insurance in the regimes considered in this paper.

Lender of Last Resort. In several contexts it will be necessary to introduce the concept of a lender of last resort (LLR). In practice, a central bank often combines the functions of the banker's bank and the LLR. In this paper, the LLR is treated separately for analytical convenience.

The balance sheet of the LLR is shown in Figure 4. It is assumed that the LLR is capitalized by lump sum taxes levied on firms. Loans made by

Figure 4

Balance Sheet of the Lender of Last Resort

Assets	Liabilities
Deposits at Banker's Bank	
Loans to Banks	
	Net Worth

the LLR take the form of transfers of deposit liabilities of the banker's bank. These assumptions allow for the separation of the portfolios of the banker's bank and the LLR.

The reason for introducing the LLR is that its presence and expectations about the availability of its loans can fundamentally affect payments activities.

Payments Networks. Payments networks are groups of two or more banks. The minimal network containing two correspondent banks is simply an established pattern for sending and receiving payments. By contrast, networks involving more than two banks are cooperatives or clubs in which members share decision-making authority. These networks establish and maintain communication links between members. They may also establish and maintain monitoring, switching, control, and data processing facilities.

B. Organization of Payments Activities -- No Credit to Firms and No Intersettlement Period Credit to Banks

This section examines the existence of equilibrium settlement points, bank portfolios, and payments system organization under two simplifying assumptions: no credit can be granted to firms and no interbank credit can be granted to finance settlement. Interbank credit is allowed only within a settlement interval. The restrictive assumptions are dropped in Section D. To analyze the effect of the structure of banking on payments systems, three types of regimes are examined: an economy without banks, an economy with one bank, and finally an economy with many banks.

1. No Banks

In a model without banks, gold is the only monetary asset in the economy. Use of the monetary asset for exchange dominates pure barter. The cost of transporting gold per occurrence is assumed to reflect both fixed and per unit costs of transport.

Because of the real costs of making payments, trading firms have an incentive to sell goods on credit and settle net amounts due at a later date. If firms delay payment, anticipated goods transactions will occur that partially or fully offset initial amounts due between pairs of firms. This netting of transactions over time between firms can significantly reduce net amounts to be settled, thus reducing the real costs of making payments. The extent of the real savings, of course, depends on the pattern and timing of

trade between firms and the underlying structure of settlement costs.

In general, there is a limit to the length of time firms will forego settlement. If expected bankruptcy losses from delaying settlement grow over time for at least one firm in any pair, then (discounted) expected losses will ultimately outweigh any share of (discounted) real savings. At that point, the other firm will demand that net transactions be settled.

2. One Bank

The introduction of "bank money" is a source of real savings to the economy.[1/] Firms will deposit gold in the bank until the opportunity cost -- return on gold foregone -- of holding the bank's liabilities equals the value of the flow of services from its deposits. That value comes from the reduced protection and transportation costs for bank liabilities as compared with gold.[2/] Some private sector transactions will still be settled with gold transfers. However, the transfer of bank liabilities will generally provide a lower-cost means of settlement.

The introduction of "bank money" can also alter the settlement practices of the nonbank sector. Since settlement is less costly with bank liabilities, the savings in settlement costs to firms from the netting of transactions over a given interval will be increased. Firms will react to

---

1/See, for example, the discussion in Johnson [5].

2/A similar argument can be made for bank notes that are essentially liabilities of a bank to the bearer of the notes.

lower costs of transfers by shortening settlement intervals to reduce expected losses from insolvencies.

The use of bank liabilities for payment makes clear that payment is a process rather than an event. When gold is exchanged for goods, it is natural to think of payment as occurring when possession of the gold changes from the buyer to the seller of goods. When bank liabilities are used for payment, the bank must first be alerted that ownership of a quantity of its liabilities is to be changed from one customer to another. An interval of time passes before the bank may process any payments. Finally, at the end of the interval, a recipient of a payment ends up with an unconditional right to withdraw gold from the bank. Only when the recipient has this unconditional right is payment said to be final.

The timing of finality in a one bank economy reflects the cost structure of payments processing for the bank. If costs are minimized by processing and posting changes in accounts periodically, "batch" processing, then payments will only be final at the end of a period. If costs are minimized by the continuous, "on-line", processing of payments, then finality may be instantaneous.

The single bank economy provides a benchmark for the analysis of risk in payments systems. Recall that in this section the bank does not make customer loans and, therefore, there is no risk that the bank will become insolvent because of defaults by customers. Moreover, since the bank does not invest in anything but gold, it cannot become illiquid; depositors can always retrieve their gold. Hence there is no payments system risk in this single bank economy.

3. Multiple Banks

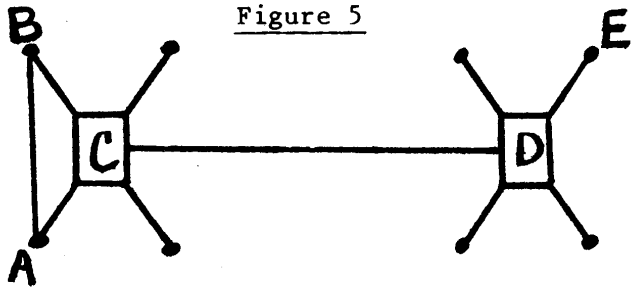
The organization of payments arrangements can be much more complex in an economy with multiple banks. If a buyer and seller are assigned to the same bank, then the payment for goods takes place as in a one bank economy. If buyers and sellers are assigned to different banks, then bank payments may take place through a sequence of bilateral interbank transfers, through interbank transfers in a multilateral payments network, or both.

The existence of a banker's bank further enhances real resource savings from exchanges of bank liabilities in the multibank economy. Banks may deposit gold in exchange for banker's bank liabilities that can be transferred between banks. Consequently, settlement of interbank payments does not require that gold be transported between banks and in general the transfer of banker's bank liabilities will dominate gold shipments as a means of settlement.

Bilateral Payments Networks. In a single bank economy, the processing of bank payments is internalized by the one bank. In a multibank economy, external organization is required for bank payments even if a banker's bank exists.

Figure 5 depicts the most basic kind of external payments network, the bilateral system. For example, suppose a customer of bank A wants to make

Figure 5



payment to a customer of bank B. The payment is made through a communication between the banks. The communication, it should be noted, is of a special type. Specific information must be sent and received in a verifiable and recognizable form according to established agreements and procedures. Because there may be significant costs to establishing communication links between banks, many banks in the economy may be linked only indirectly through other banks. In Figure 5 for example, banks A and E are linked only through the "switching" banks C and D. Moreover, if the pattern of settlement tracks the communication links, then a payment from a customer of bank A to a customer of bank E would produce a sequence of bilateral settlement transactions by banks A and C, C and D, and D and E.

A number of interbank payments issues are raised by bilateral systems. The first issue is the existence of interbank settlement intervals. As in the case of firms that trade with one another over time, banks that exchange payments over time can economize on settlement costs by settling net payments periodically. If settlement must be in gold, the banks save on transportation and protection costs. If settlement is in banker's bank liabilities, the banks save on transaction costs generated by the need to acquire or dispose of an inventory of banker's bank liabilities. In addition, because the banker's bank pays no interest on its liabilities, profit-maximizing behavior by banks, other things being equal, will entail lengthening settlement intervals in order to minimize the opportunity cost of holding balances through the netting of payments.

The time limit on interbank credit is more difficult to determine than in the case of nonbank trade credit. If banks can become insolvent, then settlement intervals would be based on the minimization of settlement costs and expected losses from insolvencies as in the extension of trade credit. Equilibrium settlement intervals may be determinate if settlement costs are a decreasing function of the length of the settlement interval and if the expected losses are an increasing function of the interval.

If payments to, and from, the nonbank sector are not final until interbank settlement is complete, the nonbank sector's options influence the duration of interbank settlement intervals. In a two bank economy in which each bank has one firm, for example, interbank settlement intervals cannot exceed the interfirm settlement intervals that would have existed in the absence of the banks. If interbank intervals were longer, the firms would withdraw their gold from the banks and resort to settlement through gold shipments. In general, determination of the nonbank sector constraint on interbank settlement can be complex. The constraint depends on underlying interfirm trade patterns, expected losses to creditor firms from insolvencies, and the customer mix at various banks. The constraint, however, does exist.1/

Settlement risk can be a major factor in determining interbank settlement by bringing expected losses from insolvency back into bank

1/If payments are final before settlement, the threat by firms to withdraw gold from banks may ultimately constrain interbank settlement intervals.

calculations. The risk of insolvency exists in a multiple bank model with interbank but no customer credit. Insolvency might be induced by illiquidity. As banker's bank liabilities earn no return, a bank enhances its profitability by choosing to hold larger proportions of its asset portfolio in gold. However, unanticipated net payments requiring costly immediate conversions of gold to banker's bank liabilities could leave a bank insolvent. If the bank has insufficient net worth to cover time-dependent costs of converting gold, then the bank would be unable to complete settlement.1/

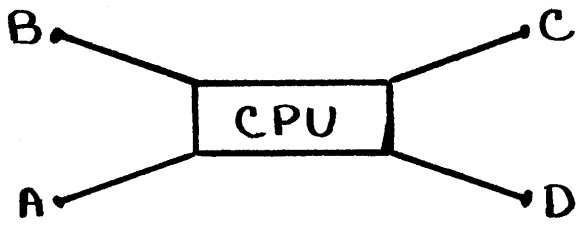
The problem of settlement risk raises the issue of the transmission of insolvency that is discussed more fully in the next two sections. In an economy with multiple banks, unanticipated net payment outflows can cause banks to become illiquid. This result is impossible in a one bank economy unless gold is withdrawn from the bank.2/ Moreover, when portfolio adjustment costs --costs of converting gold into deposits at the bankers' bank -- exceed the equity in a bank, the bank becomes insolvent.

Multilateral Communications Networks. A network with multilateral communications links but bilateral settlement is the conceptual bridge between the bilateral system just described and a fully multilateral system. Figure 6 is a schematic of the communications arrangements. When bank A wants to send or receive payment communications from bank D, it does so through a

1/The major source of time-dependent costs is the inability to sell unlimited amounts of gold without affecting its market price. Such capital market imperfections critically affect the cost of liquidating portfolios of financial assets.

2/It is generally assumed in this paper that withdrawals of gold by firms reflect only changes in needs for payments services, not "runs" on banks.

Figure 6



central processing unit [CPU]. Communication is multilateral in the sense that there are many banks on the network, each linked to all through a central facility.

It is important to recognize that settlement may remain bilateral even though a centralized communication facility is established. The analysis of bilateral settlement systems above would apply fully to such systems. The effect of centralized communication, however, may be to reduce the costs of establishing particular bilateral payments links, thus reducing the length of payment chains. Banks would still determine settlement terms, including the extent and duration of interbank credit, bilaterally. If, for example, one bank refuses to extend credit to a bank wanting to make a payment through it, then payment could still be made by combining the payment message with an immediate transfer of a settlement asset.1/

Multilateral Net Settlement Systems. Multilateral net settlement systems work differently than other payments networks. Net bilateral positions of individual members are consolidated into aggregate net debit or

1/Another option is to make the payment through a third bank that will grant credit.

credit positions vis-a-vis all other members of the network. Aggregate net debtors then discharge their consolidated net debt by transferring the settlement asset, banker's bank liabilities, to consolidated net creditors at prearranged settlement times.

The economic rationale for net settlement systems is twofold. First, savings of real resource costs are possible with the faster, more certain, and more efficient processing of payments through centralized communication and processing facilities. Depending on the extent of bilateral settlement links between banks, some of these economies also result from networks that settle bilaterally but have centralized communication. Second, however, multilateral settlement systems permit real savings in the settlement process. Members of the network not only net all bilateral positions over time to the extent permitted by the length of the settlement interval, but also net the cross-section of bilateral positions at the time of settlement. As a result, a smaller volume of settlement services are purchased. Fewer banker's bank liabilities need be held by participants in multilateral net settlement arrangements to complete the same volume of gross payments.

Net settlement systems are essentially cooperatives or clubs in which members benefit by realizing savings in the costs of transmitting and settling their customers' payments.[1/] The start-up (equipment and organization) costs for a network are likely to make it necessary that some critical group of banks agree in advance to participate in the formation of a network. Once established, network configurations are likely to be built

1/The terms "cooperative" and "club" are meant to be suggestive. The actual legal organization of payments networks may take a variety of forms.

using the Pareto principle. That is, banks will be added to a network as long as no bank's total costs increase and at least some banks' costs decrease. However, if admissions to networks are irrevocable, there are likely to be some participants whose costs are higher because certain banks were admitted and others were not. An inability to plan and coordinate network development may result in local rather than global cost minimization.1/

Thus particular banks would not be added to a network if they would increase processing and communications costs for one or more banks on the network. Banks that are perceived to increase the risks in interbank lending to finance payments on the network may also be denied membership in some arrangements. Whether or not a bank would be denied access depends on the aggregate net credit the bank would receive, the perceived probability that the bank would fail to settle, expected losses to others from the failure, and the size of cost savings from having the bank join the network. While this calculation can be very complex, two concrete points should be noted. First, banks about which incumbent network participants have little information may be excluded from a network because of perceived, rather than real, risks -- a form of discrimination. Second, the existence of an outside agent, such as the LLR, that guarantees settlement may eliminate the consideration of risk from network design. Bank participation in net settlement networks would tend to be more extensive than without a guarantor.

As with membership, settlement intervals are also determined collectively in a net settlement system. Short settlement intervals are less

1/Banks' recognition of the inefficiencies produced by this admissions procedure could be reflected in resignations from networks and the establishment of new arrangements.

likely to result in anticipated netting of transactions over time and, in general, require larger real expenditures for coordination among network participants. Thus, the costs of most banks in the network would be likely to decrease if the interval were lengthened. At some point, some participants will perceive that expected losses from the insolvency of a member outweigh the savings from a longer settlement interval as in the case of bilateral settlement. However, a credible assurance of settlement and the payment of interest (at the rate gold earns) on banker's bank liabilities, may cause interbank settlement intervals to become indeterminate. For this case, to have determinate intervals, an asset earning a higher expected return than gold or penalty rates on lender of last resort loans would be needed to induce determinate intervals.

In a multilateral settlement system with settlement risk, expected losses from the insolvency of a member are not necessarily based on bilateral net credit relationships as in bilateral payments networks. This point is discussed more fully in Part II. However, it should be noted that a payments network might adopt one of a number of rules for loss sharing in the event that a member becomes insolvent and unable to settle. Three examples of such rules are: (1) surviving net creditors bear losses in proportion to net credit extended, (2) surviving net debtors bear losses in proportion to net credit received, and (3) all surviving members bear losses equally or according to some agreed formula. Since these types of rules affect the distribution of expected losses among network members, they also influence the composition of networks.

C. Interbank Adjustment and Risk

The purpose of this section is to focus on interbank adjustment and risk due to payments networks in a simple multibank economy. To that end, the assumption of no bank credit to firms, or to banks over settlement points, continues to hold. It is assumed that net creditors, both in bilateral and multilateral settlement systems, bear losses in proportion to net credit extended when their debtors fail to settle; payments are final when received, generally before settlement. Finally, it is also assumed that there are increasing marginal costs of short-notice portfolio adjustment for firms and banks.

A point made earlier deserves elaboration here. Even when banks do not make loans to the nonbank sector, bank failure remains possible and can be transmitted to other banks through net credit extensions. Banks make portfolio choices about the relative proportions of gold and banker's bank liabilities that will be held just prior to the end of a settlement interval. Banker's bank liabilities earn no interest but have value because the holder avoids short-notice settlement, portfolio adjustment costs, or both. If insufficient banker's bank liabilities are held, a bank with unexpected net payment outflows will have to settle in gold or convert gold to banker's bank liabilities. If adjustment costs exceed the equity of a bank, it will become insolvent and fail to settle. Interbank net credit is the means by which one bank's insolvency can be transmitted to other banks.[1/] Presumably, banks make portfolio decisions, including decisions to extend net

1/For simplicity, insolvent banks are treated as if they disappear upon insolvency. The possibility of creditor recoveries after insolvency proceedings is ignored.

credit, taking insolvency probabilities into account. To the extent that portfolio choices result in relatively high proportions of gold and net interbank credit, and a corresponding low proportion of banker's bank liabilities, in the portfolios of net creditor banks, the insolvency of a net debtor bank can generate portfolio adjustment problems for its net creditors at settlement.

It is important to recognize that in an economy with no bank credit to firms, there is a partial separation of the bank and nonbank sectors. Since there are no loans to firms, credit shocks originating in the nonbank sector cannot be transmitted to banks via default on customer loans. Moreover, in the process of trying to settle payments, disturbances originating in the banking system are not transmitted directly to firms. If payments are final prior to settlement, a bank not receiving settlement from its net debtor banks has no recourse to its customers. If payments are not final before settlement, no customer adjustments at settlement will be required because of the assumed absence of bank credit to firms. However, firms may suffer the loss of part or all of their deposits as a result of a bank's insolvency.

The pricing of interbank risk also raises interesting problems. In the absence of credit from the banker's bank or the lender of last resort, the pricing of interbank credit between settlement points reflects two factors: opportunity cost and expected losses from insolvency. If the banker's bank paid interest on its liabilities, net extensions of credit must be priced to reflect foregone earnings less the marginal savings from the netting of transactions that is made possible by the credit. On the other hand, if no

interest is paid on banker's bank liabilities, savings from the netting of transactions and the deferral of portfolio readjustments into gold would continue to affect the price of credit. The primary determinant of credit prices in the latter case, however, would be expected losses from bankruptcy.

For given creditor portfolios, the expected losses to creditor banks from the insolvency of a net debtor will in general depend on the asset mix of the debtor bank, its net worth, and in particular its consolidated net debtor position vis-a-vis other banks. Since expected losses from an insolvency include the adjustment costs of net creditors, portfolios of the net creditors are also important to a full determination of expected losses. The portfolios of net debtors are not only important in the determination of potential losses, but also important in the determination of the probabilities of insolvency. Moreover, these probabilities can change rapidly during a settlement period as consolidated net debt positions change. As discussed above, the insolvency of one bank can induce other insolvencies creating indirect exposures in a banking system. The degree of indirect exposure in turn depends on the portfolio mix and capital of every bank in the system. Indirect exposure also depends critically on the bilateral net debt positions among banks in the system that are the channels for transmitting insolvency. These bilateral positions, of course, are also subject to rapid change during a settlement interval. Abstracting from creditor adjustment costs, the bilateral net credit positions of individual banks are the potential losses to which direct and indirect probabilities of insolvency are attached in order to compute expected losses.

The effect of the information requirements needed to compute expected losses is likely to be the adoption of credit indicators by banks in payments networks. If bilateral net positions were used as indicators of credit risk, a high bilateral credit price at a given time would create an incentive for a bank attempting to make payments to spread those payments across time and across network participants to the extent possible. Another type of credit indicator is the aggregate net debt position of individual banks vis-a-vis other banks in an economy's payments networks.1/ The price of credit from every bank might be tied to this indicator and vary over time. If prices were based on the same indicator, then a bank wanting to make payments would not have as great an incentive to spread payments across banks. The incentive to spread payments over time, however, would remain.

The lender of last resort and the banker's bank may also alter the pricing of intra-settlement period credit. If liquidity loans are freely-available from the LLR at no charge, expected losses to payments network participants from the insolvency of a member are nil and the market price for interbank (intrasettlement period) credit would be zero. Alternatively, if the banker's bank were to provide unlimited intrasettlement period credit at no charge, the price of interbank (intrasettlement period) credit would also be zero.

1/The coordination problems for this type of indicator might be solved if an economy has one net settlement system. With multiple systems and additional bilateral networks, the coordination problems could become very serious.

D. Effects of Bank Credit to Firms and Intersettlement Period Credit to Banks

This section discusses the effects of bank credit to the nonbank sector as well as interbank intersettlement period credit in multibank economies. The loans and some of their immediate effects are described first. The transmission of credit and liquidity shocks is then discussed. Part II analyzes the allocation of risk in a multibank economy. In anticipation of that discusion, a bank's decision to revoke provisional payments is also examined.

Loans are created when a bank customer draws on a line of credit negotiated and established in advance. The draw can be through the overdraft of a deposit account or some other form of automatic creation of a bank asset. Loans are used to finance payments imbalances of firms and banks. However, since payments imbalances are to some degree within the control of both banks and firms, the loans actually finance the level, terms, and timing of economic activity as well as ex post payments imbalances.1/

Interbank loans. Interbank loans that extend over one or more settlement points may be granted. These loans do not exist in a single bank economy. In a multibank economy, such loans may also provide an important source of real savings to an economy. Interbank loans may be used to finance unanticipated net payment outflows that would otherwise require costly bank portfolio adjustments such as gold shipments. In extreme cases, interbank loans may be used to avoid adjustment costs that would result in insolvencies for one or more banks in the economy.

1/Recall, however, that the level of real trade as well as prices is assumed to be exogenous.

Loans to Firms. Banks may also grant loans to firms. Many of the characteristics of loans to firms are also characteristic of interbank loans. A credit relationship is established between a bank and a firm when the firm draws on its credit line by ordering a payment to another agent in the economy that the bank has an obligation to settle. At the point the bank's obligation arises -- before or at settlement -- equity holders of the bank and possibly depositors incur two types of risk. First, the loan may not be repaid. Second, the bank's portfolio will become less liquid creating a greater likelihood that burdensome portfolio adjustments may be needed at settlement points. In one sense, firms demand loans in order to avoid the risk of costly portfolio adjustments, i.e., liquidity risk. To bear this risk, borrowing firms must compensate the bank. Firms are able to pay for loans out of portfolio adjustment costs saved. In sum, loans to firms can be a source of real savings to the economy as a whole.1/

The presence of loans to firms may affect settlement intervals in the bank and nonbank sectors. Since bank loans further reduce the costs of interfirm settlement, firms shorten trade credit intervals. For a given time profile of payments, banks shorten their settlement intervals to increase the probability and size of debtor positions of firms at settlement times. Banks are able to profit from any increased loan volume required by firms to finance payments. Shortening settlement intervals, however, increases processing costs, hence offsetting additions to bank revenue from increased loans. Moreover, firms are able to avoid some of the effect of shortened interbank intervals by monitoring their payments more intensively.

1/In a model with endogenous trade, bank loans may also permit increased trade for a given stock of gold.

One particular type of lending to firms is very important in economies in which payments are not final until settlement. If a firm receives a payment that is not final -- a "provisional" payment -- its bank may permit that firm to treat the payment as if it were final and to make transfers based on provisional payments received. The bank is in essence making a loan that is callable at settlement. If a sending bank fails to settle the provisional payment, the receiving bank may call upon the firm to restore "funds". A bank making these loans would treat provisional payments as security for the loan. Naturally, a bank with large concentrations of payments from one bank may find that a large amount of security can be wiped out if the sending bank fails to settle.

Transmission of Shocks. The existence of loans to firms changes the initial assignment of risk in a multibank economy. In a single bank economy, the risk of default on customer loans -- credit risk -- is born by the equity holders of the bank and to some degree by depositors. In a multibank economy, as already stated, liquidity risk exists in payments networks and liquidity shocks may be transmitted among banks through the granting of net intra-settlement period credit.

When banks grant loans against provisional payments, liquidity and credit shocks originating in or transmitted to the banking system may be directly passed to the nonbank sector. The failure of a sending bank to settle its payments may cause receiving banks to call loans that have been secured by provisional payments. The call of these loans can create severe portfolio adjustment problems -- liquidity shocks -- for the firms.

Revocation of Provisional Payments. The next part of this paper examines the allocation of risk in a multibank economy. Part of the analysis turns on the revocability of provisional payments at settlement. Hence it is useful at this point to focus on a bank's decision to revoke a provisional payment.

Given the failure of a sending bank to settle its payments, the receiving bank has an option to call the loans secured by provisional payments. Portfolio adjustments are not costless for either firms or banks. A firm would be willing to pay up to the amount of its adjustment costs in order to avoid the immediate repayment of call loans. Because of access to interbank intersettlement period credit, however, banks may have much lower costs of portfolio adjustment than firms[1/] Given differential adjustment costs, there is a profitable opportunity for banks to undertake adjustment and charge some firms for the service. Thus, banks may find it profitable not to call loans to some firms. Other firms with sufficient amounts of liquid assets may voluntarily repay and undertake adjustments. However, in either case, initially only those firms with outstanding call loans are required to bear the costs of adjustments in response to a settlement failure. Second-round adjustments among firms may involve revisions of trade-credit practices by those firms who have had to undertake portfolio adjustments in response to failure in interbank settlement arrangements.

1/Access to a lender of last resort may also cause banks to have lower adjustment costs.

II. Assignment of Risk Absent Supervisory Intervention

This part and the next examine the allocation of risk associated with payments activities. This part considers the allocation of risk between firms and banks absent regulatory or supervisory intervention. The analysis turns on the concept of finality of payment. In addition, the allocation of risk within the banking system is addressed by examining different mechanisms for allocating losses from failures to settle.

Allocation of Risk Between Firms and Banks. A payment is defined as final when a bank has unconditionally passed access to settlement assets to recipients of payments from other banks. Finality before settlement may result from individual contracts between firms and banks, collective rules of multilateral net settlement systems, or rules of law.

If payments are provisional, firms assume the risk that they will have to make portfolio adjustments should payments be revoked. If these payments are revoked, firms that have withdrawn funds from their banks must restore those funds. To the extent that restoration entails rapid portfolio changes, substantial costs may be borne by the nonbank sector. Even if provisional payments are not revoked, banks pass on the costs of their portfolio adjustments to the nonbank sector.

Firms do not necessarily avoid the problem of revocation by refusing to take out call loans against provisional payments. If a firm's portfolio or payment plans are based on the assumption that provisional payments will become final at settlement, then the revocation of these payments may still create significant adjustment costs for firms. Bank loans, of course, may be available to finance adjustment if banks can be adequately compensated.

It should be emphasized that with provisional payments, firms bear the ex post costs of adjustment regardless of the reason for the failure to settle. The failure may have originated in the bank or nonbank sector and may have been transmitted between and within the bank and nonbank sectors.

The effect of finality of payment is to assign the banking system the responsibility for the ex post liquidity adjustment costs created by failures to settle. Ex ante the banking system will charge the nonbank sector for the expected costs of these adjustments, a charge equivalent to an insurance premium for guaranteeing the completion of payment. The nonbank sector may still bear some residual risk that it will have to undertake ex post adjustment.

Allocation of Risk Among Banks. There are various mechanisms for assigning losses in the event of bank failure. When payments are final, the assignment only involves the banking sector.

Mechanisms include negotiated agreements among participants in payments networks as to the allocation of losses in the event of the failure of a member. Mechanisms also include legal processes and rules that would operate in the absence of negotiated agreements or explicit supervisory regulations. Of course, mechanisms also include any supervisory regulations that allocate losses. In general, the existence of a LLR will affect the operation of any of these mechanisms and alter the allocation of risk.

Different types of rules for the allocation of losses may be embodied in agreements governing multibank payments networks. In general, these rules reflect two decisions on the part of banks entering into the agreement. First, banks must agree on the criteria for identifying the group of banks that will bear losses in the event of a failure to settle by a network participant. Second, banks must agree on the formula for allocating losses among the group that has been identified to bear losses. One example of an allocation formula would be the sharing of losses among network participants who are net creditors of the failed bank in proportion to the amount of net credit extended. A second example of an allocation formula would be the sharing of losses equally among all participants in the network. Many other sharing formulas are conceivable. These two examples are sufficient to illustrate the important connection between rules for sharing losses and risk taking within a payments network.

Efficient risk assignment requires that decisions to take risk will be related to the losses that may be incurred. If losses are allocated to net creditors on the basis of net credit extended, then all banks that are potential net creditors will have an incentive to price extensions of net credit on the basis of risk, set limits to the amount of net credit extended, or both. When the specific linkage between the taking of risk and the bearing of losses is weakened, individual network participants will either be under-compensated, or overcompensated, for the additional risk that their payments activities may add to the particular payment network. For example, if there

is a prior agreement that losses from a failed bank will be shared equally among all network members, then the members will make decisions about the pricing and taking of risk based on average rather than marginal expected network losses. Because these two loss measures will generally differ, members may take risks for which they are not fully compensated as a group and will concentrate too much risk (for the given compensation structure) in the payments network.

III. Assignment of Risk with Supervisory Intervention

Prudential constraints may be placed on banks' participation in payments arrangements by supervisory authorities. Two types are discussed in this part. First, authorities might impose bilateral credit limitations on members of a payments network. For example, constraints on bilateral net credit exposures on a network might be based on some percentage of a bank's equity. Second, authorities might impose aggregate limitations on each bank's net position vis-a-vis all other banks in a payments network.1/ In addition, authorities might require that both types of constraints be satisfied at any given time.

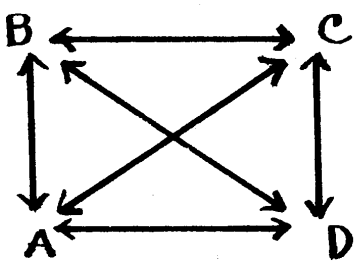
A discussion of the presence of a lender of last resort, given supervisory oversight, follows the examination of bilateral and aggregate limitations.

1/The problems of coordinating regulation across multiple networks in an economy is not explicitly treated in this paper.

Bilateral Limits. Alternative types of regulation can have quite different effects on a payments network. To amplify this point, two types of bilateral limits are examined. First, up to a bank's regulatory limit, the terms of interbank credit might be left to the market. Alternatively, banks might be required to provide free interbank credit up to the regulatory limits of other members as a precondition for their participation in particular payments arrangement.

Figure 7 is a schematic of the bilateral paths in a four bank multilateral payments network. There are six payment paths, every member has

Figure 7



bilateral contact with every other member.[1/] It is possible that supervisory bilateral constraints would not affect the payments of network members. If constraints exceed demands for credit in the free credit case, or market determined credit in the market rationing case, the constraints would not be binding.

If supervisory constraints are binding, then different regulatory schemes can have different effects. Suppose the market were allowed to

1/If N is the number of banks in a network, there are N(N-1) bilateral paths. Because one bank's credit is another bank's debit, there are at most N(N-1)/2 binding bilateral supervisory constraints in a network at a given time.

allocate credit up to the supervisory limit. Also suppose that a payment path, e.g., from bank A to bank B, were blocked by a binding supervisory constraint on bank A's bilateral net debt position to B. If a customer of bank A wants to make payment to a customer of bank B, bank A has a choice. It can reroute the payment or obtain an interbank loan. Rerouting through a bank without a blocked path, e.g., through C, however, amounts to obtaining an intrasettlement period loan from C who in turn receives an extension of credit from B. Alternatively, A might draw on interbank credit lines with bank B or some third bank such as C. When B transfers the proceeds of the loan to A, or C transfers to B who transfers to A, the supervisory constraint on A would be lifted. Assuming the loan were for the amount of A's customer's payment, the constraint would be lifted sufficiently to allow completion of the payment.

Three important points emerge from this example. First, if market constraints on interbank credit are not binding, more restrictive supervisory constraints on intrasettlement period positions in payments networks may be avoided. Unless general interbank relationships were also substantially restricted, restrictions on the rerouting of payments might also be easily avoided.1/ Second, the avoidance of supervisory constraints would presumably take place using the lowest cost method of avoidance. When credit is rationed by price, subject to the information problems discussed in Part I, low cost avoidance is likely to mean the use of low cost credit. Assuming that the cost of credit is related to risk, then the pattern of avoidance would at least spread risk across a network. Of course when payments are

1/In the presence of multiple networks, interbank deposit accounts, and a network providing finality on transfer, stopping avoidance of supervisory constraints could prove a difficult regulatory problem.

completed by rerouting or involve multiple transfers, message processing costs will rise on the network.

Contrast the regulation in which a bank's participation in a particular network puts it under an obligation to provide free credit up to a supervisory limit. There is no incentive to obtain credit from banks with the smallest exposure to the borrower.

Aggregate limits. Supervisory authorities might impose limits on the net position of each bank in a payments network vis-a-vis the aggregate of all other banks in a network. Such limits might be imposed, for example, in order to limit bank risk from participation in a payments network without attempting to regulate bilateral bank relationships.

Aggregate limits, however, suffer from the same problems of avoidance as bilateral limits. In the event that an aggregate limitation becomes binding, a bank need only obtain an interbank loan from a member of the network whose aggregate constraint was not binding and have the loan "proceeds" transferred over the network.1/ Thus credit extensions that permit the bypassing of aggregate supervisory limits would simply switch the form in which risk was assumed from net credit on a payments network to general interbank net credit. Market forces, subject to the supervisory constraints, would be responsible for limiting and distributing payments network risk.

Interaction of Loss Sharing Arrangements and Supervisory Constraints. It is important to recognize the possibility of interaction between limits on participants' intrasettlement period positions in a payments

1/Since aggregate net debt equals aggregate net credit on a network, not all aggregate constraints can be binding at one time.

network -- various forms of aggregate or bilateral limits -- and loss sharing arrangements. The interaction may affect both the overall efficiency of multilateral payments networks and the composition of participants.

The interaction can be highlighted by a comparison of two cases. In both, it is assumed that bank participants in a payments network are subject to limits on their net debt positions. These limits are assumed to be equal to a given percentage of each participant's equity. Operating under this rule, any network member may draw credit from any other member, i.e., make a payment to that member, up to its own net debt limit. Thus a large bank might receive credit from a small bank member in an amount well in excess of the small bank's equity. However, the significance of such credit extensions differ according to agreed loss sharing arrangements.

Two possible loss sharing arrangements are (1) the net creditors of a failed member are responsible for its net debt at settlement in proportion to net credit extended and (2) the net debtors of a failed participant are responsible for its net debt in proportion to net credit received. That is, under (1) the net recipients of payments would be responsible for losses, while under (2) the net senders of payments would be responsible for losses.

Networks with the combination of aggregate net borrowing limits -- based on a percentage of equity -- and loss sharing formula (1) tend to discourage smaller banks from joining the network because it is difficult for them to control potential losses. Large banks are allowed to send payments based on their own equity, creating potentially large losses for small bank net recipients.

The disincentive for small banks to join a network does not exist when net borrowing limits are combined with loss sharing formula (2). Since senders of payments bear losses, small banks can limit their potential losses by controlling the timing, routing, and volume of payments sent. Access to the payments network tends to be rationed more efficiently since banks bear risk in proportion to the risk they add to the network.

Effects due to the Presence of a Lender of Last Resort.

The reaction of payments network members to the existence of a guarantor of settlement is likely to be the discounting of risks from payments activities. If payments are final and net creditors bear network losses, net creditors will have little or no incentive to use price or quantity rationing to limit risk exposure in the presence of a guarantor. If payments are final and net debtors bear network losses, net debtors will have little or no incentive to resist sending payments in order to hold down losses from the failure of a member to settle. If payments are provisional, then a guarantor of settlement affects banks' credit evaluations of customers wanting to withdraw or transfer provisional payments. Since settlement is guaranteed, a bank receiving a payment may assume that it has no risk from allowing a customer to retransfer the funds.

In the absence of a lender of last resort, market forces will determine the use of payments services by the nonbank sector, and the organization of and membership in payments networks among banks. The allocation of risk within payments networks as well as the allocation of risk between bank and nonbank sectors will also be determined by market forces.

In the presence of a lender of last resort, market forces will still determine the use of payments services, the structures for providing the services, and the bearing of risk in connection with the provision of those services. This market allocation, however, will reflect the perceptions of the bank and nonbank sectors that net settlements of payments are guaranteed. In this alternative allocation, more payments services will be demanded by, and provided to, the nonbank sector. In addition, more nonbank sector risk will be incorporated into payments activities. Payments networks will also tend to be larger.

An important question for public policy is the rationale for supervisory controls on payments networks. As described above, the presence of a lender of last resort can seriously alter the behavior of banks in their payments network relationships and portfolio choices. Taking the presence of a LLR as given, suppose that the goal for it is to insulate the real economy from the rapid transmission of financial shocks at a minimum cost to society. One way to minimize the cost of performing the LLR function is to couple supervisory constraints on payments networks to the provision of lender of last resort services. This would be consistent with Arrow's suggestion that external constraints may be needed on economic behavior to improve efficiency when moral hazard is present.1/

1/See Arrow [1].

IV. Conclusion

A general framework for analyzing issues relating to payments networks was established in Part I. Parts II and III examined some of the issues relating to the allocation of risk and regulatory intervention using the general framework as a point of departure.

At this point, it may be useful to reiterate several conclusions that emerge from all three parts. First, the periodic net settlement of payments by banks in a multibank economy is a source of real resource savings to the economy. Second, payments networks are not isolated entities in an economy; they are part of broader interbank relationships. Analysis of the effects of regulatory policies aimed at payments networks must take these broader interbank relationships into account. Finally, the presence of a lender of last resort can fundamentally change the organization of payments activities as well as general interbank relationships. This presence is both a problem and a rationale for supervisory policy.

References

[1] Arrow, K., Essays in the Theory of Risk Bearing. Chicago: Markham Publishing Co., 1971.

[2] Association of Reserve City Bankers, Report on the Payments System. Washington, 1982.

[3] Coats, W. and Frankel, A., "The Effects of Prohibiting Reserve Account Overdrafts," Research Papers in Banking and Financial Economics. Federal Reserve Board, September, 1980.

[4] Garbade, K. and Silber, W., "The Payment System and Domestic Exchange Rates: Technological Versus Institutional Change", Journal of Monetary Economics, January, 1979.

[5] Johnson, H., "Inside Money, Outside Money, Income, Wealth and Welfare in Monetary Theory," Journal of Money, Credit, and Banking, February, 1969.

[6] Salant, S. and Henderson, D., "Market Anticipations of Government Policies and the Price of Gold, "Journal of Political Economy, August, 1978.