Federal Reserve Bank
of Minneapolis

Summer 1989

Quarterly Review

Federal Reserve Bank of Minneapolis

Quarterly Review

Vol. 13, No. 3 ISSN 0271-5287

This publication primarily presents economic research aimed at improving policymaking by the Federal Reserve System and other governmental authorities.

Produced in the Research Department. Edited by John H. Boyd, Kathleen S. Rolfe, and Inga Velde. Graphic design by Terri Gane, Public Affairs Department.

Address questions to the Research Department, Federal Reserve Bank, Minneapolis, Minnesota 55480 (telephone 612-340-2341).

Banking Without Deposit Insurance or Bank Panics: Lessons From a Model of the U.S. National Banking System

V. V. Chari
Senior Economist
Research Department
Federal Reserve Bank of Minneapolis

Bank panics were a recurrent phenomenon in the United States until the 1930s. Between 1864 and 1933, there were seven instances of systemwide runs on banks, and four of these led to widespread suspension of convertibility of deposits into currency.[1] Since 1933, however, while there have been runs on individual banks, there have been no bank panics. This turnaround has been widely attributed to deposit insurance and close regulation of banking. If this accomplishment had come at little or no social cost, a study of bank panics would be of purely historical interest, with no relevance for contemporary policy. But the recent experiences of the Federal Savings and Loan Insurance Corporation and the Federal Deposit Insurance Corporation have decisively shown that deposit insurance is not costless.

A major problem with deposit insurance is known as *moral hazard* (Kareken and Wallace 1978; Kareken 1981, 1983; Kane 1985; Benston et al. 1986; Boyd and Rolnick 1989). When bank deposits are insured, depositors have no incentive to monitor their banks' activities. Consequently, banks can, without penalty and with the possibility of greater profit, hold riskier portfolios than they otherwise might. This distortion in banks' asset choices can be mitigated by regulation and monitoring by the central bank and the deposit insurance agencies. The obvious argument against close regulation is that unfettered markets can allocate resources better than bureaucrats can. Therefore, the benefits of deposit insurance (such as eliminating bank panics) must be balanced against the costs of distorting bank decisions. In order to evaluate these costs and benefits, we need a model, or framework, within which to evaluate alternative policies.

Such a model must explain why there were bank panics in the United States until the 1930s. Neither Great Britain nor Canada had federal deposit insurance, and neither experienced systemwide bank runs as the United States did in the 19th and early 20th centuries. Clearly, then, deposit insurance is not the only arrangement that forestalls bank panics. Given this fact, it is desirable to have a model which is consistent with the sharp differences in the performance of the banking system in such similar economies as those of the United States, Great Britain, and Canada.

A promising route is to ask whether specific institutional arrangements between 1864 and 1933 contributed to bank panics. I will focus on the National Banking System in the United States from 1864 to 1913. I argue that two key features of this system played a large role in bank panics. One is that banks were prohibited from branching between states and, to a substantial extent, within states. The other key feature is that banks were permitted to meet their reserve requirements partly by deposits at national banks in designated

[1] I follow Sprague (1910) and Friedman and Schwartz (1963) in identifying bank panics in 1873, 1884, 1890, 1893, 1907, 1930, and 1933. Convertibility was suspended in 1873, 1893, 1907, and 1933.

reserve cities. Prohibiting interstate banking meant that banks were subject to community-specific variations in the demand for currency and could not diversify this risk by issuing deposits directly in other communities. Permitting reserves to be held as deposits in other banks had a subtler and more insidious effect. Such reserves could earn interest as banks in reserve cities reloaned them. However, in order to earn the interest demanded in a competitive marketplace, banks in the reserve cities held illiquid portfolios. The system as a whole thus held a more illiquid portfolio than it might have otherwise. The effect was to create a system prone to bank panics.

In this paper, I construct a model that is consistent with the institutional features of the U.S. National Banking System. My model builds on the work of Diamond and Dybvig (1983), Smith (1986), Bhattacharya and Gale (1987), and Wallace (1988) and explains why panics occurred in the United States but not in Great Britain or Canada. Essentially, this is because, unlike the United States, the other countries had no prohibitions on interstate branching. Some may argue that, without such prohibitions, the banking industry would become too concentrated,[2] so that other panic-prevention measures, such as deposit insurance, might be preferred. But my model suggests that bank panics can be eliminated without removing the prohibitions on interstate branching. I argue that a combination of effective reserve requirements, a wise discount window policy by the central bank, and a policy of occasionally restricting cash payments can eliminate bank panics. In other words, to be stable, the banking system does not need deposit insurance. That's good news for those concerned about the moral hazard problems of deposit insurance. It means a stable banking system can also be efficient.[3]

The plan of the paper is as follows. First I describe the National Banking System and its bank panics. Then I build a series of progressively more complicated models to arrive at one which captures the key features of that system. Next I describe these models more formally (in a section which readers not interested in technical details can skip without loss of continuity). And finally, I summarize the paper and draw policy implications.

Bank Panics Under the National Banking System

To review the major bank panics under the U.S. National Banking System, I rely heavily on Sprague's (1910) history of these crises. Sprague was a leading contemporary authority on bank panics. For the National Monetary Commission, he wrote a history of the panics under the National Banking System. Sprague's carefully documented work has been extensively used and cited, by Friedman and Schwartz (1963), among others. Sprague discusses five panics, which occurred in 1873, 1884, 1890, 1893, and 1907. In 1873, 1893, and 1907, suspension of convertibility was widespread. The most serious of these crises was in 1907 and led to the formation of the Federal Reserve System.

The System

Before I discuss the crises, I briefly review the main features of the National Banking System. The National Banking Act of 1864 set restrictions on banks chartered under it. These included restrictions on their branching, their capital and reserves, and their portfolios. My primary interest is in the branching and reserve restrictions. (For a general discussion of the U.S. National Banking System, see White 1983.) The National Banking Act prohibited branching across state lines (*interstate* branching), and state legislatures generally prohibited branching within a state (*intrastate* branching). Reserve requirements varied according to a bank's location. Banks in nine designated reserve cities had to hold a reserve of 25 percent of their deposits, and those elsewhere (*country* banks) had to hold 15 percent of theirs. Two-fifths of a country bank's reserves had to be held as currency and specie in their vaults, and up to three-fifths could be held as deposits in reserve city banks. In 1874, a three-tier system was established under which banks in eighteen new reserve cities could hold one-half of their required reserves as deposits in banks in New York, the central reserve city. In 1887, the law was changed yet again, and Chicago and St. Louis became central reserve cities. From my perspective, it is important to note that a substantial portion of the reserves held by country banks were, effectively, loans due from other banks.

Deposits at reserve city banks served several important purposes. When country banks had excess funds, these could be deposited with city banks and earn

---

[2]The number of banks per capita was substantially smaller in Canada than in the United States. See the paper by Williamson in this issue of the *Quarterly Review*.

[3]Stability can, of course, be achieved in many ways, but unlike my way, most others cannot also achieve efficiency. A 100 percent reserve requirement, for example, proposed by Friedman (1960), would likely eliminate bank panics. Such a proposal, however, ignores the fact that banks provide (and perhaps ought to provide) services other than facilitating transactions. One of those services is transforming illiquid assets into liquid claims. Since my model has an explicit role for banks in providing liquidity, it can be used to analyze stability and efficiency simultaneously.

interest. They also allowed banks to clear claims against one another. Perhaps most importantly, in a unit banking (no branching) system, they allowed country banks to hold a more diversified portfolio of assets. However, allowing these deposits to count toward reserve requirements led to a system which was more illiquid than was apparent.

The Panics

I turn now to the bank panics themselves. Rather than review them chronologically, I will describe the key features I think they shared.

□ *Duration*

The panics were generally short, lasting from one to three months. The longest was in 1907; the shortest, in 1884.

□ *Timing*

Most of the panics occurred in the fall of the year. As Sprague remarks (1910, p. 127), "With few exceptions all our crises . . . have occurred in the autumn, when the western banks, through the sale of the cereal crops, were in a position to withdraw large sums of money from the East." A more accurate interpretation, I think, is that the demand for currency in agricultural areas fluctuated seasonally, being particularly high in the spring and fall. This view is supported by the fact that the deposits of country banks in New York were generally low in the spring, rose in the summer, and fell to their lowest level in October (Sprague 1910, p. 22). Of course, the magnitude of this seasonal movement was not predictable.

□ *Proximate Cause*

Contemporary observers viewed the loss of confidence in banks as originating in the failure of one or more financial institutions in New York City. Sprague (1910, pp. 49, 110, 142, 164, 252) cites these failures: 1873, Jay Cooke & Co.; 1884, Grant & Ward; 1890, Decker, Howell, & Co.; 1893, the National Cordage Company; 1907, the Knickerbocker Trust Company. Of course, none of the observers mentioned the many failures during the period which did not lead to panics, so it is difficult to judge the importance of this view.

□ *Response of the Banking System*

When banks perceived that a crisis was coming, they turned to their clearinghouses to set up a mechanism to protect themselves. The particular mechanism they resorted to was the issuing of clearinghouse loan certificates. Clearinghouses issued these certificates against securities deposited by banks. (See Gorton 1985 and Gorton and Mullineaux 1987 for a more detailed description of the role of the clearinghouses.)

For example, in 1873, banks in the New York Clearing House Association issued the following statement: "Any bank in the Clearing House Association may, at its option, deposit with a committee of five persons, to be appointed for that purpose, an amount of bills receivable or other securities, . . . [and] said committee . . . shall be authorized to issue . . . certificates of deposit bearing interest at 7 per cent per annum . . . to an amount not in excess of 75 per cent of the securities" (Sprague 1910, p. 45). The committee was also authorized to issue certificates up to par value against U.S. government bonds or gold. These certificates could be used to settle unfavorable balances at the clearinghouse.

In effect, clearinghouses attempted to act as a central bank might have. While they could not issue legal tender, they attempted to go as far as they could in that direction.

Until the panic of 1893, these certificates were used solely in interbank settlements. However, in the panics of 1893 and 1907, loan certificates were issued in small denominations by banks to their customers. These small-denomination certificates were, in turn, backed by large-denomination clearinghouse certificates. This issue of paper money was clearly illegal, but the federal government looked the other way. These currency substitutes were issued in relatively large amounts. For example, at the height of the 1907 crisis, in December, \$750 million of clearinghouse instruments of various types were issued. This is a large amount compared to the stock of currency and reserves, which was \$3,069 million (White 1983, p. 79).

Toward the end of the National Banking System period, financial innovations made it more difficult for clearinghouses to coordinate actions among financial institutions. The most important of these innovations was the formation of trust companies. The trust companies received deposits of money in trust to purchase securities of business firms and generally engaged in a variety of banking activities. The assets of these companies grew rapidly in the last 20 years of the National Banking System period (White 1983, pp. 38–40), primarily because they were subject to lower reserve requirements. In 1903, however, the New York Clearing House Association adopted a rule which required trust companies clearing through its member banks to increase their reserves. Rather than meet that condition, nearly all the trust companies gave up clearinghouse privileges.

The most important exception was the Knickerbocker Trust Company. When this company ran into difficulties in 1907, it sought a loan from its clearing agent, the National Bank of Commerce. When that bank refused to make the loan, a run on the trust company began which rapidly spread to all the banks. Sprague (1910, p. 252) argues forcefully that if Knickerbocker had been a bank and a member of the clearinghouse, it would have received a loan and the crisis been averted.

Note the parallel here with arguments today that financial innovation outside the purview of the Federal Reserve System endangers the financial system as a whole. But remember, it is far from established that the failure of one financial institution necessarily endangers the entire system.

□ *Suspension of Convertibility*

As pointed out earlier, not every panic led to a suspension of the convertibility of deposits into currency. But in 1873, 1893, and 1907, the clearinghouses judged conditions serious enough to suspend conversion. The suspension was carried out, for example, in 1873 when the New York Clearing House Association adopted the following resolution: "That all checks when certified by any bank shall be first stamped or written 'Payable through the Clearing House' " (Sprague 1910, p. 54). This resolution did not mean that no cash payments would be made, but it did place the decision to pay currency firmly with the clearinghouse committee. Still, suspension was never quite complete. Generally, banks were allowed to issue limited amounts of currency to depositors and to employers to meet payrolls.

□ *Currency Premium*

Perhaps the most interesting feature of the panics documented by Sprague is the size of the currency premium. When banks suspended payment, money brokers set up business in every major city, offering to buy and sell currency at a premium for deposits. These premia were quoted in the newspapers, and Sprague collected them. The size of these premia is one indication of the severity of the crises. In 1873, the premium averaged 1 percent; in 1893 and 1907, it averaged 2 percent (Sprague 1910, pp. 57, 187, 280). On no date did it exceed 4 percent. Since depositors could obtain currency simply by selling their deposits to brokers, the currency premium reflects the wealth lost by depositors.

□ *Effect on Real Activity*

Bank panics do not affect financial institutions alone. The resulting contraction in loans, as well as the difficulty of securing currency to meet payrolls and make transactions, may also have serious effects on the nonfinancial sectors of the economy. But the post–World War II experience should convince us that eliminating bank panics does not eliminate business cycles. There is some evidence that fluctuations in economic activity were not dramatically larger before World War II than after (Romer 1989). Furthermore, there is reason to believe that bank panics do not cause recessions (Gorton 1988).

However, examining the behavior of real activity during panics is still interesting. Generally, real activity declined fairly substantially. For example, during the 1873 panic, loans by national banks fell 9 percent (Sprague 1910, p. 83). In the panic of 1907, though, these loans fell only 2 percent (Sprague 1910, pp. 305, 308, 310).

□ *Contemporary Policy Implications*

In 1873, the New York Clearing House Association formed a committee to recommend policy changes "to increase the security of . . . business" (Sprague 1910, p. 91). The main reform suggested by the committee was to prohibit the payment of interest on deposits. I suspect that the main reason for this recommendation was to form a cartel to eliminate price competition. Of course, the committee did not openly use that argument. Instead, it pointed out that "currency is superabundant in summer . . . and can not . . . [be kept] idle without loss of the interest paid to its owners . . . . There is consequently no resource but to loan it in Wall street upon stocks and bonds, in doing which so much of the nation's movable capital passes for the time into fixed and immovable forms of investment and its essential character is instantly changed" (Sprague 1910, p. 92). In effect, the committee argued that competitive forces drive the system to more illiquidity than is socially optimal.

But why should this be so? Surely, depositors can avoid banks which have illiquid portfolios if they want to. The committee had a response to this argument: "The abandonment of the practice of paying interest upon deposits will remove a great inducement to divide . . . reserves between cash in hand and deposits in cities, and make the banks throughout the country what they should always be, financial outposts to strengthen the general situation" (Sprague 1910, p. 97). Here we may see the one valid argument for prohibiting the payment of interest on deposits. The payment of interest effectively encouraged banks to evade reserve requirements. As the committee argued, "The aggregate [reserves] held by all the national banks of the United States does not finally much exceed 10 per cent of their direct liabilities, without reference to the large amount of debt which is otherwise dependent upon the same reserves" (Sprague 1910, pp. 96–97). If the committee was right, the actual reserves fell far short of the statutory requirement.

The recommendations of the 1873 committee went largely unheeded. Similar recommendations were made and ignored in subsequent panics until Congress reformed the system in 1913.

Modeling Banking Informally . . .

Now I construct a series of progressively more complicated models, ending with one which captures the key features of the National Banking System. I build very heavily on Diamond and Dybvig's (1983) model, which itself has two key features.[4] One is that individuals are uncertain about when they will want to make expenditures. The other is that real investments yield high returns over long horizons, but low returns if they are liquidated early.

The Diamond and Dybvig Model

I begin by describing my version of Diamond and Dybvig's model. This model has three periods, which I refer to as the *planning* period (period 0), the *intermediate* period (period 1), and the *final* period (period 2). There are N people in the economy. In the planning period, each person is endowed with one unit of a good—say, seed corn. There are two technologies available. One I call *illiquid.* Using this technology, people plant seed corn in the planning period. Part or all of it can be dug up and eaten in the intermediate period, when it yields a rate of return of one unit. The remainder grows at rate $R > 1$ into the final period. The other technology I call *liquid.* With this technology, a unit of investment also yields a unit of output in the intermediate period. Part or all of it can be consumed then, and the remainder yields a rate of return of only one unit into the final period. This technology is clearly dominated by the illiquid technology. Since the liquid technology plays a role only in my discussion of reserve requirements, I suppress it until then.

The economy has two kinds of people: *impatient* and *patient.* Impatient people want to consume in the intermediate period. Patient people are indifferent about consumption in the intermediate and final periods; that is, they are indifferent to a choice between consuming one unit in period 1 or one unit in period 2. No one knows in the planning period whether any particular person will be impatient or patient. Each person believes that the probability of being impatient is t. Therefore, if the number of people in the economy is large, tN will be impatient and $(1-t)N$ will be patient. Individuals learn their own type in the intermediate period.

Consider, now, the problem facing this society in the planning period. One possibility is *autarky;* each person could act alone. Each person could then look forward to consuming one unit in the intermediate period (if impatient) or R units in the final period (if patient). Note that, because $R > 1$, no patient people would consume early. However, everybody might be better off if they pool their resources, in an *insurance* arrangement, for example. Suppose everybody agreed that impatient people are entitled to consume more than one unit (say, x_1 units each) in the intermediate period and patient people to share the remaining proceeds in the final period. The amount, x_2, received by each patient person is, then, given by

$$(1) \qquad N(1-t)x_2 = R(N-tNx_1)$$

where the term in the parentheses on the right is what is left over in period 1 after the impatient are paid off.

Alternatively, we can rewrite equation (1) as

$$(2) \qquad x_2 = R(1-tx_1)/(1-t).$$

Note from equation (2) that if $x_1 > 1$, then $x_2 < R$. An insurance arrangement might be preferred to autarky because people are willing to forgo some of their consumption if they turn out to be patient in return for greater consumption if they turn out to be impatient. The precise insurance contract entered into depends, of course, on the preferences of the two types and on the value of t. Under reasonable restrictions on people's attitudes toward risk, it can be shown that an *optimal* contract has the following properties:

$$(3) \qquad 1 < x_1 < x_2$$

$$(4) \qquad x_2/x_1 < R.$$

---

[4]My description of Diamond and Dybvig's (1983) model closely follows that of Wallace (1988), who develops a detailed, useful, and entertaining analogy with a camping trip. My description is, therefore, brief.

How might the society described here implement this contract? The following scenario might help. In the planning period, everybody plants corn in a common field. The society has access to a programmable robot which guards the field and can harvest the corn. The robot also provides services similar to a cash machine. That is, in the intermediate period, people arrive randomly, and the robot gives x_1 units to each person who demands corn in that period. The robot records the identity of each person who withdraws and pays all others x_2 units in the final period.

Given these rules, turn now to behavior in the intermediate period. Impatient people are best off consuming in this period and so will arrive at the field to collect their shares. Patient people's problem is somewhat more delicate. If any individual patient person believes that no other patient people will show up, then that person is best off waiting to consume in the final period, since $x_2 > x_1$. (Recall that patient people care equally about consumption in the two periods.) Such outcomes are described as *individually incentive compatible,* since no individual has an incentive to deviate from the rules of the game. However, what if a patient person believes that all other patient people intend to show up in the intermediate period? Since $x_1 > 1$, if that happened, then not all the claims could be met in the intermediate period and, more important, there would be no seed corn in the final period. So any individual patient person is best off showing up at the field in the intermediate period and hoping to be early enough in line to get some consumption rather than none; that is, patient people are best off acting as if they were impatient. The point is that this story has potential for explaining bank panics. The fear that everybody might withdraw early is enough to cause everybody to want to withdraw early.

But we should proceed slowly. An exceedingly simple mechanism could prevent bank panics. Suppose we program the robot to pay x_1 units to the first tN people who show up in the intermediate period and nothing to anyone else who shows up in that period. Such a mechanism resembles *suspension of convertibility* of deposits into currency, a widespread practice throughout the National Banking System period. With the rules of the game augmented by suspension, impatient people would still show up in the intermediate period to claim x_1 units of consumption. What about the patient? Regardless of any individual patient person's beliefs about the actions of other patient people, that person is assured of x_2 units in the final period. Since $x_2 > x_1$, early withdrawal never pays off for a patient person. Therefore, no phenomenon resembling bank panics can possibly occur.

I now list the key properties and implications of this model.

One is that suspension of convertibility never occurs. The mere threat to suspend is enough to ensure that no patient people will withdraw early.

Also, the model explains why banks hold illiquid portfolios. Since $x_1 > 1$, short-term deposits have higher yields than the technology permits. Therefore, not all deposits can be paid off. The cost of these high, early payments is that $x_2 < R$, so that long-lived deposits have lower yields than the technology permits. People willingly accept these relatively low, late returns for relatively high, early returns. We could, of course, require that $x_1 = 1$ and $x_2 = R$; such a policy is technologically feasible, but undesirable.

In addition, as Jacklin (1987) and Wallace (1988) have emphasized, the arrangement described in the model is inconsistent with well-functioning asset markets in the intermediate period. To see why this is so, suppose that, with such an arrangement in place, one person deviates in the planning period and plants corn autarkically. If that person turns out to be patient, then the return in the final period is R, which is greater than x_2. If the person turns out to be impatient, then all the person needs to do is contact a patient person who is part of the arrangement and offer the following trade: Go to the robot, withdraw your deposits, and give them to me in exchange for my cornfield. A patient person would be more than willing to make such a trade since the person's return in the final period would then be R units, which is greater than x_2. The impatient person would receive x_1 units, which is greater than the crop currently available. Therefore, the model's arrangement is viable only if certain kinds of asset markets function poorly.

However, the assumption that asset markets do not exist should not disturb us too much. As Wallace (1988, p. 16) argues, "Inconsistency between banking, on the one hand, and well-functioning markets, on the other, should not be surprising. Almost any story about the role of banking has implicit in it that markets are costly to participate in or are incomplete in some way." One specific way to model incompleteness of markets, as Wallace does, is to assume that people are isolated from each other in the intermediate period. (I return to this theme below.)

Finally, the result that threatened suspension eliminates all bank panics is surely too strong. It appears to rely on the assumption that a fixed fraction t of people are impatient. What if this fraction were, instead, random, so that uncertainty, or risk, enters the model in an aggregate sense? Diamond and Dybvig (1983) have argued that a policy of suspension could lead to some impatient people not consuming in the intermediate period and is therefore undesirable. They argued that when the fraction of impatient people is random, payments to each depositor should be contingent on the total number of withdrawals in the intermediate period. But with an isolation assumption or a sequential service constraint taken seriously, payments to a person in the intermediate period cannot be made contingent on the number of future withdrawals.

The Wallace Model

Wallace (1988) made some headway in characterizing the optimal insurance contract when aggregate risk exists. I now provide a full characterization and show that, with random withdrawals, bank panics are possible and suspension does not eliminate them entirely.

To construct a model with aggregate risk, I modify the environment more or less as Wallace did. Suppose that the number of impatient people in the population can take one of two values, t_1N or t_2N, with $t_2 > t_1$. The probability of each of these events is given by p_1 and p_2, respectively.

Consider, first, the case where, at the beginning of the intermediate period, everybody gets to know the state of the world; that is, whether the number of impatient people will be t_1N or t_2N is public information. What does the optimal contract look like? In general, making the same payments to impatient people in both states will be undesirable because the risks to be shared in the two states are different. So all we need do is program the robot to make payments contingent on the two states. Denote the payments in the two states $(x_1(t_1), x_2(t_1))$ and $(x_1(t_2), x_2(t_2))$. Under appropriate assumptions on preferences, these allocations satisfy

(5) $\quad 1 < x_1(t_1) < x_2(t_1) < R$

(6) $\quad x_2(t_1)/x_1(t_1) < R$

and

(7) $\quad 1 < x_1(t_2) < x_2(t_2) < R$

(8) $\quad x_2(t_2)/x_1(t_2) < R.$

In effect, we solve for the optimal contract separately in the two states, just as we did in the economy without random withdrawals. Again, we can program our robot to suspend conversion just as in the nonrandom case. However, with a random number of withdrawers, the point at which suspension occurs is state-contingent: in state 1, convertibility is suspended after payments are made to t_1N withdrawers; in state 2, after payments are made to t_2N withdrawers. Such a policy again avoids panics since in both states patient people get a higher return by waiting.

Turn now to the case where the aggregate number of impatient people is not public information. In this case, when the first person shows up at the cornfield, the robot does not know whether t_1 or t_2 has been realized. The robot must, therefore, be programmed to pay the same amount to the first person in both states. Similarly, if $t_1N \geq 2$, the robot does not know which state has been realized by the arrival of the second person. So the second person must get the same amount in both states. Of course, in principle, the amount received by the second person could be different from that received by the first person; we will show, though, that, in the optimal contract, these amounts are the same. In any event, the point is that payments to individuals can depend on the number of predecessors in line, but not on the number of successors.

What does the optimal contract look like now? An equal amount is paid to the first t_1N withdrawers. Paying the same amount to such individuals in both states is desirable because people are risk averse in the model, and it is feasible. *Risk aversion* here simply means that people prefer constant consumption to a lottery which has the same average value. Since we can feasibly give everybody the average value of consumption, it is optimal to do so. However, consider the problem we face when the next depositor shows up, the one after the first t_1N withdrawers. If we assume that only impatient people show up in the intermediate period, we know that the state is t_2. If we pay this person the same amount as we paid the first t_1N withdrawers, we make the patient people worse off. In general, then, we will pay this person a different amount. Proceeding as we did earlier, we pay a constant amount to the remaining withdrawers, who number $(t_2-t_1)N$.

The optimal contract is now characterized by four numbers. These are the amount paid to the first t_1N withdrawers, denoted x_{11}; the amount paid to the next $(t_2-t_1)N$ withdrawers, x_{12}; the amount paid to patient people in the final period if t_1 has occurred, x_{21}; and the amount paid to patient people in the final period if t_2 has occurred, x_{22}. (Here the first subscript denotes the period; the second, the state.)

Technological feasibility now requires that the contract satisfy these conditions:

(9) $$(N-t_1N)x_{21} = R(N-t_1Nx_{11})$$

(10) $$(N-t_2N)x_{22} = R[(N-t_1Nx_{11}) - (t_2-t_1)Nx_{12}].$$

The term in the parentheses on the right side of equation (9) is the total amount of seed corn left after payments to the impatient in state 1. These resources grow by R into the final period, when they are distributed to the patient. Similarly, the term in brackets on the right side of equation (10) is the total payments to the impatient in state 2. Notice that the first t_1 impatient people are paid x_{11} and the next $t_2 - t_1$ are paid x_{12}. These resources, again augmented by R, are paid to patient people in the final period.

Turn now to the individual incentive-compatibility conditions that an optimal contract must satisfy. Recall that with nonrandom withdrawals, the optimal contract satisfied $x_2 > x_1$ and patient people never withdrew early. I set up the problem with random withdrawals as follows. At the beginning of the intermediate period, people must decide whether or not to visit the cornfield. If they decide to visit, they arrive there randomly. They see the current rate being paid for withdrawals and decide whether or not to withdraw. Obviously, the impatient will visit the cornfield and withdraw. What about the patient? If p_2 is sufficiently small, then under the optimal contract, patient people are best off not visiting if they believe other patient people will not visit. Essentially, this is because, in the optimal contract, $x_{11} < x_{21}$. (Note the similarity with the nonrandom case.) Since p_2 is small, most of the time a patient person receives a smaller amount by visiting and withdrawing than by waiting until the final period. The optimal contract also has $x_{12} < x_{22}$. Therefore, even if a patient person does visit, that person is best off not withdrawing. If the current payment is x_{11}, the patient person is best off waiting until the final period to withdraw, because most of the time the future payment will be higher. If the current payment is x_{12}, again the person is best off waiting.

The optimal contract has several significant features. First, as in the nonrandom case, here the arrangement explains illiquid bank portfolios. Second, as is also true in the nonrandom case, this arrangement is inconsistent with well-functioning asset markets in the intermediate period. Third, and the most questionable aspect of the model, here returns depend on the amount of prior withdrawals. Wallace (1988) has argued that this is consistent with partial suspension of convertibility. Historically, when conversion was suspended, limited withdrawals were usually allowed. Therefore, returns did depend on the withdrawer's position in line. From this perspective, the model can be interpreted as partially suspending conversion after t_1N withdrawals. Wallace argues that, in this sense, the model is consistent with observed institutional arrangements. However, I argue below that the assumption of large, aggregate fluctuations in the demand for currency is implausible.

Are bank panics possible in this environment? Clearly, the optimal contract specifies suspension once t_2N people have withdrawn. But what if a patient person sees a current return of x_{11} and believes that everybody will attempt to withdraw? If the person is right, then waiting until the final period will bring the person x_{22} instead of the x_{11} available immediately. If x_{11} is greater than x_{22}, a patient person is obviously best off withdrawing immediately. I show below that there are examples of economies in which, under the optimal contract, x_{11} is greater than x_{22}. Therefore, with suspension, bank panics are possible.

Notice that, in the event of a panic, some of the impatient are left empty-handed. Being in line after t_2N withdrawals have been made, they receive nothing in the intermediate period, which is when they want to consume. The bank panics, therefore, impose *social costs:* some people's consumption decisions are distorted away from the desirable outcome.

Note that beliefs about other people's actions are essential in generating the bad outcome. If any patient person believes no other patient people will withdraw, then that person is best off waiting and accepting the small odds of receiving a low future return. If a patient person believes all other patient people will withdraw, then immediate withdrawal is best. In this sense, the model captures the notion that panics are created by a lack of confidence in the banking system (here, a lack of confidence in the behavior of other depositors).

Such panics can be avoided at a social cost. For example, we could suspend conversion after t_1N depositors have withdrawn. Since $x_{11} < x_{12}$, no patient people will withdraw early. However, some impatient people may not receive any payment in the intermediate period. Therefore, such an arrangement imposes a cost on society. The benefit is that bank panics are avoided. If the probability of the t_2 state is small and that of a bank panic sufficiently large, a scheme of suspension after t_1N withdrawals may well dominate an arrangement which ignores the possibility of runs.

Alternatively, we could design other contracts which eliminate bank panics. For example, we could choose a low value for x_{11} and a higher value for x_{22}. Such a contract avoids bank panics, but may well be undesirable since it forgoes the insurance benefits obtainable from the optimal contract.

A Model With Community Risk

The model with random withdrawals and aggregate risk described above has at least two implausible features. One is the assumption that the economy is subject to large swings in the desired amount of currency; the aggregate U.S. data, particularly since World War II, simply don't support that. The other implausible feature is the result that bank panics can be avoided only at a (possibly high) social cost. This is difficult to square with the apparently satisfactory performance of the banking industry in the United States for over 40 years since World War II—not to mention those of Great Britain and Canada throughout the 19th and 20th centuries.

Still, the idea that the demand for currency can vary within communities is not implausible. In the second half of the 19th century, an important source of these variations was agriculture. The demand for farm loans rose during the planting season and fell at the harvest. Since cash was required for many farm transactions, the demand for currency in agricultural communities was high at both planting and harvesting times and low at other times of the year (Sprague 1910, pp. 19–20). Because of the prohibition of interstate branching, these community-specific fluctuations were also bank-specific.

I model the bank-specific fluctuations as follows. Suppose that there are a large number of communities, each of which looks exactly like the modified-Wallace environment with aggregate risk. That is, there are N people in each community, and the fraction of impatient people in each community can take on one of two values, t_1 or t_2, with probabilities p_1 and p_2, respectively. However, because there are a large number of such communities, the aggregate economy does not have random withdrawals. That is, the fraction of the economywide population who are impatient is $t = p_1t_1 + p_2t_2$. Therefore, if the communities could pool their resources, their collective economy would look exactly like the Diamond and Dybvig model.

Assume, however, that there are barriers to such resource-pooling. Specifically, assume that people living in any community can only plant corn there. (This assumption is meant to capture the idea that a Minnesotan would rather use a bank in Minnesota than one in Puerto Rico.) The robots can, though, transfer resources between cornfields in the intermediate and final periods. The model thus has an interbank borrowing and lending market. The information assumption within each community is exactly that made above: individuals know only their own type. The information assumptions across communities are subtler. The fraction of early withdrawers in any community is unobservable by other communities. Furthermore, so are the payments to depositors.

Given these information assumptions, what kinds of contracts will be offered to depositors? I claim that the contracts will be exactly the same as in the single-community environment with aggregate risk. Note that offering such contracts is feasible. Now suppose that some other contract were offered. The results will be the same under any mechanism, but they are clearest if we simply assume that communities can borrow and lend from each other in the intermediate period at some interest rate r. I claim that $r = R$. Suppose, by way of contradiction, that $r < R$. Then no community with a realization of t_1 will lend because, by retaining the corn in the ground, it gets a larger return. Suppose that $r > R$. The same argument tells us that then no community will borrow. Therefore, r must equal R. But if this is true, then no community can do better than it does by acting autarkically. In other words, optimal contracts are likely to display features resembling bank panics.

In this model with community risk, however, a simple mechanism can eliminate bank panics. The mechanism involves an inferior technology: the so-called liquid, or *reserves,* technology. Recall that this technology produces one unit of output in the next period from one unit of investment in the current period. I assume that in the planning period the amount invested by each community in the reserves technology is observable. Say that each community agrees to invest a fraction t of its resources in the reserves technology (where t still equals $p_1t_1 + p_2t_2$). Depositors who withdraw in the intermediate period receive x_1 units, and those who withdraw later receive x_2 units, where the pair (x_1, x_2) solves the programming problem in the Diamond and Dybvig economy without aggregate risk. All communities agree to borrow and lend at the interest rate x_2/x_1 in the intermediate period. (Recall that $x_2 > x_1$, so that $x_2/x_1 > 1$.)

I argue that this agreement is incentive compatible. Any community which has t_1 withdrawers is happy to lend its remaining reserves at this interest rate, and communities which have t_2 withdrawers borrow at this interest rate. If the demand for withdrawals exceeds the supply of funds—that is, if more than a fraction t of individuals in the entire economy chooses to withdraw—then convertibility is suspended. Given this arrangement, it is clear that the allocations are feasible and incentive compatible and do not involve bank panics.

This risk-sharing arrangement yields the best possible outcomes and avoids bank panics. It depends very heavily on the observability of investment in the reserves technology. Every community has an incentive to invest entirely in the illiquid technology and borrow at x_2/x_1 in the marketplace because $x_2/x_1 < R$. So if good outcomes are to be sustained, there must be some means of collectively enforcing adequate investment in the reserves technology.

. . . And Formally

Here I present a detailed technical description and analysis of the models with aggregate and community risk.

The Model With Aggregate Risk

The model with aggregate risk has three periods—labeled 0, 1, and 2—with one good per period. The periods are referred to as the *planning, intermediate,* and *final* periods, respectively. The economy has N people, and each person is endowed with one unit of the planning period good.

□ *Technology*

All production occurs at a central location, and everyone in the economy has access to two technologies.

One is the *illiquid* technology. I denote the illiquid technology set by Y. Let y_t be an output (if negative, an input) of the period t good. The triplet (y_0, y_1, y_2) is in the illiquid technology set if it equals $(-y, \lambda y, R(1-\lambda)y)$ for some $y \geq 0$ and all $\lambda \in [0, 1]$. Here, the investment in the planning period is y; a fraction λ of the investment is withdrawn in the intermediate period, yielding an output of λy; and the remainder is withdrawn in the final period and returns $R(1-\lambda)y$. I assume that $R > 1$.

The other technology is, of course, *liquid*. The liquid technology set, denoted by L, is described by the triplet $(l_0, l_1, l_2) = (-l, \lambda l, (1-\lambda)l)$ for some $l \geq 0$ and all $\lambda \in [0, 1]$. The liquid technology yields one unit of the good in either period for each unit of the investment.

Notice that the illiquid technology dominates the liquid technology.

□ *Preferences*

People in this economy care only about consumption c_t in the intermediate and final periods. Each person is one of two types. *Type 1* people, whom I shall also call *impatient,* have preferences given by the utility function $u(c_1 + \delta c_2)$, where δ is close to zero. *Type 2* people, also called *patient,* have preferences given by the utility function $u(c_1 + c_2)$. I assume that the utility function displays relative risk aversion greater than unity; that is, $-cU''(c)/U'(c) > 1$. An example of a utility function satisfying this condition is $c^{1-a}/(1-a)$ for $a > 1$. An example of one failing to satisfy the condition is $\log c$. I also assume that $\lim_{c\to 0} U'(c) = \infty$.

The fraction of the population that is Type 1 is a random variable which takes on one of two values, t_1 or t_2, where $0 < t_1 < t_2 < 1$, with probability p_1 and p_2, respectively. Individuals learn their own type at the beginning of the intermediate period. Therefore, during the planning period, all people are identical. Each person seeks to maximize the expected utility of consumption.

□ *Information*

Although people learn their own types, no one else does. A person's type is private information.

□ *Isolation*

In the intermediate period, people are isolated from each other, but each person can choose to contact the central location at which all production occurs. The decision to contact the production location must be made at the beginning of the intermediate period. This decision is private information to each person. Once people have made this decision, they arrive at the production location at some random instant during the intermediate period. They consume as soon as they arrive there. People arrive at the production location at a constant rate of one person per unit of time, independent of the number of people who have chosen to contact the production location.[5] Once the stock of contacting people is exhausted, the arrival rate drops to zero.

The constant arrival rate implies that the aggregate state of the economy cannot be inferred from the arrival rate. Because I assume that people arrive at the production location randomly and because they consume the instant they arrive, I preclude the possibility of setting up a credit market in the intermediate period. I make

[5] As an example of this sort of constant arrival rate, think of people traveling to the production location on a crowded highway which they enter from several ramps. Each entrance ramp allows only one car per green light. Thus, the arrival rate is independent of the number of people in line.

this assumption to capture the notion that a demand deposit provides the holder with the ability to withdraw at any time. The assumption is also consistent with the notion that people hold liquid assets because they may need to consume at times and places at which accessing asset markets is difficult or costly.

□ *Sequential Service*

The period 0 investment can be made collectively, by a group (or all) of the individuals. The amount invested grows according to the technology. A programmable robot acts like a cash machine in this economy. It makes payments to withdrawers in periods 1 and 2 and can check a person's account to make sure that person is entitled to make a withdrawal. Payments to individuals can depend only on the number of people who have previously withdrawn, not on the number of prospective withdrawers. I call this last assumption the *sequential service constraint.* In the Appendix, I describe a class of environments for which every efficient mechanism respects the sequential service constraint.[6]

□ *Efficient Outcomes*

Here I analyze efficient outcomes in the model with aggregate risk. *Efficient* outcomes are those which maximize expected utility over the planning period subject to the technology and information constraints. I restrict attention to *symmetric* efficient outcomes, those which treat everybody in period 0 the same.

I now describe the *space of allocations.* At the beginning of period 1, after types are realized, people must decide whether or not to visit the production location. Let V_i be an indicator variable which indicates whether or not a person of Type i—where $i = 1, 2$—visits the production location, and let $V_i = 0$ indicate a decision to visit and $V_i = 1$, a decision not to visit. Once a decision to visit is made, people arrive at the production location randomly and announce their types (possibly falsely). Payments to a person in the intermediate period can depend only on the number of reported Type 1 predecessors and on that person's reported type. Denote the number of predecessors by s, the reported type of a person by i, and the payment in period 1 by $x_1(i, s)$. In the final period, the aggregate number of people reporting as Type 1 in the intermediate period is also known. Denote the payment function in the final period by $x_2(i, s, t)$. The space of allocations is, then, a pair of numbers (V_1, V_2), which represent decisions to visit the production location for each type, and a pair of consumption allocation functions $(x_1(i, s), x_2(i, s, t))$.

Here I focus on the case where only Type 1 people visit the production location. (In the Appendix, I show that an allocation mechanism with this property is efficient for a class of environments.) Of course, if only Type 1 people visit, then Type 2 people consume nothing in the intermediate period. From this fact, it also follows that Type 2 people receive payments in the final period which take on one of two values, depending on the number of Type 1 withdrawals in the intermediate period. I denote payments to Type 2 people in the final period, when the number of Type 1 people is t_1 or t_2, by x_{21} and x_{22}.

The resource constraints are, then, written this way:

$$(11)\quad \int_0^{t_1} x_1(1, s)\, ds + R^{-1}\Big[\int_0^{t_1} x_2(1, s, t_1)\, ds + (1-t_1)x_{21}\Big] = 1$$

$$(12)\quad \int_0^{t_2} x_2(1, s)\, ds + R^{-1}\Big[\int_0^{t_2} x_2(1, s, t_2)\, ds + (1-t_2)x_{22}\Big] = 1$$

$$(13)\quad x_1 \geq 0$$

$$(14)\quad x_2 \geq 0.$$

Equation (11) is the resource constraint if the number of Type 1 people is t_1. The first term on the left is the sum of payments to Type 1 people in period 1. The term in the brackets is the sum of payments to Type 1 and Type 2 people in period 2, the final period. Note that payments in period 1 can depend only on the number of predecessors, denoted s. But payments in period 2 can depend both on when a person shows up at the production location and on the aggregate number of Type 1 people. Equation (12) is the resource constraint in state t_2. Equations (13) and (14) embody the assumption that people's endowments are zero in periods 1 and 2.

The expected utility of a typical person in the planning period is given by

---

[6] An *efficient* mechanism is one which yields highest planning period utility among all mechanisms which respect the information and technology restrictions on the economy. You may wonder whether any mechanisms which respect the isolation assumption fail to respect the sequential service constraint. There is at least one such mechanism. It requires everybody to show up at the production location and report their type on arrival. Even though the arrival rate is constant, in a truth-telling equilibrium, it is possible very early in period 1 to infer the relative proportions of patient and impatient types in the population and, therefore, the aggregate state of the economy. This kind of mechanism can be ruled out by imposing a small cost of traveling to the production location. I impose such a cost in the environment described in the Appendix.

$$U(x_1, x_2) = p_1\Big[\int_0^{t_1} U(x_1(1,s) + \delta x_2(1,s,t_1))\,ds + (1-t_1)U(x_{21})\Big] + p_2\Big[\int_0^{t_2} U(x_1(1,s) + \delta x_2(1,s,t_2))\,ds + (1-t_2)U(x_{22})\Big]. \tag{15}$$

Consider now the problem of maximizing (15) subject to (11)–(14). The properties of the solution are easily established by studying three of the first-order conditions. For $s \in [0, t_1]$, the first-order condition with respect to $x_1(1, s)$ is

$$p_1 U'(x_1(1,s) + \delta x_2(1,s,t_1)) + p_2 U'(x_1(1,s) + \delta x_2(1,s,t_1)) - \lambda_1 \leq 0. \tag{16}$$

Condition (16) holds with strict equality if $x_1(1, s) > 0$. The Lagrange multiplier λ_1 is that associated with constraint (11). The first-order condition with respect to $x_2(1, s, t_1)$ for $s \in [0, t_1]$ is

$$p_1 \delta U'(x_1(1,s) + \delta x_2(1,s,t_1)) - (\lambda_1/R) \leq 0. \tag{17}$$

Condition (17) holds with strict equality if $x_2(1, s, t_1) > 0$. An immediate implication of (16) and (17) is that if $\delta < 1/R$, then $x_2(1, s, t_1) = 0$. Similarly, $x_2(1, s, t_2) = 0$. Therefore, as is to be expected, Type 1 people consume nothing in the final period when δ is sufficiently small. So, in what follows, I will set these variables equal to zero.

With respect to $x_1(1, s)$ for $s \in [t_1, t_2]$, the following condition holds [again, with strict equality if $x_1(1, s) > 0$]:

$$p_2 U'(x_1(1,s)) - \lambda_2 \leq 0 \tag{18}$$

where λ_2 is the Lagrange multiplier associated with constraint (12). Given the assumption that the marginal utility of consumption goes to infinity as consumption goes to zero, (16) and (18) hold with equality. An immediate implication of (16) and (18) is that $x_1(1, s)$ is constant for $s \in [0, t_1]$ and also for $s \in [t_1, t_2]$. That is, Type 1 people who arrive at the production location with t_1 or fewer predecessors receive a constant payment—say, x_{11}—and those who arrive with more than t_1 predecessors receive another constant payment—say, x_{12}.

We can also write the first-order conditions for x_{21} and x_{22}. I collect all the first-order conditions and write them as follows:

$$U'(x_{11}) = p_2 U'(x_{12}) + p_1 R U'(x_{21}) \tag{19}$$

$$U'(x_{22}) = R U'(x_{12}). \tag{20}$$

Equations (11), (12), (19), and (20) uniquely characterize the solution to the maximization problem. The properties of this solution can be understood by looking at some special cases. Let the utility function be $x^{1-\gamma}/(1-\gamma)$ with $\gamma > 1$, and let p_2 be very close to zero. Then the solution is (approximately) given by

$$x_{11}^{-\gamma} \approx R x_{21}^{-\gamma} \tag{21}$$

$$x_{22}^{-\gamma} \approx R x_{12}^{-\gamma}. \tag{22}$$

From equations (11) and (12), it is straightforward to show that

$$x_{22} \approx x_{11}(1-t_1)R^{1/(\gamma-1)} \div [(t_2-t_1)R^{-1/\gamma} + (1-t_2)/R]. \tag{23}$$

It is possible to find parameter values with the property that $x_{22} < x_{11}$. For example, an economy with $R = 4$, $t_1 = 1/2$, and $t_2 = 3/4$ has this property. This possibility will be important when I turn to the issue of bank panics.

Before doing that, I need to verify that the solution is incentive compatible, or that Type 2 people do not have an incentive to visit the production location in period 1. I show that incentive compatibility is satisfied if p_2 is sufficiently small. Notice from (19) that in such a case $U'(x_{11}) \approx RU'(x_{21}) > U'(x_{21})$. Since $U'(\cdot)$ is decreasing, $x_{11} < x_{21}$. Then, with p_2 sufficiently small, by visiting the production location, Type 2 people would trade an almost sure return of x_{21} for a smaller return of x_{11}. Therefore, Type 2 people do not visit the production location, and the allocation is incentive compatible.

I turn now to the possibility of multiple equilibria given this allocation mechanism. To recapitulate, the mechanism specifies a payment of x_{11} to the first t_1 withdrawers, x_{12} to the next $t_2 - t_1$ withdrawers, and zero thereafter. If there are t_2 withdrawers, those who do not withdraw receive x_{22} in the final period. Suppose now that everybody believes that everybody else will attempt to withdraw. Suppose each person adopts the following strategy: Visit the production location, and if the current payment rate is x_{11}, withdraw; if it is x_{12} or zero, do not withdraw. By following this strategy, a person can be assured of at least a payment of x_{22} in the final period, and if the person is lucky, a payment of x_{11}.

I have demonstrated that there are economies for which $x_{11} > x_{22}$. Therefore, for such economies, multiple equilibria in the form of bank panics are possible.[7]

To summarize this section on the model with aggregate risk, I have

- Constructed an explicit model in which sequential service is an essential ingredient of the description of the technology.
- Solved for the optimal contract for such a model and shown that multiple equilibria are possible.
- Modeled sequential service so that credit markets are ruled out in the intermediate period.

The Model With Community Risk

Now let's replicate the aggregate-risk economy many times, so that we have a large number of *communities*, each of which is exactly like the aggregate-risk economy. Think of these communities as uniformly distributed over the interval [0,1]. Assume that a law of large numbers holds, so that p_1 and p_2 are now interpreted as the fraction of communities with t_1 and t_2 Type 1 people, respectively.

Within each community, the technology is the same as that described above. Also assume that the period 0 good is specific to each community and cannot be transported across communities. However, the period 1 and 2 goods can be shipped across communities. Assume, as well, that the amount invested in each technology is not observable.

The community information and isolation assumptions are also the same as the aggregate assumptions described earlier. In addition, assume that people cannot cross community boundaries. The realization of the number of Type 1 people is not observable across communities, and the consumption of any person is not observable to anyone else.

Within each community, assume that a local planner chooses allocation functions. These functions are unobservable to other local planners. What is observable is the net amount of period 1 and 2 goods shipped to a community. The amounts shipped are functions of the reported type of a community. Let $(M_1(t_1), M_1(t_2), M_2(t_1), M_2(t_2))$ be the intercommunity transfer functions, where $M_1(t_1)$ denotes the amount shipped (possibly negative) to a community in period 1 if it reports t_1 Type 1 people; $M_1(t_2)$, the amount shipped to a community in period 1 if it reports t_2 Type 1 people; and so on. Resource feasibility requires that

$$(24)\quad p_1M_1(t_1) + p_2M_1(t_2) = 0$$

$$(25)\quad p_1M_2(t_1) + p_2M_2(t_2) = 0.$$

In the economy with aggregate risk, we could safely assume that there was no investment in the liquid technology since it was dominated in return. For now, continue to assume that no community invests in the liquid technology; later I will argue that no community would choose to invest in this technology. Given this assumption, the resource constraints in each community are

$$(26)\quad (1-t_1)x_{21} = R[(1-t_1)x_{11} + M_1(t_1)] - M_2(t_1)$$

$$(27)\quad (1-t_2)x_{22} = R[(1-t_1)x_{11} - (t_2-t_1)x_{12} + M_1(t_2)] - M_2(t_2).$$

The left sides of (26) and (27) are period 2 payments. On the right sides, the terms in the brackets are the amounts left over after payments to Type 1 people and shipments from other communities.[8]

Given the intercommunity transfer functions M_1 and M_2, the problem faced by each local planner is to maximize expected utility of the community residents subject to the resource constraints (26) and (27) and the information and incentive constraints.

I now argue that this is necessarily true:

$$(28)\quad RM_1(t_1) - M_2(t_1) = RM_1(t_2) - M_2(t_2) = 0.$$

Equation (28) has the nice interpretation that the interest rate for intercommunity borrowing and lending must be R. Note that $M_1(t_1)$ is the amount loaned (if positive) in period 1 and $M_2(t_1)$ is the amount repaid by borrowers in period 2. The ratio is R.

To see this result, suppose that $RM_1(t_1) - M_2(t_1) < 0$. Then, from (24) and (25), $RM_1(t_2) - M_2(t_2) > 0$. Now consider the following strategy for a community. Regardless of the true state, report that the state is t_2. Since $RM_1(t_1) - M_2(t_1) < RM_1(t_2) - M_2(t_2)$, we can introduce some slack into a resource constraint (26) and increase payments to Type 2 people. Therefore, such an allocation is not incentive compatible. A similar argument

---

[7]Of course, I have not explicitly modeled sunspots or other features which trigger the bad equilibrium outcome. For further discussion of this issue, see Postlewaite and Vives 1987. In particular, if the probability of bank panics is high, then people might choose not to set up any collective arrangement.

[8]Strictly speaking, this way of writing the resource constraint is consistent with the goods-in-process aspect of the technology only under the assumption that goods invested in period 0 continue to grow at rate R into period 2 no matter where they are shipped, as long as they are not consumed.

establishes that if $RM_1(t_1) - M_2(t_1) > 0$, then the allocation is not incentive compatible. We have the desired result.

Clearly, intercommunity borrowing and lending is a technology with exactly the same properties as the illiquid technology. The problem of the local planners is exactly the same as the problem in the economy with aggregate risk. It should also be clear that the liquid technology will not be used in any essential way. Since investments in this technology are unobservable and since the intercommunity interest rate cannot be different from R, no community invests in this technology.

Now let's make one change in the community-risk economy. Assume that the amount invested in the liquid technology by each community is observable. Retain all the other assumptions. I will argue that there is a mechanism which implements the *full-information allocations.* That is, it will produce the same allocations as the best mechanism in an economy where the types of all people are observable and people are not isolated.

To show this result, I first describe the full-information allocations. Since there is no aggregate risk in this economy, the fraction of Type 1 people equals $p_1t_1 + p_2t_2$, which I denote by t. The problem is to maximize expected utility subject to the resource constraints. Clearly, Type 1 people will not consume in period 2, and it is easy to show that Type 2 people will not consume in period 1. Therefore, the problem is to

(29) $\max\ tU(x_1) + (1-t)U(x_2)$

subject to

(30) $tx_1 + (1-t)(x_2/R) = 1.$

This is exactly the problem Diamond and Dybvig (1983) studied. Because the relative risk aversion of the utility function is greater than unity, the optimal allocation satisfies

(31) $x_1 < x_2$

(32) $x_2/x_1 < R.$

Now consider the following mechanism. Each community invests an amount tx_1 (per capita) in the liquid technology. All communities can borrow and lend in the intermediate period at an interest rate x_2/x_1. Each community promises to pay x_1 units to those who withdraw in period 1 and x_2 units to those remaining in period 2. I claim that these allocations are incentive compatible.

Consider the problem faced by a community which has promised to pay x_1 and x_2 and is faced with t_2 Type 1 people. It pays off the first t Type 1 people from the liquid technology. It borrows $(t_2 - t)x_1$ in the market to pay off the remaining Type 1 people. In period 2, the illiquid technology yields $R(1-tx_1)$ units of output, out of which it pays x_2 units to each of its $(1-t_2)$ Type 2 people and $(t_2-t)x_2$ to its lenders. Since, from (30), $R(1-tx_1) = (1-t)x_2$, this transaction is feasible. This argument also shows that, with any other borrowing (or lending) strategy, the community will be unable to meet its commitments to its people.

To complete the argument, we should ask whether any local planner would benefit by choosing different allocations than the pair (x_1, x_2). But recall that these allocations yield the highest utility in the class of all mechanisms. Therefore, no planner would choose anything else. An important point, I repeat, is the observability of investment in the liquid technology. If this investment were not observable, every community would be better off by not investing in the liquid technology and then borrowing, if needed, at the interest rate x_2/x_1 in the marketplace. Lastly, I address the multiple equilibrium problem. It should be clear that a policy of suspension of convertibility if all communities wish to borrow eliminates panics since, from (31), $x_2 > x_1$.

Summary and Policy Implications

In this paper I have constructed a series of models of bank panics. The substantive contribution is to show that such a model can be constructed with rational, optimizing people. The technical contribution is to construct a model in which sequential service is a central property of an efficient mechanism rather than an unmodeled imposition upon otherwise efficient arrangements. The resulting model was constructed to capture two key features of the U.S. National Banking System: the limits on bank branching and the ability of banks to circumvent reserve requirements. These features meant that the interbank borrowing and lending market did not function very well. The result was that banks could not diversify away withdrawal risk. Therefore, banks were susceptible to runs and the banking system, to panics.

A central feature of the National Banking System is that it inhibited portfolio diversification, on both the asset and the liability sides of bank balance sheets. Jagannathan and I (Chari and Jagannathan 1988) and

Williamson (in this issue) have argued that what matters for bank runs is the failure to diversify risk on the asset side. I have argued here that what matters is, rather, the failure to diversify withdrawal risk. What made me change my mind was the fact that the rate of bank failure has been remarkably low since the onset of deposit insurance (Williamson, this issue). If risk on the asset side were so important, that insurance would have made little difference to the rate of bank failure.

In this paper, I have shown that efficient outcomes can be sustained by an appropriate choice of reserve requirements and by a well-functioning interbank borrowing and lending market. I will argue further here that an effective central bank is also likely to be necessary. For one thing, conditions and the appropriate level of reserves change over time. (In the model, think of the aggregate number of Type 1 people changing.) If reserve requirements were set legislatively, changing them would take quite a bit of time. Therefore, delegating the task to a central bank might be preferable.

Furthermore, the model suggests that the central bank's discount window policy plays a key role in promoting efficient banking arrangements and preventing bank panics. To understand the role played by the discount window in the model, suppose that all banks maintain their reserves at the central bank. In the intermediate period, the central bank promises to lend at a gross interest rate of x_2/x_1 against assets which mature in the final period. Note that this interest rate is higher than the gross rate of return on reserves (which is unity), but lower than the rate of return on the secured assets (which is R). Therefore, the central bank seems to be subsidizing banks by allowing borrowing at the discount window at below-market interest rates. Here, appearances do not deceive. The central bank provides an important insurance role by allowing such below-market borrowing. The premium for this insurance contract is paid in the form of reserves held at the central bank. I have argued that such a mechanism is efficient and can eliminate bank panics. In addition, central banks today can create reserves by printing money. Therefore, a promise by the central bank to lend unlimited amounts against sound assets is credible. In this sense, central banks can prevent panics arising from a lack of confidence in the banking system.

The advantages of using these central bank policies instead of deposit insurance should be obvious. Deposit insurance creates moral hazard problems that can be mitigated only by regulations which effectively reduce bank risk-taking. The appropriate amount of risk for banks is hard to measure and harder to enforce. Lending by the central bank of the kind discussed here is not subject to the same sorts of moral hazard problems, for two reasons. One is that, to the extent such lending is done against sound assets, it does not subsidize risk-taking. The other reason is that, since depositors are not insured, they have an incentive to monitor banks to ensure that they don't take excessive risk. Monitoring by depositors rather than by regulators is desirable because it generally allows for greater diversity in risk among banks. Depositors could choose among banks which offer different menus of risk versus return. Such a choice is effectively denied by the current system.

I have thus far not mentioned the well-known U.S. bank panics between 1929 and 1933. These panics, I think, are best explained by Friedman and Schwartz (1963), who blame the Federal Reserve System for not suspending convertibility until it was too late. Remember that in my model, as in Diamond and Dybvig's (1983), suspension is an essential component of the efficient arrangement; if there are legal constraints against suspension, panics can occur.

Some may argue that the considerations I have advanced are purely theoretical and that policy reforms such as eliminating deposit insurance would return us to the National Banking System. But were conditions under that system so terrible? Recall, for example, that, during the panics of that period, the currency premium was roughly 2 percent. Today, a 2 percent loss of wealth in the U.S. stock market is a 50 point decline in the Dow Jones industrial average. While this kind of decline does not occur everyday, when it does, we certainly do not call it a *panic* or rush to reform the stock market. To put the matter somewhat differently, three times in 50 years the National Banking System imposed on depositors a loss of about 2 percent. Did that system perform so much worse than the current system does?

Appendix
On the Optimality of Sequential Service Mechanisms

Three questions naturally arise from the models in the preceding paper:

- What are the allocations when both types of people visit the production location but payments depend only on the number of Type 1 predecessors?
- Why do the payments depend only on the number of Type 1 predecessors rather than, say, the relative numbers of predecessors of each type?
- Why must people consume as soon as they arrive at the production location? Why not have everybody wait to consume until the total number of Type 1 people is known?

I answer these questions here.

What Are the Allocations When Both Types Visit?

Consider the allocations when both types of people visit the production location. As in the paper, here payments can depend only on the person's reported type and the number of predecessors reporting they are Type 1. I now develop the resource constraints when the state is t_1. Recall that people arrive at a uniform rate of unity at the production location. At some instant r in calendar time, the fraction of Type 1 people is t_1 and the fraction of Type 2 people is $1-t_1$. The number of Type 1 predecessors is rt_1. Therefore, total payments at instant r are given by $t_1x_1(1, rt_1) + (1-t_1)x_1(2, rt_1)$. Therefore, payments to all people in period 1 are given by $\int_0^1[t_1x_1(1, rt_1) + (1-t_1)x_1(2, rt_1)]\, dr$.

A change in the variables, so that $s = rt_1$, makes this the resource constraint in state t_1:

$$\text{(A1)}\quad \int_0^{t_1}[t_1x_1(1,s) + (1-t_1)x_1(2,s)]\, ds/t_1 + R^{-1}\{\int_0^{t_1}[t_1x_2(1,s,t_1) + (1-t_1)x_2(2,s,t_1)]\, ds/t_1\} = 1.$$

Note that s denotes the number of Type 1 predecessors. Similarly, the resource constraint in state t_2 becomes this:

$$\text{(A2)}\quad \int_0^{t_2}[t_2x_1(1,s) + (1-t_2)x_1(2,s)]\, ds/t_2 + R^{-1}\{\int_0^{t_2}[t_2x_2(1,s,t_2) + (1-t_2)x_2(2,s,t_2)]\, ds/t_2\} = 1.$$

The expected utility of a typical person in the planning period is

$$\text{(A3)}\quad U(x_1, x_2) = \sum p_i\{t_i\int_0^{t_i}[x_1(1,s) + \delta x_2(1,s,t_i)]\, ds/t_i + (1-t_i)\int_0^{t_i} U(x_1(2,s) + x_2(2,s,t_i))\, ds/t_i\}.$$

I turn now to the incentive-compatibility constraints. If people know the time at which they themselves arrive, then they also know that $t=t_2$ if $s>t_1$. However, if $s \in [0, t_1]$, then people perceive the probabilities that $t=t_1$ and $t=t_2$ as p_1 and p_2, respectively. Therefore, we have the following incentive constraints: If $s \in [t_1, t_2]$, then

$$\text{(A4)}\quad U(x_1(1,s) + \delta x_2(1,s,t_2)) \geq U(x_1(2,s) + \delta x_2(2,s,t_2))$$

$$\text{(A5)}\quad U(x_1(2,s) + x_2(2,s,t_2)) \geq U(x_1(1,s) + x_2(1,s,t_2)).$$

And if $s \in [0, t_1]$, then

$$\text{(A6)}\quad \sum p_i[U(x_1(1,s) + \delta x_2(1,s,t_i))] \geq \sum p_i[U(x_1(2,s) + \delta x_2(2,s,t_i))]$$

$$\text{(A7)}\quad \sum p_i[U(x_1(2,s) + x_2(2,s,t_i))] \geq \sum p_i[U(x_1(1,s) + x_2(1,s,t_i))].$$

An allocation (x_1, x_2) is *efficient* if it maximizes (A3) subject to (A1), (A2), and (A4)–(A7). It is straightforward to show that, if p_2 is sufficiently small, then the allocations described in the paper are efficient. Therefore, requiring that all people visit the production location does not alter the set of efficient allocations. These allocations obviously dominate the allocations if only Type 2 people are required to visit the production location and coincide with the efficient allocations if only Type 1 people are required to visit the production location.

The key assumption here is that allocations depend only on the number of Type 1 predecessors. If requiring all people to visit is costless and the allocations are incentive compatible, then the relative proportions of the two types can be uncovered within a very short time at the beginning of the period. Since the state of the economy is then known, the sequential service constraint need not be respected. Therefore, I modify the economy in a small way and show that efficient allocations in the modified economy require only Type 1 people to visit.

Why Does Only Type 1 Visit?

Suppose that traveling to the production location has a small cost, ϵ. Let (x_1^s, x_2^s) denote the allocations which respect sequential service. Then expected utility under the efficient sequential service mechanism is given by

$$\text{(A8)}\quad U^s = p_1[t_1U(x^s_{11}) + (1-t_1)U(x^s_{21})] - p_1t_1\epsilon + p_2[t_1U(x^s_{11}) + (t_2-t_1)U(x^s_{12}) + (1-t_2)U(x^s_{22})] - p_2t_2\epsilon.$$

Consider now the set of allocations that are realized if the state of nature becomes known in an infinitesimal time at the beginning of period 1. Denote these allocations by $x_1(t_1)$, $x_1(t_2)$, $x_2(t_1)$, $x_2(t_2)$. Expected utility under this mechanism is given by

$$\text{(A9)}\quad U^k = p_1[t_1U(x_1(t_1)) + (1-t_1)U(x_2(t_1))] - p_1\epsilon + p_2[t_2U(x_1(t_2)) + (1-t_2)U(x_2(t_2))] - p_2\epsilon.$$

Note that under this mechanism, everybody visits the production location. Given that we are here restricted to nonrandomized visit decisions, the only way to uncover the aggregate state is to require that everybody visit.

Suppose now that p_2 equals zero. Clearly, $U^s > U^k$. Therefore, by continuity, for p_2 sufficiently close to zero, $U^s > U^k$.

Why Not Wait to Consume?

The last issue addressed here is the requirement that consumption occur immediately on arrival at the production location. I address this issue by modifying preferences so that efficiency requires immediate consumption. Suppose that the preferences of Type 1 people are now given by $U(\int\delta(r)c_1(r) + \delta c_2)$, where $c_1(r)$ denotes consumption at instant r in period 1, c_2 denotes consumption in period 2, and $\delta(r)$ is a decreasing function. If $\delta(r)$ decreases rapidly enough, consuming as early as possible will be efficient. Since consumption cannot occur before arrival at the production location, consuming as early as possible once a person arrives there is efficient.

References

Benston, George J.; Eisenbeis, Robert A.; Horvitz, Paul M.; Kane, Edward J.; and Kaufman, George G. 1986. *Perspectives on safe & sound banking: Past, present, and future.* Cambridge, Mass.: MIT Press.

Bhattacharya, Sudipto, and Gale, Douglas. 1987. Preference shocks, liquidity, and central bank policy. In *New approaches to monetary economics,* ed. William A. Barnett and Kenneth J. Singleton, pp. 69–88. Proceedings of the Second International Symposium in Economic Theory and Econometrics. New York: Cambridge University Press.

Boyd, John H., and Rolnick, Arthur J. 1989. A case for reforming federal deposit insurance. In *Federal Reserve Bank of Minneapolis 1988 annual report,* pp. 3–16. Minneapolis: Federal Reserve Bank of Minneapolis.

Chari, V. V., and Jagannathan, Ravi. 1988. Banking panics, information, and rational expectations equilibrium. *Journal of Finance* 43 (July): 749–61.

Diamond, Douglas W., and Dybvig, Philip H. 1983. Bank runs, deposit insurance, and liquidity. *Journal of Political Economy* 91 (June): 401–19.

Friedman, Milton. 1960. *A program for monetary stability.* New York: Fordham University Press.

Friedman, Milton, and Schwartz, Anna J. 1963. *A monetary history of the United States, 1867–1960.* Princeton, N.J.: Princeton University Press (for NBER).

Gorton, Gary. 1985. Clearinghouses and the origin of central banking in the United States. *Journal of Economic History* 45 (June): 277–83.

__________. 1988. Banking panics and business cycles. *Oxford Economic Papers* 40 (December): 751–81.

Gorton, Gary, and Mullineaux, Donald J. 1987. The joint production of confidence: Endogenous regulation and nineteenth century commercial-bank clearinghouses. *Journal of Money, Credit, and Banking* 19 (November): 457–68.

Jacklin, Charles J. 1987. Demand deposits, trading restrictions, and risk sharing. In *Contractual arrangements for intertemporal trade,* ed. Edward C. Prescott and Neil Wallace, pp. 26–47. Minnesota Studies in Macroeconomics, Vol. 1. Minneapolis: University of Minnesota Press.

Kareken, John H. 1981. Deregulating commercial banks: The watchword should be caution. *Federal Reserve Bank of Minneapolis Quarterly Review* 5 (Spring-Summer): 1–5.

__________. 1983. Deposit insurance reform or deregulation is the cart, not the horse. *Federal Reserve Bank of Minneapolis Quarterly Review* 7 (Spring): 1–9.

Kareken, John H., and Wallace, Neil. 1978. Deposit insurance and bank regulation: A partial-equilibrium exposition. *Journal of Business* 51 (July): 413–38.

Kane, Edward J. 1985. *The gathering crisis in federal deposit insurance.* Cambridge, Mass.: MIT Press.

Postlewaite, Andrew, and Vives, Xavier. 1987. Bank runs as an equilibrium phenomenon. *Journal of Political Economy* 95 (June): 485–91.

Romer, Christina D. 1989. The prewar business cycle reconsidered: New estimates of gross national product, 1869–1908. *Journal of Political Economy* 97 (February): 1–37.

Smith, Bruce D. 1986. Bank panics, suspensions, and geography: Some notes on the "contagion of fear" in banking. Manuscript. University of Western Ontario.

Sprague, O. M. W. 1910. *History of crises under the National Banking System.* Report by the National Monetary Commission to the U.S. Senate. 61st Cong., 2nd sess., S. Doc. 538. Washington, D.C.: Government Printing Office.

Wallace, Neil. 1988. Another attempt to explain an illiquid banking system: The Diamond and Dybvig model with sequential service taken seriously. *Federal Reserve Bank of Minneapolis Quarterly Review* 12 (Fall): 3–16.

White, Eugene Nelson. 1983. *The regulation and reform of the American banking system, 1900–1929.* Princeton, N.J.: Princeton University Press.