BEBR

FACULTY WORKING PAPER NO. 89-1541

College of Commerce and Business Administration

University of Illinois at Urbana-Champaign

March 1989

The Role of Demandable Debt in Structuring Optimal Banking Arrangements

Charles W. Calomiris
Northwestern University

Charles M. Kahn, Associate Professor
Department of Economics

THE ROLE OF DEMANDABLE DEBT IN STRUCTURING OPTIMAL BANKING ARRANGEMENTS

by

Charles W. Calomiris

Department of Economics
Northwestern University

and

Charles M. Kahn

Department of Economics
University of Illinois, Urbana-Champaign

Demandable-debt finance of banks warrants explanation because it entails costs of bank suspension, liquidation, and idle reserve holdings. An explanation is developed in which demandable debt provides incentive-compatible intermediation where the banker has comparative advantage in allocating investment funds, but may act against the interests of uninformed depositors. Demandable debt attracts funds by giving depositors an option to force liquidation. Its usefulness in transacting follows from information sharing between monitors and non-monitors.

Revised, February 1989

We thank Lee Alston, Herbert Baer, Kyle Bagwell, Ben Bernanke, Sudipto Bhattacharya, Doug Diamond, Gary Gorton, Monica Hargrave, Charlie Jacklin, two anonymous referees, and participants in the joint Northwestern-University of Chicago theory seminar, and seminars at the University of Illinois, the Federal Reserve Bank of Chicago, Purdue University, SUNY Stony Brook, the National Bureau of Economic Research, and the Garn Institute for helpful comments. This work was partially funded by the National Science Foundation under grant SES-8511137.

THE ROLE OF DEMANDABLE DEBT IN STRUCTURING OPTIMAL BANKING ARRANGEMENTS

For centuries the vast majority of externally financed investments have been funded by banks, for which demandable debt instruments (bank notes and checking accounts) have been the principal source of funds.[1] The goal of this paper is to explain the emergence of demandable-debt banking historically as the primary means of external finance in the economy.

Demandable debt warrants explanation because in several respects it appears more costly than available alternative contracting structures. By issuing demandable debt, banks created a mismatch between the maturity of assets and liabilities. This mismatch left them exposed to the possibility that depositors would attempt to withdraw more funds than a bank could supply on short notice. When this occurred, the consequences were costly. Individual banks which did not meet their obligations were forced into expensive procedures (liquidation or receivership) that would not have arisen in an equity-based or maturity-matched contracting structure.[2] If depositors en masse attempted to withdraw funds from the entire banking system, banks as a group were forced to suspend convertibility of their liabilities into specie on demand. Such suspension was also disruptive and costly.[3] To defend against either of these undesirable consequences, banks had to hold a proportion of their assets in idle reserves to insulate themselves from excessive withdrawals.

Given these costs, demandable debt seems inferior to both maturity-matched debt and equity contracting. However, in this paper we show that demandable debt has an important advantage as part of an incentive scheme for disciplining the banker. In effect, demandable debt permits depositors to "vote with their feet"; withdrawal of funds is a vote of no-confidence in the activities of the banker. Without the ability to make early withdrawals depositors would have little incentive to monitor the bank. The ease with which banks may be forced into liquidation, far from being an unfortunate consequence of the contracting structure, turns out to be central to the structure: we show that by submitting to the threat of liquidation under appropriate circumstances the banker can reduce his cost of capital.

I. EXPLANATIONS FOR DEMANDABLE DEBT

Recent theoretical work on the role of banks has tended to divide into two categories. One category emphasizes the role of banks as providing flexibility for depositors in the timing of consumption. The other category, to which our paper belongs, emphasizes the incentive problem inherent in the divergence of interest between a bank's depositors and its managers.[4] For reasons indicated below, we believe that accounts which ignore the incentive problem facing the banker do not adequately explain why banks historically settled on demandable debt.

Liquidity Explanations for Demandable Debt

In the past several years, the preeminent theoretical analyses of banks, bank runs, and bank regulation have assumed that the economic role of demandable debt is to provide liquidity to risk averse depositors who are uncertain about the timing of their future consumption demand.[5]

In this category of models, bank runs, when they occur, are an unfortunate and undesirable side-effect of a contract whose whole purpose is to provide liquidity. In consequence, these papers and successive papers in the area have gone on to examine the role of government regulation in eliminating bank runs, and the possibility that other private arrangements might exist which would provide liquidity without the undesirable side effect.

These models provide a concise formalization of the fact that banks provide transactions flexibility and a coherent account of bank runs. But because they ignore the incentive problem facing the banker, they are unable to account for several important institutional features of demandable debt.

First of all, the major consequence of a run on an individual bank was to place the bank into receivership. If the major cost of a bank failure was the inability of demanders to receive timely payments, then it is paradoxical that liquidations were associated with long delays in access to bank assets by creditors.[6] Historically, bank loans were of short duration (typically, a matter of months). If rapid access to assets had been the goal, a mechanism that simply prevented banks from making new loans until all depositors had been paid off would have been superior to putting banks into receivership.

Second, in the absence of incentive constraints on the part of the banker, the optimal arrangement in liquidity-based accounts always involves suspension of convertibility rather than expensive liquidation. But suspension was not an option available to individual banks; it was only an alternative for the financial-system as a whole, in the face of system-wide panics. Individual banks that could not satisfy creditors' fears about solvency were not permitted to suspend; they were forced to close.[7]

Third, studies of actual bank failures give fraud a prominent place in the list of causes. For example, Smead finds, in his study of bank failures in the 1920s, that three of the nine most common causes of bank failure involve fraudulent or questionable activities by the banker: loans to officers and directors, outright defalcation, and loans to enterprises in which officers and directors were interested. Studies of nineteenth century banking indicate that fraud and conflicts of interest characterize the vast majority of bank failures for state and nationally chartered banks.[8]

Finally, receivership resulted from a critical mass of depositor withdrawal orders, and was invoked because of information about bank asset values, not because of exogenous liquidity needs of individual depositors. In cases of massive exogenous demand for an individual bank's assets by small depositors, banks avoided failure by appealing to other banks for loans of reserves; however when large informed depositors (including other bankers) concluded that a bank was in trouble, they would precipitate a run, depleting the bank's reserves and forcing it to be placed in receivership.[9]

These considerations make it apparent that the liquidation of banks -- which was part and parcel of demandable-debt contracts -- was designed to place the assets of banks beyond the reach of the banker. The rationale for prohibiting banks from suspending at their own discretion may be the discipline that it imposed on the behavior of the banker. Thus a model of demandable debt with bank liquidation through receivership should account for the desirability of taking control of the bank away from the banker at the option of depositors.

The "sequential service constraint" (first-come first-serve rule) for bank withdrawals, which allowed informed depositors to receive repayment before banks were placed into receivership also warrants explanation. In cases other than banking, payments of bankrupt firms to creditors in anticipation of bankruptcy are not allowed, and creditors may be forced to relinquish such payments during the bankruptcy proceeding. Why in the case of banking should those who run the bank first receive preferential treatment in liquidation states?

Demandable Debt as an Incentive Scheme

The second category of models of the role of banks begins with the assumption that bankers have an informational advantage in determining which projects are most worthy of financing. Therefore the banker has a comparative advantage in allocating funds for investment, but he also may have the ability to act against the interests of uninformed depositors.[10]

We show that demandable debt can provide an incentive-compatible solution to this problem in the presence of costly information. The right to take one's money out of the bank if one becomes suspicious that realized returns are low makes it in the depositor's interest to keep an eye on the bank. If enough depositors agree with this negative assessment of the bank's future, liquidation will be called for and the bank will close. The demandable debt contract allows the banker to pre-commit to a set of payoffs he otherwise would not be able to offer depositors.[11]

Not all depositors need monitor the banker. We argue that the first-come first-serve ("sequential service") rule of demandable debt provides compensation for those who choose to invest in information and thus avoids free-riding. We view bank intermediation, therefore, as a three-sided relationship. The monitors pay the costs of vigilance, but receive the benefit of knowing that they will be "first in line" (and thereby receive a higher payment than other depositors) should it become necessary to withdraw their funds from the bank. The depositors who do not monitor are willing to pay the price of being last in line in "bad" states, because they receive a benefit in return: the active monitors keep the banker in line and thereby provide a benefit to the passive depositors. Depositors need not reveal whether they are active or passive -- the same contract works for both types.[12]

The physical structure we assume includes the following important features: (1) The bank is operated by a monopolist with special access to a profitable investment opportunity which yields either a good or a bad realization. (2) There is potential for cheating by the banker which takes the form of his absconding with a proportion of the bank's assets after the investment realization. (One can think of this more generally as costly ex post fraudulent behavior which the banker undertakes whenever it is more profitable to do so than to make the promised payments to depositors). (3) Depositors face different costs of obtaining a signal which allows them to predict profitability. (4) An authority exists who will enforce contracts (some of which may stipulate conditions for bank liquidation) and who can act as receiver for liquidated banks. (5) Depositors have a reservation level of return on their endowments below which they will not invest funds with the banker.

The profit maximizing banker will act to maximize social gain by selecting a contract that achieves beneficial intermediation (investment in profitable enterprises), while avoiding as much as possible the costs associated with absconding or liquidating. We find that the demandable debt contract is optimal for a range of parameter values. The potential for costly liquidation may be more than offset by the social gain which comes from enhanced investment opportunities.

It is useful to place these results in the context of two related papers. Chari and Jagannathan (1988) provide an example of an information-based run[13] for a model which has many features in common with ours. In their account, banks provide demandable debt in a risk neutral environment. Individuals obtain private information regarding the solvency of the bank. Runs are based on this information.[14]

One important difference between their formulation and ours is the fact that they assume an (exogenously imposed) negative externality from liquidation of the bank's assets. Therefore when individual depositors choose to run the bank, non-runners suffer in the process. Since there are no incentive problems on the part of the banker, suspension of convertibility always improves the outcome. Thus the creation of a liquidation technology is not efficient in their model.

Chari and Jagannathan recognize that their model is designed to investigate bank runs, not to justify the existence of banks. In our model, there is a <u>positive</u> externality from running the bank: when the depositor observes a bad signal, he calls for liquidation, thereby salvaging some of the bank's value. The bank's structure is designed to internalize this positive externality, and allow non-monitoring depositors to compensate monitors for the benefits they provide. Because of this monitoring function, runs on the bank in our model provide social benefits, rather than social costs.[15]

Bernanke and Gertler (1987) provide a simple macroeconomic model in which bankers are subject to moral hazard and depositors desire liquidity. Because of the moral hazard problem, bank owners must invest their own assets in the bank; there is a limit to the amount of deposits that they can obtain. Because of the lack of information about the profitability of projects, only debt contracts are incentive compatible; because depositors desire liquidity, demandable debt is best. Bernanke and Gertler explicitly assume that costly monitoring and punishment of defaulting bankers are not possible; therefore in their model there is no room for shutting down banks that fail to pay depositors on demand. In effect this means that the only possible debt contracts are those in which payments are low enough that default occurs with probability zero: the bank's equity must be sufficient to make this guarantee possible.

Our model demonstrates the gains to be had from allowing the demandable debt contracts to be less than certain, and use of costly monitoring to reduce the danger of malfeasance on the part of the banker. In this case, demandable debt becomes desirable even in the absence of liquidity demand. It also explains the usefulness of the sequential service constraint in banking, giving preferential treatment to the depositors who are first to request redemption.

The remainder of the paper is organized as follows: Section II demonstrates the value of a demandable debt contract in the case of a single investor contracting with the banker monopolist. Here a run corresponds to a demand by the investor for liquidation of the bank. Section III examines the case where different monitors receive different (i.i.d.) signals. In this case it pays to have more than one depositor monitoring the bank, because the quality of signals in the aggregate improves with the number of monitors. Banks find it advantageous to hold reserves, as well as to make loans, because the reserves provide a buffer that allows banks to reduce the likelihood of unwarranted liquidation. An optimal threshold of withdrawal orders is chosen at which the bank is liquidated, and relative payoffs ensure that the optimal number of monitors invest in receiving signals.

The final section offers some perspectives on our results. We consider several limitations inherent in our approach. In addition we briefly and informally indicate how solving the incentive problem facing the banker will also make the banker's liabilities more transactable. Formal models combining the incentive problem and the issue of transactability are an important field for further research.[16]

II. THE MODEL WITH A SINGLE DEPOSITOR

Physical Structure

A banker has an opportunity available to him as an investment. However he lacks sufficient capital to take advantage of the opportunity. The investment opportunity costs one dollar. Each potential depositor has one dollar to invest. We will let S represent the total expected return available for a dollar's investment elsewhere in the economy. We assume all agents are risk neutral; thus any scheme the banker develops will have to yield a depositor that same expected return.

The investment opportunity yields an uncertain payoff which may take one of two values T_1 or T_2, with $T_2 > T_1$. The probability of the high outcome is γ. The realization is unknown by all parties at the outset, and is observable ex post only by the banker. Thus there is no way to make a contract tied directly to the value of T_i. (This analysis extends naturally to the case where T has a continuous distribution).

Let period 3 be the date at which the payoff is realized and the loan is to be repaid. We assume that in period 3, immediately before repayment, the banker has the opportunity to abscond with the funds. Absconding is socially wasteful -- for concreteness we will assume it reduces the realization T_i by the proportion A, where A is between zero and one.

Although the act of absconding reduces the size of the pie that is divided between the banker and the depositor, it places the banker beyond the reach of the law. Therefore he is no longer constrained to repay the loan as initially promised. Thus any promise to pay the depositor an amount P is actually an option of the banker either to pay P, or to leave town with his assets diminished by the proportion A.

The losses from absconding may be interpreted in a variety of ways. They may represent the cost of engaging in fraud (payments to co-conspirators) or the costs (foregone earnings) of placing the bank's resources in a form that allows theft. The latter interpretation requires a richer, multiperiod model than the one we provide, in which bankers allocation decisions depend on last period earnings.

It should be readily apparent that the temptation to abscond will be greater with lower realizations of T_i. In deciding whether to abscond, the banker compares the "tax" on absconding AT_i with the promised funds due the depositor. If the absconding tax is less, then absconding is more profitable than paying up. Historical evidence confirms the greater prevalence of fraud in times of low returns to bank investments.[17]

Because of the threat that the banker has of absconding -- a threat against which he cannot commit himself -- it will generally be necessary for the banker to increase the payment offered to a depositor by a "default premium" as protection against those states in which the depositor will in fact receive nothing.

Note that the addition of a default premium can in turn increase the probability of default, by making it desirable for the banker to abscond in good states as well. For example, suppose

$$S > AT_1$$

so that any payment promised to the depositor must be sufficiently large so as to incur absconding in the low realization -- that is, a promise to pay P will be honored no more than γ of the time. Suppose also that

$$\gamma T_2 + (1-\gamma)(1-A)T_1 > S$$

so that the investment would be socially desirable (even taking into account the loss from absconding in the low realization). Then if

$$S > \gamma AT_2$$

there is <u>no</u> way to promise the depositor enough expected payment to make him willing to invest, despite the social desirability of the project; the promised payment would have to exceed AT_2, making it desirable for the banker to abscond all the time.

Because of the loss of socially desirable opportunities, it is useful to have a method of thwarting absconding. One such method is the liquidation of the bank in period 2. Liquidation means that the bank's assets are taken over by a receiver, controlled by a court. This is an expensive process, not the least because the court-appointed and -controlled receiver is likely to be less able to realize the full potential of the assets. On the other hand, the fact that the assets are no longer in the banker's control preempts any decision by him to abscond with the funds.

We assume that liquidation reduces the value of the assets by the proportion L, so that L can be regarded as the tax due to liquidation. For a complete characterization of the process of liquidation, it is necessary to take some stand as to the maximum that can be feasibly paid to the depositor in the case of liquidation. We call this value M, and we assume that[18]

$$AT_2 > M > AT_1 \tag{1}$$

so that the amount that can be guaranteed to the depositor in a liquidating contract is greater than the maximum amount that can be guaranteed in a non-liquidating contract. We also assume that

$$L < A \tag{2}$$

so that liquidation is less wasteful socially than is absconding.[19]

In some cases it may be desirable <u>always</u> to put the assets of the bank into liquidation rather than risk the banker's absconding. We call such an agreement a "simple liquidation contract", as opposed to a "simple non-liquidation contract" which states a promised repayment and leaves it to the banker to abscond or not.

The more interesting case however, is one in which the depositor, based on his own information, is given the option of demanding liquidation or not. Specifically, suppose that by paying a cost I the depositor is able to receive a signal σ in period 1 as to the likelihood of a high (T_2) or low (T_1) realization. The action of investing in the signal, and the result are private. The signal σ works as follows: It takes on one of two values {g, b} (for "good" and "bad").[20] The probability of a high realization, contingent on the signal is ρ_σ;

$$\rho_g > \gamma > \rho_b \tag{3}$$

We will use the indicator variable $e \in \{0,1\}$ to represent the depositor's choice: $e = 1$ if there was an investment in the signal, 0 otherwise.

In summary, the physical structure of our model is as follows: There are three periods. In period 1 the depositor may invest in receiving a signal. In period 2 the bank may be liquidated. In period 3, the loan is repaid to the depositor, unless the banker decides to abscond (which he can only do if the bank has not been liquidated.)

<u>The Contracting Structure</u>

Contracts are arranged in period zero. The monopolist banker offers the profit maximizing contract among those which yield the depositor at least S in expected returns. (If no such contract exists, or the best such contract yields negative profits, then none is offered).

The universe of contracts in this structure is as follows: A contract is a function from a space of announcements Σ into <u>outcomes</u>. An outcome is a pair (P,Λ), where $\Lambda \in \{0,1\}$ is an indicator variable equaling 1 if liquidation is mandated and 0 otherwise. P is the mandated repayment. (Of course P will only be received if the banker does not abscond).[21]

If the contract only specifies one outcome, we'll call it a "simple contract." Otherwise we call it a "compound contract." We have already described the two kinds of simple contracts -- the simple liquidating contract and the simple non-liquidating contract. A straightforward application of the revelation principle demonstrates that for the single depositor case, contracts need never contain more than two outcomes because the signal the depositor may observe has only two values. We can identify the announcements in a compound contract with assertions by the depositor that he has observed one or the other signal. Thus a compound contract consists of a quartet $(P_b, \Lambda_b, P_g, \Lambda_g)$.

Each contract generates a sequential game in which the depositor chooses the level of investment in information gathering e and the announcement he makes as a function of the signal he receives. The banker chooses whether to abscond as a function of the announcement made by the depositor and the realization on the investment. An <u>optimal</u> contract is one for which there is a sequential equilibrium which generates maximum profits consistent with the depositor's receiving expected returns equal to the amount S.

<u>Theorem 1</u>: The optimal contract in the problem takes one of the following four forms:

a. A simple non-liquidating contract.

b. A simple liquidating contract. In this case,

$$AT_1 < P \leq M. \tag{4.b}$$

c. A compound contract composed of two simple non-liquidating contracts. ($\Lambda_b = \Lambda_g = 0$). In this case

$$P_b \leq AT_1 \text{ and } AT_1 < P_g \leq AT_2. \tag{4.c}$$

d. A compound contract composed of one simple liquidating contract and one simple non-liquidating contract. ($\Lambda_b = 1$; $\Lambda_g = 0$). In this case

$$AT_1 < P_b < P_g \leq AT_2 \quad . \tag{4.d}$$

If the optimal contract is a compound contract, then the depositor invests in the signal; if it is a simple contract he does not. In the case of compound contracts, absconding occurs if and only if the signal was g but the low value outcome T_1 was realized.

We'll call contract d "demandable debt." It works as follows: After making the deposit the depositor invests in learning what the likely outcome will be. If he receives the bad signal, he opts for the simple liquidating contract. This delivers a payment with certainty. If he receives the good signal he opts for the non-liquidating contract. This promises a higher payment but runs the risk of the banker's absconding.

Contract c works in virtually the same way. The only difference is that the guaranteed payment in the case of a bad signal is sufficiently low that the banker will never wish to abscond and so it is not necessary to use liquidation to hold him in place. Since liquidation always involves social costs, it is not difficult to demonstrate that in any case where contract c is feasible, it dominates contract d. We will (with prejudice) describe contract c as a "nuisance contract."

Next we provide a characterization of when the various contracts will be observed. We do so under the assumption that the signal is "accurate" (that is, ρ_g is high and ρ_b is low, so that the signal is a good predictor of the state) and the signal is "cheap" (so that I is small). It is easily demonstrated that if the signal is sufficiently inaccurate or sufficiently expensive, a compound contract is not useful.

Theorem 2: If the signal is sufficiently cheap and accurate, then there exist values S^*, $\hat{S}$, such that the optimal contract depends on the required returns S in the following way: For $S \leq AT_1$ the simple, non-liquidating contract is optimal. For $S \in (AT_1, S^*]$, the nuisance contract is optimal. For $S \in (S^*, \hat{S}]$, demandable debt is optimal. For $S > \hat{S}$, no contract is feasible.

In other words, demandable debt will be observed when the returns that depositors can receive in alternate investments is relatively high. Precise bounds for "sufficiently accurate" and "sufficiently cheap" are provided in the proof in the appendix.

III. MULTIPLE DEPOSITORS WITH INDEPENDENT SIGNALS

In this section we develop a model for the case in which a number of depositors enter into contracts with the banker. As before, each depositor has \$1 to invest, and the banker has one "project" he can pursue The project costs Y and yields a total return of YT_i, which takes one of two values. Any deposits the banker receives in excess of Y can be used to yield the same competitive return S that depositors have available to them on their own. Deposits in excess of Y will be identified with "reserves."

We make the following natural assumptions about the difference between the two forms of bank assets, "project" and "reserves":

If the bank is liquidated, the value of the project decreases by the proportion L; the value of the reserves are unchanged.[22]

If the banker absconds then he takes the projects with him and receives $(1-A)YT_i$. The depositors retain the entirety of the reserves.[23] We strengthen assumption (2) as follows:

$$L < A(T_1/T_2) \qquad (5)$$

There are Z individuals available to enter into a contract with the bank. Of these individuals, K can receive signals by investing at a cost I; for the remainder, the cost of receiving a signal is prohibitive.[24] Signals are i.i.d. conditional on T_i. For any individual a "bad" signal is associated with reduced likelihood of the high productivity state, so, $\rho_b < \rho_g$ as before.

Supposing that all K individuals have invested in the signal, let N be the number who receive the "bad" realization. Given the i.i.d, structure, N is a sufficient statistic for T_i, and the probability that the realization is T_2 decreases with N.

The Contract from the Banker's Viewpoint

We start by examining only the incentive problem for the banker, taking the behavior of all depositors as given. We will return to the individual depositors' incentives in the succeeding section. For now, we assume that all K individuals who can invest in obtaining the information do so and report it truthfully.[25] A contract specifies an aggregate payment P and a liquidation decision Λ as functions of the number of depositors who announce observations of the bad signal. (In the subsequent section we will investigate a scheme for dividing aggregate payments among the depositors). Note therefore that the contract is the direct generalization of the contract in the previous section to a case of multiple signals.

After the announcement of the signals, the game tree is as before: if a liquidation is not mandated, the banker makes a decision whether to abscond. The following table describes the payoffs on each of the three nodes of the game tree:

	Banker Receives	Depositors Receive
Liquidation:	$(1-L)T_iY + (Z-Y)S - P$	P
No Liquidation		
Banker Absconds:	$(1-A)T_iY$	$(Z-Y)S$
Banker does not Abscond:	$T_iY + (Z-Y)S - P$	P

The optimal contract maximizes the banker's expected profits subject to three restrictions:

1) The expected payments to the depositors equal their aggregate reservation level:

$$SZ + KI$$

-- that is, all depositors must be compensated for the opportunity cost of their funds; in addition, any monitors must be compensated for the cost of monitoring.

2) In the case of liquidation, actual payment cannot exceed what is assumed feasible; as before we suppose that a liquidated investment Y pays off at most MY to the depositors. Thus the total payment to depositors out of the project and the reserves is

$$P \leq MY + (Z-Y)S \qquad \text{if } \Lambda = 1.$$

3) Finally we must consider the banker's incentive to abscond. If liquidation does not occur, then the banker will prefer to abscond whenever

$$AT_iY < P - (Z-Y)S.$$

If the inequality is reversed the banker prefers not to abscond.

As before, we define $\hat{S}$ to be the least upper bound of feasible expected returns to depositors from the project; if the required rate of return exceeds $\hat{S}$, no contract is feasible. The appendix calculates $\hat{S}$ explicitly.

Our first result is that for required returns which are sufficiently high, (but less than $\hat{S}$), the optimal contract calls for liquidation when the number of bad signals is high, and not when the number of bad signals is low. When the number of bad signals is low, there is a positive (but small) probability that the banker will abscond.

Theorem 3: For an interval of values of S ($\underline{S}$, $\hat{S}$], the optimal contract has the following form: there exists $\underline{N}$ such that

$$\text{If } N > \underline{N},\quad \Lambda(N) = 1 \quad \text{and} \quad P(N) = MY + (Z-Y)S$$

$$\text{If } N < \underline{N},\quad \Lambda(N) = 0 \quad \text{and} \quad P(N) = AT_2Y + (Z-Y)S$$

In other words, the contract has informed agents announce whether their signal was bad. If more than a critical number $\underline{N}$ announce bad signals, the bank is liquidated. If fewer than $\underline{N}$ announce bad signals the bank is not liquidated, and the banker chooses to abscond if the productivity draw was low.[26]

Note that Z is arbitrary in this contract. As Z increases the optimal P increases one-for-one: additional deposits beyond those invested in the project are held in reserves and returned to the depositors with certainty.[27]

Depositor Incentives

It remains to be shown that the total aggregate payment to depositors specified in the previous section can be divided among depositors in such a way as to maintain the incentives for low cost information depositors to invest in the signal, and to report it truthfully. In this section we derive a demandable debt contract which achieves this goal.

We make the following assumptions about the population of monitors and the signals:

Distributional assumptions: There are large numbers of potential depositors Z and potential monitors K. The cost of monitoring I is small. The probability of any one monitor receiving a bad signal is small. The probability of a bad realization of T is small (although the losses can be large).

In modeling a bank, each of these assumptions seems to us natural. The assumptions allow us to model the distribution of the number of bad signals as a Poisson distribution. More precise criteria for "small enough" or "large enough" are indicated in the appendix.

Note that as long as I is sufficiently small, it is always optimal to have all the potential monitors engage in investment.

The contract for all depositors is identical. Ex post depositors will pick one of two announcements within the contract. Since there are three information possibilities: observing "g", observing "b" or not making an investment, there will have to be some pooling in the outcomes. We will build a contract in which it is incentive compatible for the depositors who have made no investment to pool with those who have observed the good draw.

Each depositor's payoff depends on his announcement and the signal (if any) he observes. We let the symbol $EU(\sigma,\hat{\sigma})$ denote the expected return for a depositor who observes signal σ and announces signal $\hat{\sigma}$.

Individual depositors are subject to two sorts of constraints: participation constraints (i.e., the contract must give expected returns which are sufficient for depositors to participate) and incentive constraints. From the point of view of the individual depositors the contract must satisfy the following requirements:

1) Always announcing "g" gives an expected return of S, which exceeds the expected return from always announcing "b." This means that depositors

with high costs of gathering information will be willing to participate in the contract in the manner specified.

2) Announcing the observation truthfully gives a return of S + I, which exceeds the return from lying. If conditions in 1 are satisfied as well, then individuals with cost of I for investing are willing to make the investment in monitoring and report truthfully.

These constraints for individual depositors can be written as follows:

$$\lambda EU(\hat{g},g) + (1-\lambda)EU(\hat{g},b) - S \geq \lambda EU(\hat{b},g) + (1-\lambda)\ EU(\hat{b},b)$$

$$\lambda EU(\hat{g},g) + (1-\lambda)EU(\hat{b},b) - S + I \geq \lambda EU(\hat{b},g) + (1-\lambda)EU(\hat{g},b)$$

where λ is the prior probability of signal g.

The scheme we consider has payments of a particularly simple form: Any depositor announcing "b" receives the payment R with certainty. We can call an announcement "b" a "withdrawal of funds." If more than $\underline{N}$ depositors announce "b" the bank is liquidated; otherwise it is not, and the banker has the option of absconding. In any event, those depositors who do not announce "b" evenly split the aggregate payment to depositors described in the previous section, less the funds withdrawn. We call this scheme a "standard demandable debt contract."

Under a standard demandable debt contract, of course,

$$EU(\hat{b},b) = EU(\hat{b},g) = R.$$

However, for depositors who do not withdraw their funds the payment depends on the number of depositors N who do withdraw, and on whether the banker absconds. The following table describes the payments for a depositor who announces "g":

	Payoff to depositor who announces "g" if number of depositors announcing "b" is Less than $\underline{N}$	Greater than $\underline{N}$
And project realization is T_1	$\frac{AT_2 + (Z-Y)S - RN}{Z-N}$	$\frac{MY + (Z-Y)S - RN}{Z-N}$
T_2	$\frac{(Z-Y)S - RN}{Z-N}$	

For example, if more than $\underline{N}$ depositors withdraw funds, then the bank is liquidated and according to the contract, the total payment to depositors P is MY + (Z-Y)S; that quantity, less the withdrawn deposits RN is split among the remaining depositors Z-N, yielding the quantity in the rightmost column of the table. The remaining numbers are calculated in a similar fashion.

Given the probabilities of the realizations of T_i, and the probability of each signal contingent on T_i it is a straightforward matter to calculate $EU(\hat{g},b)$ and $EU(\hat{g},g)$. For this scheme, the incentive and participation constraints reduce to the following:[28]

$$EU(\hat{g},b) = R - I/(1-\lambda)$$

$$S > R$$

In the appendix we demonstrate that when an aggregate contract of the sort described in the previous section is optimal, it can always be implemented with a demandable debt scheme:

Theorem 4: Under the distributional assumptions and the conditions of the previous theorem, the optimal outcome can be achieved with a simple demandable debt contract.

The role of reserves in our model warrants discussion. By holding reserves the bank is able to guarantee early payment to a small number of monitors (those who receive bad signals) without forcing the bank to be placed into receivership. Reserves allow the bank to commit to the sequential service constraint (early withdrawals by those who run the bank), which supports the implementation of the contract between bankers and depositors. More familiar justifications for bank reserve holding includethe usefulness of reserves in meeting stochastic demands for conversion into gold (say, due to foreign transactions needs of depositors), or the contribution of reserves to an optimally diversified portfolio of bank assets. Our model adds to these transactions and portfolio motivations for holding reserves an "incentive-compatibility" demand for reserves.

III. SUMMARY, EXTENSIONS, AND LIMITATIONS

Summary

We have argued that historical demandable-debt banking can be understood as the optimal means of incentive-compatible intermediation in an environment of asymmetric information with potential for fraudulent behavior on the part of the banker. Monitoring by some depositors, and runs by monitors who receive bad signals, ensure sufficiently high payoffs to depositors in states of the world which would otherwise lead to malfeasance by the banker.

Agency problems are inherent in banking. Depositors entrust their endowments to bankers, who decide how to invest them, and have essentially unfettered immediate control over the depositors' funds. We capture this agency problem in a simple way by allowing the potential for "absconding" by the banker. The banker has the ability to remove funds from the bank. Absconding is socially wasteful; if the banker steals funds from the bank he uses a "leaky bucket," so that the amount he actually receives is less than the amount stolen.

If the required return for depositors is sufficiently high, then the banker may find it attractive to abscond rather than make the promised payment to depositors. Anticipating this, depositors will be unwilling to entrust their funds to the banker, and efficient intermediation will not take place. In other words, the possibility for a banker to abscond may make it difficult for him to attract depositors to his bank.

We introduce a liquidation technology that allows depositors, at a cost, to prevent the banker from absconding, and makes it possible for the banker to attract depositors. We show that under some circumstances the optimal arrangement has the depositor choose whether to liquidate the bank, contingent on a costly signal he receives. In good states, it will pay for the banker not to abscond and to pay the depositor as promised; in bad states, absent a liquidation announcement, the banker will abscond rather than pay as promised. Thus when monitors receive bad signals they call for liquidation.

If the signal is perfect and costless to the depositor, liquidation would only occur when there are bad loan investment realizations. If the signal is imperfect and costly, but not prohibitively so, it still makes sense to use the contingent liquidation contract, even though on occasion monitoring depositors may make errors in judging when to "run the bank," and force the bank to liquidate unnecessarily. Banks can fail either because the banker absconds, or because the depositor initiates a run on the bank. The purpose of a run is to prevent absconding from taking place.

In the case of multiple depositors, the bank uses reserves to offer guaranteed payments to early withdrawers, and to insulate itself from a few bad idiosynchratic signals. At the same time, under circumstances that probably would lead to costly absconding, depositors as a group are likely to order liquidation preemptively. The number of monitors and the threshold at which a bank liquidation is called for will be chosen optimally to minimize total expected costs of liquidation, absconding, and monitoring.

Transactability as a Motivation for Demandable Debt

Thus far we have argued that demandable debt intermediation may arise in order to permit profitable investment opportunities to be realized. In our models, there is no demand for transactability; therefore assets are valued entirely based on expected return. An important feature of demandable debt instruments historically, however, has been their use as a medium of exchange. In this section we briefly consider the implications of our model for the liquidity of demandable debt.

It is important to note from the outset that transactable instruments need not be demandable. Post-dated bills of exchange and post-dated bank notes were physically transactable instruments that existed in the nineteenth century in the United States (Dewey 1910). Their primary difference from demandable debt was that they could be redeemed, not on demand, but only on the date of maturity. Since such instruments could be maturity matched, they would seem to have none of the disadvantages of demandable debt. Nonetheless, demandable debt out-competed these as a medium of exchange.

In order to explain the relative liquidity of demandable debt one must explain why the ability to redeem a bank note or deposit on demand makes people more willing to accept it as a means of payment. We argue that under demandable debt, monitors and non-monitors alike are better informed of the market value of the debt instrument at all times.

The fact that "the bank is open" (that monitors have not called for a liquidation) is revealing to non-monitors. In the simplest, one-monitor case, the fact that the bank is open is fully revealing because the signal that the monitor receives takes one of two values. In the multi-monitor case, the fact that the bank is open is not fully revealing; it only indicates that fewer than the threshold number of bad signals have been announced. Even this information, however, places a lower bound on the value of the bank's liability.[29] If the liquidity of an asset depends on the extent to which information about its value is shared, then one would expect demandable debt to have been more liquid than other contracts with which it competed.[30] Thus it may be possible to view the liquidity of bank claims as a by-product of the solution to the agency problem.

While we argue that the transactability of demandable debt enhanced its attractiveness, it is interesting to note that demandable debt banking pre-dates the transactability of demandable debt. Thus the desirability of demandable debt contracting does not seem to have depended crucially on the transactability of the instruments. For example, historians of the Roman banking panic of 33 A.D. point out the importance of demandable debt in causing massive bank disintermediation. Roman banks, however, did not issue transactable claims like modern bank notes and checks.[31]

The "liquidity premium" which demandable debt enjoys can be included into our framework by reducing the level of the required return S on demandable debt by the amount of the liquidity premium. In other words, demandable debt would face a lower threshold reservation level to satisfy than the non-liquidating compound contract. This implies an expansion of the parameter values for which demandable debt is preferred to the "nuisance" contract.

It may appear paradoxical that the most illiquid assets in the economy (those which require information-intensive bank intermediation) back the most liquid liabilities. According to our model the answer to this paradox lies in the common information which monitors generate regarding the value of a demandable-debt contract. Thus even though bank assets may be more difficult to appraise, and may be appraised less exactly, the form of a demandable-debt contract makes it relatively liquid.

In other words, the answer to the puzzle of why banks use the most illiquid assets in the economy to back the most liquid liabilities can be traced to the fact that demandable debt simultaneously solves the agency problem of financial intermediation, and the asymmetric information problem of asset transactions. The economies of scope between financing information-intensive projects and providing liquidity may explain why banks out-competed other intermediaries in providing small denomination liquid assets.

Limitations of the Analysis

It is worth noting some of the important limitations of our framework. First, our goal is to explain the historic importance of demandable debt in banking. In today's more regulated environment, where for example, regulations on clearing through the Federal Reserve System has favored demandable debt instruments, and where deposit insurance makes depositor monitoring less important, demandable debt may persist simply as an artefact of regulation.

Second, our method in arguing for the disciplinary role of demandable debt is to show that a demandable debt contract is an optimal arrangement in specific environments. In general this could not possibly be the case; the truly optimal arrangement in most general situations exceeds in complexity anything ever observed in actual economies. Our ability to devise extremely complicated arrangements in models exceeds our ability to devise correspondingly complicated arrangements in real life, perhaps because the transactions costs of designing and implementing such arrangements are prohibitive. Our claim for the optimality of demandable debt is more modest: The use of demadable debt helps in solving an incentive problem that arises for real banks. It is also an extremely simple contract. In the situations examined here it in fact achieves the optimum possible under all allocations. In more complex situations it does well even if some extremely complicated alternatives might do somewhat better.

Third, our account is one of individual banks and individual bank liquidations, not of systems of banks or systemic crises. We are dealing with bank runs, not economy-wide bank panics. We are only attempting to model the operation of demandable debt in normal times, when the rules required banks to pay on demand. In historical practice, the provisions of demandable debt, including liquidation, were suspended during crisis.[32] That is to say, demandable debt was a contingent rule; it required banks to meet the threat of runs in response to idiosyncratic problems; but it allowed banks to escape convertibility on demand in the face of systemic disturbances. Only individual bank difficulties led to placing a bank in receivership. Suspension and interbank relations during panics are important as well, but doing this topic justice requires a larger analysis than the one we undertake in this paper.[33]

APPENDIX

To make the proofs more concise we will make the following modifications to the structure without loss of generality. First we can treat the choice of $e = 0$ as generating a "signal" which is pure noise. Let $\rho(\sigma,e)$ be the posterior probability of the high realization given an investment e and the receipt of signal σ. Then

$$\rho(\sigma,e) = \rho_\sigma \text{ if } e = 1$$
$$= \gamma \text{ if } e = 0.$$

From this we can generate expectations as functions of σ and e.

A <u>contract</u> is a quartet $(\Lambda_b, P_b, \Lambda_g, P_g)$ with the restriction that $P_\sigma \leq M$ if $\Lambda_\sigma = 1$. We can identify simple contracts with quartets $(\Lambda_b, P_b, \Lambda_g, P_g)$ in which $\Lambda_b = \Lambda_g$ and $P_b = P_g$. Finally, by the revelation principle, we can restrict attention to contracts in which the announcement $\hat{\sigma}(\sigma) = \sigma$ in equilibrium.

The choice of a contract generates a game in which the depositor picks a strategy consisting of investment and announcement $\{e, \hat{\sigma}(\sigma)\}$ and the banker picks a probability of absconding as a function of the depositor's announcement and the realization of T_i: $\alpha(\hat{\sigma},T_i)$. (Let $\alpha = 1$ indicate absconding, and $\alpha = 0$ indicate no absconding).

The banker's equilibrium strategy can be described simply:

$$\alpha = 1 \text{ if } \Lambda_\sigma = 0 \text{ and } P_\sigma > AT_i \,,$$
$$= 0 \text{ if } \Lambda_\sigma = 0 \text{ and } P_\sigma < AT_i \,.$$

Let the function $U(\hat{\sigma},\sigma)$ represent the expected profits of the depositor conditional on his observation σ and on his announcement $\hat{\sigma}$. Then

$$U(\hat{\sigma},\sigma) = P_{\hat{\sigma}} \qquad \text{if } \Lambda_{\hat{\sigma}} = 1$$
$$= P_{\hat{\sigma}}(1 - E\{\alpha(P_\sigma,T_i)|\sigma,e\}) \qquad \text{if } \Lambda_{\hat{\sigma}} = 0.$$

Lemma: An optimal contract generates an equilibrium with $e = 0$ if and only if the contract is a simple contract.

Proof: If $e = 0$ in equilibrium, then

$$U(\hat{\sigma},g) = U(\hat{\sigma},b)$$

because the signal conveys no information. Thus

$$U(g,g) = U(b,g) = U(g,b) = U(b,b) ;$$

otherwise the contract would not induce truth-telling. If the depositor is indifferent between announcements, but the banker is not, then the contract is sub-optimal, because banker profits could be improved by having the depositor always make the announcement the banker prefers.

Thus both depositor and banker must be indifferent between the stated outcomes for the two announcements. This requires the outcomes themselves to be identical.

The converse follows from the next lemma. Define λ to be the probability of observing the signal g, so that

$$\lambda = (\gamma - \rho_b) / (\rho_g - \rho_b)$$

Lemma: A contract generates an equilibrium with $e = 1$ if and only if

$$U(\hat{b},b) - U(\hat{g},b) \geq I/(1-\lambda) \tag{A.1}$$

$$U(\hat{g},g) - U(\hat{b},g) \geq I/\lambda \tag{A.2}$$

Proof: If the depositor does not make the investment in the signal, he has the following options: he could always report that the signal was g; he could always report that the signal was b; or he could make a randomization. In equilibrium, $e = 1$ if and only if all of these strategies are dominated by the strategy in which the depositor makes the investment and reports the signal truthfully, that is, if and only if

$$\lambda\ U(\hat{g},g) + (1-\lambda)\ U(\hat{b},b) - I \geq \lambda\ U(\hat{g},g) + (1-\lambda)\ U(\hat{g},b)$$

$$\lambda\ U(\hat{g},g) + (1-\lambda)\ U(\hat{b},b) - I \geq \lambda\ U(\hat{b},g) + (1-\lambda)\ U(\hat{b},b)$$

These two inequalities simplify to those in the statement of the lemma.

Note that these inequalities imply that if e = 1 the contract is not a simple contract. Note also that these inequalities automatically imply that truth-telling is preferred:

$$U(\hat{b},b) - U(\hat{g},b) \geq 0$$

$$U(\hat{g},g) - U(\hat{b},g) \geq 0$$

For reference figure A1 depicts $U(\hat{\sigma},g)$ and $U(\hat{\sigma},b)$ as functions of $P_{\hat{\sigma}}$ in the case where $\Lambda_{\hat{\sigma}} = 0$ and e = 1. Note that

$$U(\hat{\sigma},g) \geq U(\hat{\sigma},b) \quad \text{for all } P_{\hat{\sigma}} \tag{A.3}$$

As payment promised in the contract increases up to the level AT_1, the expected payoff to the depositor increases one-for-one. From AT_1 to AT_2 it increases less than one-for-one, since the depositor knows he will receive the payment only in the good realization which happens with probability ρ_σ. Promised payments above AT_2 are irrelevant, since the depositor knows he will never receive them. Note also that there is a unique equilibrium value of the banker's response α, except when the contracted repayment equals AT_1 or AT_2. In this case, the banker is indifferent among absconding, not absconding, and any randomization of those two responses. In this case U is a correspondence, where the particular value depends on the choice of banker's strategy in the equilibrium.

FIGURE A.1 ABOUT HERE

<u>Proof</u> <u>of</u> <u>Theorem</u> <u>1</u>: The claim that an optimal contract must conform to one of the four cases listed in the theorem is equivalent to the following claims:

a: If $P_\sigma \leq AT_1$ in an optimal contract, then $\Lambda_\sigma = 0$.

b: If the optimal contract is a compound contract, then

$\Lambda_g = 0$ and $P_g \in (AT_1, AT_2]$.

c: If the optimal contract has $e = 1$, then $P_g > P_b$.

Claim a is fairly immediate: If a contract has liquidation with a promised price less than or equal to AT_1, it is obviously suboptimal. The depositor can be guaranteed the same amount without liquidation, since the banker will never desire to abscond. And if there is no liquidation, the banker's profits are greater.

Most of claim b is derived from the following lemma:

<u>Lemma:</u> If an optimal feasible contract has $e = 1$, then $\Lambda_g = 0$, and $P_g \in [AT_1, AT_2]$.

<u>Proof:</u> If $\Lambda_g = 1$, then the return the depositor receives if he says g is independent of the true signal. If $\Lambda_g = 0$ but P_g is not in the specified interval, then the return is also independent of the signal (see figure A1). In either case, that is,

$$U(g,b) = U(g,g).$$

But by the incentive conditions (A.1) - (A.2), this implies

$$U(\hat{b},b) > U(\hat{b},g)$$

contradicting (A.3).

Claim c follows by noting first that P_b is certainly less than AT_2 (if it is equal to or greater than AT_2, then there is a lower price for which all incentive constraints are satisfied for the depositor, and which gives greater profits to the banker). If this is the case, (A.1) implies $P_b < P_g$. (Again, this can be verified in figure A1).

The final step in demonstrating claim b is to verify that any feasible compound contract in which $P_g = AT_1$ is dominated by a simple contract with P_g less than or equal to AT_1.

Proof of Theorem 2:

If $S \leq AT_1$ it is immediate that the optimal contract is a simple non-liquidating contract. Such a contract entails no social waste, since the required payback is so low that the banker never has an incentive to abscond. Therefore, assume $S > AT_1$.

It is straightforward to show that whenever a contract is feasible and optimal, the depositor receives a return equal to his outside return plus compensation for any investment he has made. If the nonliquidation contract is chosen, then profits are

$$\lambda\,(\rho_g T_2 + (1-\rho_g)(1-A)T_1) + (1-\lambda)\,(\rho_b T_2 + (1-\rho_b)(1-A)T_1) - S. \qquad (A.4a)$$

If the simple liquidation contract is chosen, then profits are

$$\lambda\,(1-L)\,(\rho_g T_2 + (1-\rho_g)T_1) + (1-\lambda)\,(1-L)\,(\rho_b T_2 + (1-\rho_b)T_1) - S \qquad (A.4b)$$

If the nuisance contract is chosen then profits are

$$\lambda\,(\rho_g T_2 + (1-\rho_g)(1-A)T_1) + (1-\lambda)\,(\rho_b T_2 + (1-\rho_b)T_1) - S - I \qquad (A.4c)$$

If the demandable debt contract is chosen, then profits are

$$\lambda\,(\rho_g T_2 + (1-\rho_g)(1-A)T_1) + (1-\lambda)\,(\rho_b (1-L)T_2 + (1-\rho_b)(1-L)T_1) - S - I \qquad (A.4d)$$

Profits in a nuisance contract are greater than profits in a demandable debt contract. If

$$\rho_g > (A-L)\,T_1 \,/\, [(A-L)T_1 + L\,T_2] > \rho_b, \qquad (A.5)$$

then the demandable debt contract dominates any simple contract for I sufficiently small.

Assuming (A.5), the optimal contract is the nuisance contract if there exists an incentive compatible one which yields the depositor a return equal to S + I. If there exists no such contract, the optimal contract is a demandable debt contract if there exists an incentive compatible one which yields the depositor a return equal to S + I. If neither of these contracts exist, the optimal contract is a simple contract if one is feasible.

In any feasible complex contract, the return to the depositor must exceed $S + I$:

$$\lambda P_g \rho_g + (1-\lambda) P_b \geq S + I. \qquad (A.6)$$

A nuisance contract satisfying Theorem 1 (one with $\Lambda_b = \Lambda_g = 0$ and P_b not equal to P_g) exists if and only if there are a pair of payments P_b and P_g satisfying restriction (c) of Theorem 1, (A.6), and the incentive compatibility restrictions (A.1) and (A.2). (The last two reduce to)

$$P_g \rho_g \geq P_b + I/\lambda \qquad (A.7)$$

$$P_b \geq P_g \rho_b + I/(1-\lambda) \qquad (A.8)$$

The following five inequalities are necessary and sufficient for the existence of such a P_g and P_b

$$\left.\begin{array}{ll} S+I \leq \lambda\, AT_2\, \rho_g + (1-\lambda)\, AT_1 & \text{(a)} \\[2ex] S+I \leq AT_1 \left(1 + \lambda \dfrac{\rho_g - \rho_b}{\rho_b}\right) - \lambda \dfrac{\rho_g}{\rho_b} \dfrac{I}{1-\lambda} & \text{(b)} \\[2ex] S+I \leq AT_2\, \rho_g - \dfrac{1-\lambda}{\lambda} I & \text{(c)} \\[2ex] \dfrac{I}{\rho_g - \rho_b} \left(\dfrac{1}{\lambda} + \dfrac{1}{1-\lambda} \right) \leq AT_2 & \text{(d)} \\[2ex] \dfrac{I}{\rho_g - \rho_b} \left(\dfrac{\rho_g}{\lambda} + \dfrac{\rho_b}{1-\lambda} \right) \leq AT_1 & \text{(e)} \end{array}\right\} \quad (A.9)$$

For each I sufficiently small, there exists $S^* > AT_1$ such that these inequalities are satisfied by all S less than S^*. In the interval $(AT_1, S^*]$ the nuisance contract is optimal.

The demandable debt contract is feasible if there exist payments P_g and P_b satisfying (A.1) and (A.2), condition (d) of theorem 1, (A.6); and the feasibility restriction that $P_b \leq M$. Necessary and sufficient conditions for such payments to exist are

$$
\left.
\begin{array}{ll}
S+I \le \lambda\, AT_2\, \rho_g + (1-\lambda)\, M & \text{(a)} \\
S+I \le M\, (1 + \lambda \dfrac{\rho_g - \rho_b}{\rho_b}) \;-\; \lambda \dfrac{\rho_g}{\rho_b} \dfrac{I}{1-\lambda} & \text{(b)} \\
S+I \le AT_2\, \rho_g - \dfrac{1-\lambda}{\lambda}\, I & \text{(c)} \\
\dfrac{I}{\rho_g - \rho_b} \left(\dfrac{1}{\lambda} + \dfrac{1}{1-\lambda} \right) \le AT_2 & \text{(d)} \\
\dfrac{I}{\rho_g - \rho_b} \left(\dfrac{\rho_g}{\lambda} + \dfrac{\rho_b}{1-\lambda} \right) \le M & \text{(e)}
\end{array}
\right\} \quad \text{(A.10)}
$$

Note that these conditions are identical to the five conditions in (A.9) with M substituted for AT_1. For I sufficiently small, the conditions are satisfied by all S less than or equal to some critical value S**, where S** > S*. In the interval (S*,S**] demandable debt is the optimal contract.

Provided that

$$\rho_g > M / AT_2 > \rho_b \qquad \text{(A.11)}$$

for I sufficiently small,

$$S^{**} = \lambda\, AT_2\, \rho_g + (1-\lambda)\, M - I,$$

which exceeds the maximum possible payout to the depositor under any simple contract. Thus for required returns beyond S**, no contract is feasible.

Thus the theorem is proved provided that the signal is sufficiently accurate in the sense that inequalities (A.5) and (A.11) hold and provided that I is small enough -- that is, if it satisfies the following eight requirements: Expression (A.4d) exceeds (A.4a) and (A.4b); inequalities (A.9d) and (A.9e) hold; the right sides of (A.9b) and (A.9c) exceed AT_1; the right sides of (A.10b) and (A.10c) exceed the right side of (A.10a). Given (A.5) and (A.11), these eight inequalities hold strictly for I = 0; therefore they hold in an interval of I small but positive.

Proof of Theorem 3:

We begin by determining the optimal contract in general.

Lemma: The optimal contract in general involves three regions. For high values of N, the contract mandates liquidation. For intermediate values of N, liquidation is not mandated, but aggregate payment is set sufficiently low that absconding never occurs. For low values of N, payment is set sufficiently high that absconding takes place in low productivity outcomes.

Proof: In this proof we explicitly include the possibility of randomized outcomes for various realizations of N. Recall that Λ is the indicator variable for a liquidation and α is the indicator variable for absconding. Let N_α, N_Λ, N_o be any triple of integers t such that.

$$\Pr(\Lambda = 1|N_\Lambda) > 0$$

$$\Pr(\alpha = 1,\ \Lambda = 0|N_\alpha) > 0$$

$$\Pr(\alpha = 1|N_o)\ 0 \text{ and } \Pr(\Lambda = 0|N_o) > 0\ .$$

(Thus, N_Λ is an N for which liquidation can occur, N_α is an N for which absconding can occur, N_o is an N for which absconding does not occur and for which liquidation need not occur). Let

X_Λ be a subset of the event $\{\Lambda = 1 \cap N = N_\Lambda\}$

X_α be a subset of the event $\{\Lambda = 0 \cap N = N_\alpha\}$

X_o be a subset of the event $\{\Lambda = 0 \cap N = N_o\}$

each with identical probability ϵ, and independent of T. Suppose $N_\Lambda < N_o$. Then by reversing behavior on X_Λ and X_o (i.e. setting $\Lambda = 0, \alpha = 0$ on X_Λ and Λ

= 1 on X_o and reversing the payoffs between the two), we do not affect the distribution of aggregate payoffs to the depositors, but since the good outcome is more likely for N_o we increase the expectation of the banker's profits.

Suppose $N_o < N_\alpha$. Then the distribution of $T_i|N_o$ stochastically dominates the distribution $T_i|N_\alpha$. Reverse the payoffs on X_α and X_o and let α be determined by the incentive compatibility condition. This change reduces the likelihood of absconding; it increases both depositors' expected payments and the banker's profits.

If liquidation occurs, P can be no greater than

$MY + (Z-Y)S$

If absconding occurs with zero probability, P can be no greater than

$AT_1 + (Z-Y)S$

If absconding occurs with less than probability 1, P can be no greater than

$AT_2 + (Z-Y)S$

Next we show that in each of the three regions described in the previous lemma, these maximal payoffs are binding. It is clear that for a maximum, constraint (3.6) is binding. If $S > AT_1$ the total payoff to depositors must be greater than AT_1 in some state, and this implies that either liquidation or the possibility of absconding must have positive probability in some state. Both liquidation and the possibility of absconding reduce social benefits. In other words, for a given required aggregate expected payment to the depositors, profits are maximized by making the region of no liquidation and no absconding (the intermediate region described in the lemma) as large as possible. If in some state the payoff is not at its maximum, then by increasing the payoff to the maximum in that state we could expand the intermediate region.

Next we determine the boundaries of the three regions. In our calculations we will treat the variable N as if it had a continuous distribution. The calculation with N a discrete random variable is analogous and the resultant conditions are identical, but the complete description in that case consumes considerable notation and is therefore omitted.

Let F(N) be the distribution of N, and let $\rho(N)$ be the probability that $T_i = T_2$ conditional on N bad realizations. Given the results so far, the choice of an optimal contract can be reduced to a choice of two numbers, $\underline{N}$ and $\bar{N}$, to maximize

$$\int_0^{\underline{N}} [\rho(N)\ T_2 + (1-\rho(N))\ (1-A)\ T_1]\ dF(N)$$

$$+ \int_{\underline{N}}^{\bar{N}} [\rho(N) T_2 + (1-\rho(N))\ T_1]\ dF(N)$$

$$+ \int_{\underline{N}}^{K} [(1-L)\ \rho(N)\ T_2 + (1-L)\ (1-\rho(N))\ T_1]\ dF(N) \tag{A2.0}$$

subject to

$$\int_0^{\underline{N}} \rho(N)\ AT_2Y\ dF(N) + \int_{\underline{N}}^{\bar{N}} AT_1Y\ dF(N) + \int_{\bar{N}}^{K} MY\ dF(N) \geq YS + KI \tag{A2.1}$$

and

$$\underline{N} \leq \bar{N} \tag{A2.2}$$

The maximand is basically the sum of the firm profits and the expected payments to the depositors. Constraint (A2.1) is essentially a transformation of the aggregate participation constraint; the righthand side is a transformation of the aggregate reservation payment for all depositors.

Define $\underline{\rho} = \rho(\underline{N})$ and $\bar{\rho} = \rho(\bar{N})$. Then the first order conditions for this maximization problem are as below:

$$[\bar{\rho}T_2 + (1-\bar{\rho})\ T_1] - [\bar{\rho}\ (1-L)\ T_2 + (1-\bar{\rho})\ (1-L)\ T_1] - \sigma\ [M - AT_1] = 0$$

$$-[\underline{\rho}\ T_2 + (1-\underline{\rho})\ T_1\ ] + [\underline{\rho}\ T_2 + (1-\underline{\rho})(1-A)\ T_1] - \sigma\ (AT_1\ -\underline{\rho}AT_2) = 0$$

where $\bar{\rho} < \underline{\rho}$ and σ is the Lagrange multiplier attached to constraint (A2.1). As the right side increases from AT_1Y in constraint (A2.1), σ increases from zero, $\bar{\rho}$ rises, and $\underline{\rho}$ falls -- in other words, the region in which there is no absconding and no liquidation shrinks. Finally, there is a critical level such that for σ greater than the critical level constraint (A2.2) becomes binding. From this point, the solution is defined by constraints (A2.1) and (A2.2) holding with equality. For YS + KI above this level, the middle region is degenerate, and the contract takes the form described in the text. When the middle region is degenerate the right side of (A2.2) is calculated from the following formula

$$W(\underline{N}) = \int_0^{\underline{N}} \rho(N)\ AT_2Y\ dF(N) + \int_{\underline{N}}^{K} MY\ dF(N)$$

where $\underline{N}$ is the boundary between the remaining two regions.

The remaining portion of the proof demonstrates that given $LT_2 < AT_1$, the set of values W = YS + KI for which the optimum has a degenerate middle region is a closed interval $[\underline{W}, \bar{W}]$ with $\underline{W} < \bar{W}$.

The lower end of the interval is found as follows: Combine the above first order conditions as follows:

$$\frac{L(T_2\bar{\rho}+ T_1(1-\bar{\rho}))}{AT_1} = \frac{M - AT_1}{\underset{-}{\rho}AT_2 - AT_1} \tag{A2.3}$$

Solve when $\bar{\rho} = \underset{-}{\rho}$ for the unique positive root. (Call it ρ^*). Then

$$\underset{-}{W} = W(N^{-1}(\rho^*))$$

Provided $LT_2 < AT_1$, as W increases beyond the critical point, the optimum is achieved by reducing the region in which liquidation occurs and increasing the region in which absconding becomes a possibility. In other words, as W increases it takes more and more individuals to call for a liquidation.

The upper end of the interval is found by determining the maximum feasible amount for $W(\underline{N})$. The maximum occurs at $N^{-1}(M/AT_2)$. By using the equation (A2.3) it is possible to verify that

$$\rho^* > M/AT_2$$

Finally, we note that $\hat{S}$ as described in the text is equal to $\bar{W}/Y$. If S exceeds this amount, no matter how small I is there is no way to pay the depositors the required amount in aggregate.

<u>Proof of Theorem 4</u>

Given the payoff structure we can explicitly write $EU(g,\hat{b})$ as follows:

$$EU(\hat{g},b) = \rho_b \left[\sum_{N=0}^{K} Pr(N|T_2)\frac{S(Z-Y) - NR}{Z-N} + \sum_{N=0}^{\underline{N}} Pr(N|T_2)\frac{AT_2Y}{Z-N} + \sum_{N=\underline{N}+1}^{K} Pr(N|T_2)\frac{YM}{Z-N} \right]$$

$$+ (1-\rho_b) \left[\sum_{N=0}^{K} Pr(N|T_1)\frac{S(Z-Y) - NR}{Z-N} + \sum_{N=\underline{N}+1}^{K} Pr(N|T_1)\frac{YM}{Z-N} \right]$$

Where $Pr(N|T_i)$ is the probability of N individuals out of K receiving bad signals conditional on the productivity draw being T_i and individual 1 already having received a bad signal. (Recall that ρ_b is the probability of T_2 conditional on an individual's observing a bad realization of the signal.)

From the incentive conditions we know that for sufficiently small I, the structure satisfies the conditions if and only if we can find Z and R such that

$$EU(\hat{g},b) = R$$

$$S > R$$

By combinining these two conditions with the formula for $EU(\hat{g},b)$, and simplifying the expressions we see that the following condition is equivalent:

$$S \left[\sum_{N=0}^{K} Pr(N|b)\frac{1}{Z-N} \right] > AT_2\rho_b \left[\sum_{N=0}^{\underline{N}} Pr(N|T_2)\frac{1}{Z-N} \right] + M \left[\sum_{N=\underline{N}+1}^{K} Pr(N|b)\frac{1}{Z-N} \right]$$

where $Pr(N|b) = \rho_b \, Pr(N|T_2) + (1-\rho_b)Pr(N|T_1)$

If Z is very large, this inequality can be approximated by

$$S > AT_2 \; Pr \; (T_2 \text{ and } N \leq \underline{N} \mid b) + M \; Pr \; (N \geq \underline{N} \mid b)$$

(Note also that Z must be greater than N and greater than $Y + RK/S$ so that all payments specified for all individuals are non negative in our contract.)

It remains only to show that for $S = \hat{S}$, calculated at the end of the previous proof, this strict inequality holds. But using the same approximation for large Z, the definition of $\hat{S}$ implies that

$$\hat{S} > AT_2 \Pr(T_2 \text{ and } N \leq \underline{N}) + M \Pr(N > \underline{N})$$

It is therefore sufficient to demonstrate that

$$AT_2 \Pr(T_2 \text{ and } N \leq \underline{N} \mid g) + M \Pr(N > \underline{N} \mid g)$$
$$\geq AT_2 \Pr(T_2 \text{ and } N \leq \underline{N} \mid b) + M \Pr(N > \underline{N} \mid b).$$

Given that $AT_2 > M$, we only need to demonstrate that

$$\Pr(T_2 \mid N \leq \underline{N}, b) \leq \Pr(T_2 \mid N \leq \underline{N}, g)$$

which in turn is equivalent to

$$\frac{\Pr(N \leq \underline{N} - 1 \mid T_2) \Pr(b \mid T_2)}{\Pr(N \leq \underline{N} \mid T_2) \Pr(b \mid T_2)} \leq \frac{\Pr(N \leq \underline{N} - 1 \mid T_1) \Pr(b \mid T_1)}{\Pr(N \leq \underline{N} \mid T_1) \Pr(b \mid T_1)}$$

Using the Poisson distribution, the ratio in the above inequality for a given T_i can be approximated by

$$\frac{\sum_{x=0}^{\underline{N}-1} f(x;\ \mu_i)}{\sum_{x=0}^{\underline{N}} f(x;\ \mu_i)} \quad \mu_i$$

where

$$f(x,\ \mu) = \mu^x e^{-\mu} / x!$$

For $\underline{N}$ sufficiently large, this quantity is decreasing in μ. From the previous theorem, increases in S increase the cutoff level of $\underline{N}$. It is therefore only necessary to determine that $\underline{N}$ can be sufficiently large without rendering the contract infeasible, that is, without allowing the quantity $\rho(\underline{N})$ to fall below M/AT_2. Direct calculation reveals that if $\underline{N}$ satisfies the following inequality, then the contract remains feasible:

$$(\mu_2/\mu_1)^{\underline{N}} \geq \frac{M}{AT_2 - M} \quad e^{-(\mu_1 - \mu_2)} \quad \frac{1-\gamma}{\gamma}$$

By setting γ, the probability of the good outcome, sufficiently large, we can make the maximum feasible $\underline{N}$ arbitrarily large, thereby guaranteeing that the incentive constraint is satisfied without the contract becoming infeasible.

REFERENCES

Akerlof, George, "The Market for 'Lemons': Qualitative Uncertainty and the Market Mechanism," Quarterly Journal of Economics, 1970, 488-500.

Benston, George J. and George G. Kaufman, "Risks and Failures in Banking: Overview, History and Evaluation," Federal Reserve Bank of Chicago, Staff Memorandum, SM 86-1, 1986.

Bernanke, Ben and Gertler, Mark, "Banking and Macroeconomic Equilibrium," in Barnett, William and Singleton, Kenneth J.,eds., New Approaches to Monetary Economics (Proceedings of the Second International Symposium in Economic Theory and Econometrics), Cambridge: Cambridge University Press, 1987.

Bryant, John, "A Model of Reserves, Bank Runs and Deposit Insurance," Journal of Banking and Finance 4, 1980, 335-344.

Calomiris, Charles W. and Charles M. Kahn "The Private Regulation of Banking by Banks," unpublished working paper, December 1988.

Calomiris, Charles W. and Schweikart, Larry, "Was the South Backward?: North-South Differences in Antebellum Banking During Crisis and Normalcy," Working paper, Federal Reserve Bank of Chicago, 1988.

Cannon, James G., Clearing-Houses: Their History, Methods and Administration, New York: D. Appleton and Co., 1908.

Chari, V.V. and Jagannathan, Ravi, "Banking Panics, Information, and Rational Expectations Equilibrium," Journal of Finance 43, July 1988, 749-763.

Comptroller of the Currency, Annual Report, 1920.

Comptroller of the Currency, "An Evaluation of the Factors Contributing to the Failure of National Banks," Unpublished Memorandum, January 1988.

Cone, Kenneth R., "The Regulation of Depository Institutions," Ph.D. Dissertation, Stanford University, 1983.

Diamond, Douglas W., "Financial Intermediation and Delegated Monitoring," R.E.Stud. 51, July 1984, 393-414.

Diamond, Douglas W. and Dybvig, Philip, "Bank Runs, Deposit Insurance, and Liquidity," J.P.E. 895, June 1983, 401-419.

Engineer, Merwan, "Bank Runs and the Suspension of Demand Deposit Withdrawals," unpublished, Queen's University, July 1987.

Fama, Eugene F., "Contract Costs and Financing Decisions," Working Paper 145, Center for Research in Security Prices, University of Chicago, July 1988.

Ferrin, A. W., "The Business Panic of A.D. 33," Moody's Magazine, August 1908, 81-2.

Gendreau, Brian C. and Prince, Scott S., "The Private Costs of Bank Failures: Some Historical Evidence," Federal Reserve Bank of Philadelphia Business Review, March/April 1986, 3-14.

Golembe, Carter H. and Warburton, Clark, Insurance of Bank Obligations in Six States During the Period 1829-1866, Unpublished manuscript, Federal Deposit Insurance Corporation, 1958.

Gorton, Gary, "Self-Regulating Banking Coalitions," Unpublished manuscript, 1989.

Gorton, Gary and Pennacchi, George, "Transactions Contracts," Finance Department, the Wharton School, University of Pennsylvania, 1988, presented at the Garn Institute of Finance Academic Symposium, August 18-20, 1988.

Grossman, Sanford J. and Stiglitz, Joseph E., "On the Impossibility of Informationally Efficient Markets," American Economic Review 70, June 1980, 393-408.

Jacklin, Charles J., "Demand Deposits, Trading Restrictions, and Risk Sharing" in Prescott, Edward C. and Wallace, Neil, eds., Contractual Arrangements for Intertemporal Trade (Minnesota Studies in Macroeconomics, vol. 1), Minneapolis: University of Minnesota Press, 1987.

Jacklin, Charles J., "Demand Equity and Deposit Insurance," Stanford University Graduate School of Business, 1988, presented at the Garn Institute of Finance Academic Symposium, August 18-20, 1988.

Jacklin, Charles J. and Bhattacharya, Sudipto, "Distinguishing Panics and Information-based Bank Runs: Welfare and Policy Implications," Journal of Political Economy 96, June 1988, 568-592.

Klein, Benjamin, "The Competitive Supply of Money," Journal of Money, Credit and Banking, November 1974, 423-453.

Kyle, Albert S., "An Equilibrium Model of Speculation and Hedging," Ph.D. Dissertation, University of Chicago, 1981.

Nicholas, Henry C., "Runs on Banks," Moody's Magazine, December 1907, 23-6.

Smead, E.L., "Bank Suspensions in 1927 and during 1921-1927," Unpublished Memorandum, Federal Reserve Board of Governors, April 11, 1928.

Townsend, Robert M., "Optimal Contracts and Competitive Markets with Costly State Verification," Journal of Economic Theory 21, October 1979, 265-293.

Villamil, A.P., "Demand Deposit Contracts, Suspension of Convertibility, and Optimal Financial Intermediation," University of Illinois, Urbana-Champaign, 1988.

Williamson, Stephen D., "Liquidity, Banking, and Bank Failures," International Economic Review 29, February 1988, 25-44.-44.

NOTES

[1]Financing through stocks and bonds, commercial paper, or longer-term bank liabilities was less important historically. Until recently, few firms had access to centralized markets in stocks or bonds, while commercial paper issues typically have been limited to the financing of trade or the short-term working-capital needs of only the most reputable enterprises.

[2]Cone (1983) shows that, in a world of full information, the risk of depositor liquidation under demandable debt is absent provided that financial intermediaries are maturity-matched.

[3]A detailed treatment of the rationale, technology and costs of systemic suspension of convertibility are beyond the scope of our paper. Although there is much interest in the role of demandable debt during banking crises, it seems important to us first to understand the role of demandable debt during normal periods of financial operation, and only afterward to examine behavior in exceptional periods in which demandable debt provisions were suspended.

We view suspension of convertibility as a relatively cumbersome alternative solution to the agency problem inherent in banking, as bankers, legislators and jurists susbstituted for the normal bank monitoring we model in this paper as occuring via demandable debt contracting. Our framework is perhaps best viewed as an explanation of the role of demandable debt in imposing "market discipline" on the behavior of banks during non-panic times.

[4]Jacklin and Bhattacharya (1988) gives a concise but useful review of these approaches.

[5]Fundamental papers which utilize this approach are by Bryant (1980), Diamond and Dybvig (1983), and Jacklin (1987). A major focus of Jacklin's work is the extent to which traded assets can substitute for non-traded bank deposits. In Jacklin (1988) the instruments considered are shares whose dividend depends on the market price of the asset. Our focus is on an era in which centralized markets for most assets are non-existent (presumably due to prohibitive information and transactions costs), so that the possibility of such instruments does not arise.

[6]The authorities sometimes gave noteholders of failed banks early access to bank assets by selling off bank bond holdings, but most demandable debt (including demand deposits) was not redeemable until court-appointed receivers and courts could value all assets and settle on a plan for distributing them among liability holders. For a model emphasizing the costs to depositors of delay in liquidation, see Engineer (1987).

[7]See Calomiris and Schweikart (1988).

[8]Smead (1928). For discussions of the role of fraud in earlier eras see Golembe and Warburton (1958), Calomiris and Schweikart (1988), and Bentson and Kaufman (1986). Data on national bank failures, by cause, can be found in the Annual Report of the Comptroller of the Currency, 1920, pp. 56-79. On the importance of fraud in more recent bank failures, see Comptroller of the Currency (1988).

[9] Nicholas (1907), p. 26, dismisses the importance of withdrawals by small depositors in causing bank liquidation. He writes, "If a bank is actually in bad shape there is far more likelihood of its initial condition being discovered by other banking institutions than by the individual depositors of the bank...A run is sometimes started in this manner...and continues until it has practically wiped out the reserves of the suspected institution, the ordinary depositors receiving their first information regarding the position of the bank when that institution is finally forced to close its doors and formally apply for a receiver." This discussion makes important points about bank runs which appear in our model below: Some depositors are informed while others are not. Only runs by informed depositors end in liquidation. Informed depositors are able to exercise their withdrawal option before uninformed depositors are able to observe the bank's difficulty (or the run).

[10] This point is emphasized by Diamond (1984) and Bernanke and Gertler (1987). Diamond's solution to the delegated-monitoring problem of financial intermediation relies on two assumptions that are absent in our framework: the existence of an ex post non-pecuniary penalty that can be imposed on the banker, and the ability of the banker to construct a riskless portfolio through diversification. The second assumption permits enforcement of the penalty even if cheating is costly to observe directly, whenever the banker fails to meet his obligations. Townsend (1979) notes that in circumstances where only one party has access to information, debt contracts - i.e., contracts not contingent on the private information will often be the only feasible alternative. For an overview of the relation between agency costs and the structure of financial contracts see Fama (1988).

[11]In our framework, the structure is sufficently simple that it is impossible for institutions to arise in which reputation solves the incentive compatibility problem. In our view, this is the correct first step. The viability of a reputation equilibrium depends crucially on the comparison of the current gain from cheating and the prospective future loss of reputation capital. When the reputation guarantees against, for example, the production of inferior quality goods, it seems reasonable to predict that reputations may be effective, since the gains from cheating are comparatively small. Bankers, however, receive a large quantity of mobile liquid wealth from depositors in advance of beginning loan operations. The temptations to a bank manager from absconding with the gold may be great, making it particularly difficult to maintain a reputation equilibrium. The peculiar vulnerability of financial institutions to damage of reputations is consistent with the concern expressed by bankers to maintain reputations for an extraordinary degree of probity and their emphasis on concomitant signals -- ornate, expensive edifices and the like.

[12]Thus, this framework captures an important feature of historical banking: Some depositors (including, for example, other banks) monitored banks' prospects and activities closely, while others deposited funds without investing resources in following the fortunes of the bank.

[13]See Jacklin and Bhattacharya (1988) for a distinction beteen information based runs and pure panics.

[14]Their model also provides for exogenous liquidity demands and for the possibility of inferring others' information by watching the length of lines at the banks.

[15] In our model, informed depositors are allowed to benefit from exercising a put option based on the information they receive. They can opt to convert their debt into goods and receive a higher payoff than the uninformed depositors. However, unlike the usual "inside-trading" scenario, the uninformed depositors also benefit at the expense of the bank. While the uninformed depositors receive a lower payoff than the informed depositors, they benefit because the bank is prevented from cheating. In the usual scenario (for example, Kyle, 1981), the uninformed either lose or the informed cannot successfully earn a return on their information production because of free riding, as in Grossman-Stiglitz (1980). We thank an anonymous referee for suggesting this comparison to us.

[16] See Villamil (1988), Jacklin (1988), Williamson (1988), and Gorton and Pennacchi (1988) for various approaches to combining the liquidity and incentive arguments.

[17] The concentration of bank fraud during times of regional or national economic decline is pronounced in national bank failure data. See the Annual Report of the U.S. Comptroller of the Currency, 1920, pp. 56-79.

[18] There are several ways we can approach the question of the maximum to be paid once the court has control. For simplicity we assume M does not very with the realization of T_i. One argument is that the value of the firm might be determined by the court, but at a very high cost. In this case, the best the court can do is to give the depositor the minimum value in the support of the distribution. In this case $M = (1-L)T_1$. Another approach would be to assume that the court physically hands the assets over to the depositor -- who then receives whatever they are worth. In this case, M would be the expected value of the assets conditional on any signal the depositor has drawn.

[19] Actual liquidation costs in the United States varied historically, depending on time, location and bank size, but seem to have been small relative to potential social losses from absconding, as our model assumes. Bankruptcy expenses averaged between three and six percent of total collections for national banks between 1872 and 1904 (Gendreau and Prince, 1986).

[20] In the single depositor case, the assumption that the signal takes only two values is <u>not</u> restrictive. In fact, the multi depositor model of the subsequent section can be reinterpreted as a single depositor model with multivalued signals.

[21] As it stands, the specification of the contract is incomplete in two technical respects. First the specification of the outcome should include a specification of the banker's response -- i.e. whether he chooses to abscond -- as a function of the announcement $\hat{\sigma}$ and of the realization T_i. However in almost all contracts the banker's response is easily discerned: He absconds if $P_{\hat{\sigma}} > AT_i$ and does not abscond if $P_{\hat{\sigma}} < AT_i$. Only in the case of indifference would it be necessary to specify his response in detail.

Secondly the contract does not include the possibility of randomized outcomes. These can be shown never to dominate deterministic outcomes.

[22] This assumption is natural given that we regard the project as requiring the banker's expertise, but the reserves as invested in publicly available technologies.

[23] An alternative assumption is that if the banker absconds, he takes the entirety of the reserves as well. The assumption in the text is natural if we regard absconding as occuring by siphoning a project into a less desirable project whose returns accrue directly to the banker. The assumption in this footnote is natural if we regard absconding as occuring when the banker piles the loot into the stagecoach and heads out of town.

[24] This is the simplest structure of supply of signals; it can be generalized. Alternatively, the cost of investing in a signal can be determined in a general equilibrium model. Since these extensions are beyond the scope of our interests we omit them.

[25] It will be clear that as long as the cost of investing in the signal is sufficiently low, it is optimal to have all individuals with cost I make the investment.

[26] If exactly $\underline{N}$ announce bad signals, the optimal contract has a randomization between liquidation and non liquidation. We omit the details.

The appendix shows that for values of S below this range, it will be useful to have <u>two</u> thresholds rather than one. For a range of values of bad signals received, it will be optimal to reduce the promised payment rather than liquidate the bank. This is analogous to the nuisance contract discussed before, and as before, it can be precluded by sufficiently high reservation levels of return.

[27] In a richer model, one could imagine banks choosing between holding reserves and investing more in higher earning projects. In this model, reserves are used exclusively for redistributing payouts between monitors and non-monitors in an incentive compatible way.

[28]The constraints initially have two equalities that must be satisfied; however given the fact that the total expected payments equal SZ + KI as they do by construction of the demandable debt contract, one of the equations is redundant: if the informed depositors are each receiving S+I, then the uninformed are automatically receiving the remainder, or S per depositor

[29]Historically, observed specie prices of bank notes published in bank note "reporters" confirm the view that non-monitors faced little price uncertainty for bank notes of banks that were open. Discounts on ante-bellum bank notes convertible on demand into specie traded in the home city at par; in distant locations the discounts for currencies mainly refelected the risk due to the time it would take to reach the city of issue. Typically, one could know the value of a bank's notes in New York by knowing the state in which the bank was located. These discounts typically remained small (between 1/8 percent and 2 percent) and were not subject to much variation. For failed banks, bank note discounts either were not quoted in bank note reporters or were subject to extreme variations across banks in the same locale and over time. See Calomiris and Schweikart (1988).

[30]See Akerlof (1970) and Klein (1974).

[31] For a summary of the events of the panic of 33 A.D. see A. W. Ferrin, "The Business Panic of A.D. 33," Moody's Magazine (August 1908), pp. 81-2.

[32]See Calomiris and Schweikart (1988) and Cannon (1910).

[33]See Gorton (1989) and Calomiris and Kahn (1988).

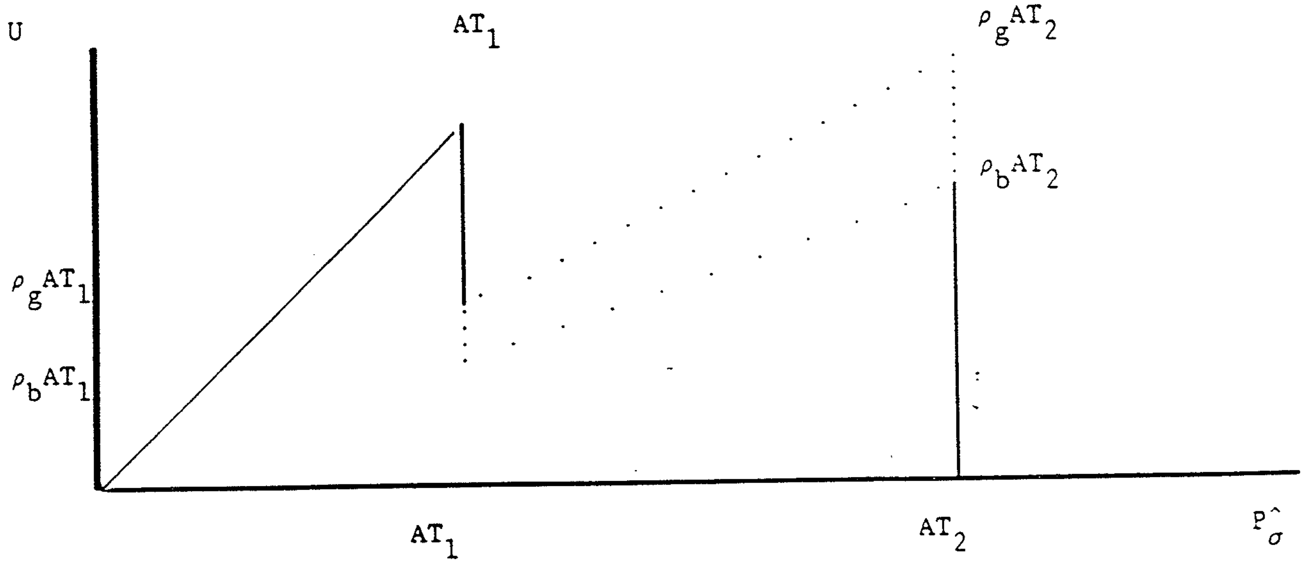

Figure A.1