JOURNAL OF FINANCIAL INTERMEDIATION **7,** 198–218 (1998)
ARTICLE NO. JF980236

Deposit Liquidity and Bank Monitoring

Jianping Qi*

Finance Department, University of South Florida, BSN 3403, Tampa, Florida 33620-5500
E-mail: jqi@coba.usf.edu

Received April 24, 1997

Why do banks fund loans embodying considerable borrower-specific information with liquid deposits? I address this question by examining the disciplinary effect of liquid deposits in a framework where banks' production of nontransferable borrower information is explicitly considered. I show that deposit liquidity motivates banks to provide greater monitoring of their loan applicants despite the general lack of observability of bank monitoring and bank loan quality. The analysis provides an important link between the two activities in which banks are viewed as "special"—their liquidity provision through demand deposits and their lending to information-intensive borrowers. *Journal of Economic Literature* Classification Numbers: G21, G28.

1. INTRODUCTION

The literature views banks as "special" intermediaries because of their provision of demand deposits. The role of such bank liabilities in the conduct of monetary policy has long been the subject of research. More recently, studies have focused on how such deposit liabilities relate to banks' provision of liquidity and to the likelihood of bank runs.[1]

* I thank Mark Carey, Joe Haubrich, George Kanatas, Raghu Rajan (the processing editor), Anjan Thakor (the managing editor), Greg Udell, Andy Winton, an anonymous referee, and seminar participants at the Federal Reserve Bank of Cleveland and the University of Houston for helpful comments and suggestions. An earlier version of this work was supported, in part, by the University of South Florida Research and Creative Scholarship Grant Program.

[1] See Niehans (1978) for an insightful discussion of bank provision of liquidity through demand deposits. The issues of bank liquidity and stability are examined in, for example, Bryant (1980), Diamond and Dybvig (1983), Gorton (1985), Gorton and Pennacchi (1990), Jacklin (1987), Jacklin and Bhattacharya (1988), Qi (1994, 1996), and Wallace (1988, 1990).

1042-9573/98 $25.00

A different stream of research has emphasized that lending makes banks "special." Fama (1985) conjectures that banks provide uniquely valuable lending services to offset the regulatory costs they initially incur and subsequently pass along to their borrowers. James (1987) provides supporting evidence. This literature attributes banks' uniqueness in lending to their efficiency in screening and/or monitoring of borrowers whose credit risk is not readily observable (see Diamond (1984) and Ramakrishnan and Thakor (1984)).[2] Thus, bank loans are very information-intensive, and this information is not easily transferable.

The extensive work within these two parallel strands in the literature has improved our understanding of the role of banks in liquidity provision and in the production of borrower information (see Bhattacharya and Thakor (1993) for a comprehensive survey). A seemingly unresolved issue concerns the link between the bank's use of liquid deposits and its funding of information-intensive loans. Why do banks fund illiquid loans with liquid deposits? Flannery (1994) suggests that because of high leverage, bank stockholders confront unusually severe asset-substitution moral hazard, and as a result, they must employ special devices to control this incentive problem. By assuming that bank asset quality is observable but not contractable, Flannery argues that liquid deposits are valuable contracting devices since changes in bank quality will be promptly reflected in financing costs. The assumption of observable bank asset quality, however, is at odds with the general view about the illiquidity of bank loans due to the bank's private information about borrowers, something Flannery (1994) recognizes as well. Thus, while Flannery's explanation is plausible that liquid deposits commit the bank to the discipline of the financial markets, the robustness of this explanation is debatable.

In this paper, I study the disciplinary effect of liquid deposits in a framework where private information production ("monitoring") by banks is explicitly considered. In my model, bank asset quality is neither observable nor verifiable, and the optimal level of bank monitoring is part of the banking equilibrium. I show that by issuing liquid deposits, banks credibly commit themselves to provide greater monitoring of their loan applicants despite the general lack of observability of bank monitoring and bank loan quality. Thus, my analysis goes beyond augmenting Flannery's (1994) basic argument of the disciplinary effect of liquid deposits by providing a formal link between deposit liquidity and bank monitoring, the two features of banking that are viewed as being special.[3] The analysis will be shown to

[2] Work that examines intermediaries' role in information production includes Boyd and Prescott (1986), Campbell and Kracaw (1980), and Winton (1995).

[3] Carey *et al.* (1996) present empirical evidence that supports a link between financial intermediaries' use of liquid liabilities and their monitoring of borrowers.

have implications for policy that addresses bank stability and for the narrow-banking proposal that calls for functional separation of banks' liability and asset services.

The need to motivate bank monitoring of credit applicants can be understood from the incentive conflict between the bank and its customers. As an intermediary, the bank provides monitoring and financing services on behalf of its depositors. With costly and unobservable monitoring, the bank has the incentive to shirk its responsibility, thereby imposing a cost on its customers by lowering its overall loan quality. Possible remedies such as portfolio diversification, compensation contracts, and capital adequacy requirements notwithstanding, my analysis suggests a simple, perhaps complementary, mechanism that utilizes the liquidity of deposits to motivate bank monitoring. This alternative is especially useful when other mechanisms are shown to be infeasible or inadequate.

The analysis proceeds in a general equilibrium framework. There are agents with pure saving need (''pure savers'') and potential borrowing need (''potential borrowers''). The project type of a potential borrower is privately revealed to be either (socially) ''good'' or ''bad.'' Outside investors must engage in costly monitoring to assess a project's type. As in Diamond (1984), with a large number of savers and borrowers, the monitoring and financing responsibilities are delegated to the ''bank.'' The key issue is the design of a mechanism that motivates costly bank monitoring.[4] This mechanism is shown to have the bank issue liquid deposits to savers as well as to potential borrowers, with the promise of funding the latter's projects if they are identified as good.

An important feature of my model is that the bank's credit applicants are also its depositors, which is common in practice. With less monitoring, the bank will erroneously reject more good loan applicants. Rejected good applicants, however, will withdraw their deposits in the interim because they have other uses for these funds that are more productive than bank deposits. A large number of such withdrawals will then convey negative information about bank monitoring, which can cause other depositors to do the same. While deposit liquidity makes early withdrawals not costly to the withdrawing depositors, these withdrawals are costly to the bank because they can cause it to fail. This threat of bank failure provides the incentive for bank monitoring.

My analysis relates to that of Calomiris and Kahn (1991), who show that demand deposits and the corresponding ''sequential service constraint''

[4] In Diamond (1984), the threat of a nonpecuniary penalty motivates bank monitoring. Besanko and Kanatas (1993) show that when banks cannot commit to a particular level of monitoring, there is a unique credit market equilibrium in which firms are financed with a combination of bank credit and external capital. In Rajan and Winton (1995), lender monitoring is motivated by debt covenants and collaterals.

motivate depositors to monitor their bank by allowing them to abandon the bank at the first sign of trouble. The idea that deposit liquidity serves to keep the bank's portfolio choice in line with depositors' preferences is in Calomiris *et al.* (1991). My focus is instead on how to motivate the bank to monitor its borrowers. Nakamura (1993) argues that the joint provision of loans and deposits makes bank lending special by allowing the bank to learn from deposits about its borrowers. Here, in contrast, it is depositors who learn about bank quality from the bank's lending decisions.

The paper is organized as follows. Section 2 establishes the basic incentive link between deposit liquidity and bank monitoring. Sections 3 and 4 extend the analysis to unobservable monitoring and possibly unknown loan quantity. Section 5 examines the feasibility of using compensation contracts to motivate monitoring. Section 6 concludes the paper and discusses its implications. Mathematical proofs are contained in an Appendix.

2. DEPOSIT LIQUIDITY AS MONITORING INCENTIVE

2.1. *The Setup*

Consider a risk-neutral economy with a single perfectly storable consumption good and three dates, $t = 0, 1, 2$. There are three categories of agents ex ante; all of them value consumption at $t = 2$. The first is a continuum of "savers" with a measure of $\alpha > 0$. Each saver is endowed with one unit of the good at $t = 0$ and nothing at any other time. The second is a continuum of "potential borrowers" with a measure normalized to one. Each potential borrower is endowed at $t = 0$ with one unit of the good and a project of initially uncertain quality. The project will become either type-g or type-b, with the realized type only privately revealed at $t = 1$. The ex ante probabilities of a project being type-g and type-b are p and $1 - p$, respectively. Anticipating the need of monitoring, I introduce a third category of agents who, endowed with nothing, are simply to offer competitive monitoring services. In anticipation of the banking equilibrium, the monitoring agent is referred to as the "banker," and the term "bank" refers to the institution set up and controlled by the banker. The banker will be assumed to enjoy private rents from being in control, and the essential problem is the design of a banking institution that motivates monitoring by the banker.

A potential borrower's project can be financed with either one unit or $I > 1$ units of the good at $t = 1$. With the greater funding level I, the per unit payoff of a type-g project at $t = 2$ is R_g with probability ϕ_g and 0 with probability $1 - \phi_g$. With the lower investment of one unit, however, a type-g project's per unit payoff is only $S_g < R_g$ with probability ϕ_g and 0 otherwise. I assume that either level of investment in a type-g project is socially good, but the greater investment is more productive, i.e., $\overline{R}_g \equiv \phi_g R_g > \phi_g S_g > 1$. As will be seen, this assumption ensures that the owner of a type-g project will seek the greater (bank) funding for his project, but if such funding is not provided, he will invest his own funds in it.[5] Unlike that of a type-g project, the per unit payoff of a type-b project is R_b with probability ϕ_b and 0 with probability $1 - \phi_b$, independent of its initial investment level being one or I units. I assume that either amount of investment in a type-b project is socially bad, and in particular, $R_b > R_g$, $\phi_g > \phi_b$, and $\overline{R}_b \equiv \phi_b R_b < 1$.[6] Further, aggregate resource is large enough to permit full funding of all projects; i.e., $I < 1 + \alpha$.

To ensure that external financing takes place with monitoring, I assume

$$p < \frac{1 - \overline{R}_b}{\overline{R}_g - \overline{R}_b}, \tag{1}$$

$$I > \frac{(1 - \phi_b)\overline{R}_g - (1 - \phi_g)\overline{R}_b}{\phi_g \overline{R}_b - \phi_b \overline{R}_g}. \tag{2}$$

By (1), investment in a randomly chosen project is socially undesirable since it implies $p\overline{R}_g + (1 - p)\overline{R}_b < 1$. Condition (2) ensures that there is no separating equilibrium with competitive external financing. Were such an equilibrium to exist, there would be an interest factor (one plus the rate of return), say, $\gamma > 0$, such that type-g potential borrowers would choose external project funding while type-b would not. That is,

$$\phi_g(R_g I - (I - 1)\gamma) > \gamma > \phi_b(R_b I - (I - 1)\gamma).$$

By rearranging terms, the above would require

$$\frac{\overline{R}_g I}{1 + \phi_g(I - 1)} > \gamma > \frac{\overline{R}_b I}{1 + \phi_b(I - 1)}.$$

But this is impossible since (2) implies

$$\overline{R}_b(1 + \phi_g(I - 1)) > \overline{R}_g(1 + \phi_b(I - 1)).$$

[5] This assumption reflects the view that good loan applicants have other uses for their funds that are more productive than bank deposits. The important implication is that rejected type-g applicants will make early deposit withdrawals.

[6] My analysis would be unchanged were the return of type-b projects to differ in the amount of investment, so long as such investment is socially bad.

In my model, costly project monitoring takes place prior to its type being privately revealed at $t = 1$.[7] A banker incurs a nonpecuniary cost of $m > 0$ for every project she monitors. Given monitoring, the banker will be able to unbiasedly rate the project $i \in \{g, b\}$, with the rating relating to its type as follows:

$$\Pr(i = g \mid \text{a type-g project}) = \pi_g,$$
$$\Pr(i = g \mid \text{a type-b project}) = \pi_b,$$

where $0 < \pi_b < \pi_g < 1$. Thus, the rating contains both types of errors, but the probability of a type-g project being (correctly) rated $i = g$, π_g, is greater than that of a type-b being (incorrectly) rated $i = g$, π_b. I assume that m is sufficiently small so that it is socially desirable for the banker to monitor and subsequently finance projects with the $i = g$ rating; i.e.,

$$m < (p\pi_g \overline{R}_g + (1 - p)\pi_b \overline{R}_b - n^*)I,$$

where $n^* \equiv p\pi_g + (1 - p)\pi_b$ is the probability of a random project being rated $i = g$.

2.2. *The Banking Mechanism*

The key issue is how to motivate proper monitoring and financing by the banker. Consider the formation of a financial intermediary (bank) by the banker. To attract depositors, the banker must convince them that she will carry out monitoring and finance only projects with the $i = g$ rating. The mechanism that will make this commitment credible is the bank's issuance at $t = 0$ of liquid deposits to savers and to potential borrowers, with the proceeds used to finance projects at $t = 1$. By "liquid deposits" I mean that interim withdrawals are not costly to the withdrawing depositors and the bank must honor such withdrawals even when this necessitates a premature liquidation of bank assets.[8] Figure 1 summarizes the sequence of events.

[7] The assumption of ex ante monitoring avoids complications that could arise from the use of prepaid monitoring costs to screen out low quality borrowers. In general, information generated from ex ante monitoring could help loan applicants understand better market conditions and their investment prospects. Similar assumptions appear in Boyd and Prescott (1986), Campbell and Kracaw (1980), and Ramakrishnan and Thakor (1984).

[8] This is essentially the sequential service constraint, or the "first-come, first-served" rule, associated with bank demand deposits. Although the deposit contracts are interpreted as demand deposits, it should be noted that in my discrete-time model with two deposit-withdrawable dates, one cannot meaningfully distinguish between demand deposits and one-period, renewable debt.

$t = 0$	Bank issues deposits
	Agents invest in deposits
	Potential borrowers seek loans
	Banker monitors projects
$t = 1$	Bank announces loan decisions
	Potential borrowers learn types
	Rejected type-gs withdraw deposits
	Project investments are made
$t = 2$	Project returns are realized
	Loans and deposits are repaid
	Banker obtains control rent

FIG. 1. Sequence of events.

Let d_1 denote the per unit payoff to deposits withdrawn at $t = 1$. If deposits remain, they will be paid in proportion to the $t = 2$ bank asset value subtracted by any contractual payments to the banker; let d_2 denote the expected per unit payoff to deposits at $t = 2$. Given the storage alternative, the deposit payoff $\{d_1, d_2\}$ must satisfy

$$\max\{d_1, d_2\} \geq 1. \tag{3}$$

The payoff must also ensure that depositors with only saving need will not prefer withdrawing to maintaining deposits at $t = 1$. That is,

$$d_1 \leq d_2. \tag{4}$$

For the moment, suppose that potential borrowers desire bank funding of projects (the conditions for this will be derived shortly). As will become clear, they will then make deposits and apply for project loans at $t = 0$. The bank promises to extend a loan of $I - d_1$ at $t = 1$ if a potential borrower's project is rated $i = \mathrm{g}$. Note that this loan amount plus that of interim deposit withdrawal, d_1, will enable the greater funding level I for the borrower's project.[9] It is not essential, however, that the borrower withdraw his deposit to supplement the bank loan at $t = 1$. The results would be unchanged were the bank to loan I and the borrower to maintain his deposit.

What is important is the deposit withdrawal decision of type-g potential borrowers who are (incorrectly) denied loans. Rejected type-g applicants

[9] In fact, type-g borrowers desire to redeem their deposits at the face value d_1 and borrow $I - d_1$ additionally to finance their projects whenever the loan interest factor r and the deposit payoff $\{d_1, d_2\}$ satisfy $d_1 \phi_{\mathrm{g}} r > d_2$. This condition is satisfied in my model.

will make interim withdrawals if they can invest their funds at a return higher than that on bank deposits. That is,

$$d_1 + \phi_g S_g - 1 > d_2. \tag{5}$$

In contrast, (correctly) rejected type-b potential borrowers will not make early withdrawals because they do not want to invest solely their own funds in their projects. Note that the bank cannot base its loan decisions on whether potential borrowers make interim withdrawals, which would induce rejected type-b applicants to do the same.

Type-g borrowers will indeed desire bank funding of projects if the loan interest factor r and the deposit payoff $\{d_1, d_2\}$ satisfy

$$\phi_g(R_g I - (I - d_1)r) > \max\{d_1, d_2, d_1 + \phi_g S_g - 1\}. \tag{6}$$

I assume that (incorrectly) accepted type-b applicants will also want to take bank loans and invest in their projects. This is the case if

$$\phi_b(R_b I - (I - d_1)r) > \max\{d_1, d_2\}.$$

Given (2), the above condition is satisfied for $\overline{R}_g$ sufficiently large.

To focus solely on the incentive effect of liquid deposits without complications of compensation contracts, I assume that the banker values being in control of a "successful" bank—one that is not prematurely liquidated at $t = 1$. Her private control rent, however, will be lost if the bank fails in the interim.[10] Further, the size of her control rent is assumed to be greater than her monitoring cost, making a direct compensation unnecessary. In Section 5, I will relax these assumptions. The results will be shown to be robust even when compensation contracts are explicitly considered.

Letting $\theta \in [0,1]$ denote the number of projects (out of measure 1) the banker monitors at $t = 0$, I now examine her choice of the (aggregate) monitoring level. Given θ, θn^* of monitored projects will be rated $i = \mathrm{g}$; recall that $n^* = p\pi_g + (1 - p)\pi_b$. If the number of project loans is $n \leq \theta n^*$, the bank will finance only among those with the $i = \mathrm{g}$ rating. Its expected payoff from these n loans is

$$L(\theta, n) = n\,\frac{p\pi_g\phi_g + (1 - p)\pi_b\phi_b}{n^*}(I - d_1)r \tag{7}$$

[10] An interpretation of the control rent is reputation capital associated with a successful bank, which would be lost were it to fail.

Note that $p\pi_g/n^*$ and $(1 - p)\pi_b/n^*$ are the posterior probabilities that a project is type-g and type-b, respectively, given the i = g rating. If the number of loans is $n > \theta n^*$, on the other hand, the bank will finance not only θn^* projects with the i = g rating, but also $n - \theta n^*$ projects that are either not monitored (and not rated) or rated i = b. Since the number of projects with the i = g or i = b rating is deterministic for a given level of monitoring, the bank has no incentive to finance projects with the i = b rating. A strategy that lowers the monitoring level while increasing lending to nonmonitored projects dominates funding projects with the i = b rating. Thus, with $n > \theta n^*$, the additional $n - \theta n^*$ projects will be drawn from those not monitored. The bank's expected loan payoff is now

$$\begin{aligned} L(\theta, n) = (&\theta(p\pi_g\phi_g + (1-p)\pi_b\phi_b) \\ &+ (n - \theta n^*)(p\phi_g + (1-p)\phi_b))(I - d_1)r, \end{aligned} \tag{8}$$

where $\theta(p\pi_g\phi_g + (1 - p)\pi_b\phi_b)(I - d_1)r$ is, as in (7), the expected payoff from lending to θn^* projects with the i = g rating, while $(n - \theta n^*)(p\phi_g + (1 - p)\phi_b)(I - d_1)r$ is that from lending to $n - \theta n^*$ non-monitored projects. In lending to the latter, the proportions being type-g and type-b are their respective prior probabilities p and $1 - p$.

Since the bank commits to honor interim deposit withdrawals whenever possible, such withdrawals take priority over loan takedowns. Even if loans were taken, the bank would call back them to satisfy the withdrawal need. Thus, the number of project loans, n, that the bank makes (and are not called back) is limited by its interim liquidity constraint,

$$nI \leq \max\{0, 1 + \alpha - wd_1\},$$

where w denotes the number of interim deposit withdrawals made by nonborrowers, i.e., savers and rejected potential borrowers.

I now quantify the expected deposit payoff at $t = 2$. If the number of early withdrawals is $w < (1 + \alpha)/d_1$, $n \leq (1 + \alpha - wd_1)/I$. Then, given θ and n, the expected payoff per unit deposit at $t = 2$ is

$$d_2(\theta, n, w) = \frac{1 + \alpha - wd_1 - nI + L(\theta, n)}{1 + \alpha - n - w}.$$

If the number of interim withdrawals has reached $(1 + \alpha)/d_1$, however, the bank will run out of all resource and will be closed at $t = 1$. In this case, no project loan is extended and deposits that have not been withdrawn prior to bank closure will receive nothing. That is, if $w \geq (1 + \alpha)/d_1$, $n = 0$ and $d_2(\theta, n, w) = 0$.

Let $w^* \equiv p(1 - \pi_g)$ denote the number of rejected type-g applicants because of the rating error. Proposition 1 establishes the conditions that ensure that the $t = 2$ deposit payoff $d_2(\theta, n, w)$ is maximized when the banker carries out full monitoring, i.e., $\theta = 1$, and finances only projects with the $i =$ g rating, i.e., $n = n^*$.

PROPOSITION 1. *Suppose the deposit payoff d_1 and the loan factor r satisfy*

$$d_1 = d_2(1, n^*, w^*), \textit{ and} \tag{9}$$

$$\frac{n^*(I-1)}{p\pi_g\phi_g + (1-p)\pi_b\phi_b} < (I - d_1)r$$
$$< \frac{(1+\alpha)(I-1)}{(1+\alpha)(p\phi_g + (1-p)\phi_b) + p(1-p)(\pi_g - \pi_b)(\phi_g - \phi_b)}. \tag{10}$$

Then $d_2(1, n^*, w^*) \geq d_2(\theta, n, w)$, and the inequality is strict if $(\theta, n) \neq (1, n^*)$.

The first part of (10) ensures that the loan payoff is sufficiently large that the bank at least breaks even on lending to projects with the $i =$ g rating. The second part, on the other hand, limits the loan payoff not to be so large that depositors no longer have the incentive to withdraw early even when inadequate monitoring, i.e., $\theta < 1$, is known. Then, with (9), if inadequate monitoring is known, depositors will make early withdrawals since $d_2(\theta, n, w) < d_1$ for $\theta < 1$. Interestingly, $d_1 = d_2(1, n^*, w^*)$ ensures that the deposit payoff is indeed liquid—early withdrawals are not costly to the withdrawing depositors.

The above analysis applies when the monitoring level θ is observable. With inadequate monitoring, depositors will withdraw early, thereby liquidating the bank and causing the loss of the banker's control rent. With full monitoring, however, recognizing that the bank has no incentive to deviate from lending to n^* projects with the $i =$ g rating, depositors with saving need will be unconcerned about deposit safety and consequently will not withdraw at $t = 1$. Only rejected type-g applicants will withdraw deposits at this time for investment in their projects. Proposition 2 establishes the conditions for the viability of the banking mechanism, given observable monitoring.

PROPOSITION 2. *Let $d_1 = 1 + \varepsilon$ and*

$$(I - d_1)r = \frac{n^*(I-1) + (1+\alpha-n^*)\varepsilon}{p\pi_g\phi_g + (1-p)\pi_b\phi_b}.$$

Then $d_2(1, n^*, w^*) = d_1$, *and for a sufficiently large* $\overline{R}_g$ *and a sufficiently small* $\varepsilon > 0$, *conditions (3) through (6) are satisfied. That is, the banking mechanism is viable.*

In Proposition 2, requiring a sufficiently large $\overline{R}_g$ ensures that type-g borrowers will desire bank funding of projects; that is, (6) is satisfied. Then, the banking mechanism is viable when the interim deposit payoff d_1 is "slightly" greater than the storage return, and the loan payoff $(I - d_1)r$ is "slightly" greater than the break-even point for lending to projects with the i = g rating, i.e., the first part of (10). Interestingly, by choosing ε arbitrarily close to 0, d_1 is arbitrarily close to 1, in which case, were the bank to be liquidated at $t = 1$, almost all deposits would be paid in full.

3. UNOBSERVABLE MONITORING

When the monitoring level θ is unobservable, the incentive link between deposit liquidity and monitoring no longer holds. In this section, I examine how the number of type-g potential borrowers who are rejected loans, and hence the number of their early deposit withdrawals, will provide an inference about the monitoring level and help restore the incentive link. In particular, given θ, θn^* projects will be rated i = g. If the number of project loans is $n \leq \theta n^*$, the bank will make loans only to n out of θn^* projects with the i = g rating. In this case, the number of rejected type-g applicants is

$$k(\theta, n) = p\left(1 - \frac{n}{n^*}\pi_g\right). \tag{11}$$

If the number of loans is $n > \theta n^*$, however, the bank will lend to θn^* projects with the i = g rating and to $n - \theta n^*$ projects that are not monitored. In the latter case, the number of rejected type-g applicants is

$$k(\theta, n) = p(1 - n - (1 - p)(\pi_g - \pi_b)\theta). \tag{12}$$

From (11) and (12), if the monitoring level is $\theta = 1$ and the number of loans is $n = n^*$, the number of misidentified type-g applicants will be $k(1, n^*) = p(1 - \pi_g) = w^*$, which is that resulting from the rating error.

Suppose for now the number of project loans can be limited to $n \leq n^*$ (I will show how this can be done in the next section). Proposition 3 shows how the number of rejected type-g potential borrowers relates to the levels of monitoring and financing.

PROPOSITION 3. *Suppose $n \leq n^*$. Then the number of rejected type-*g *potential borrowers is greater than that resulting from the rating error, i.e., $k(\theta, n) > w^*$, if and only if either the level of monitoring is inadequate, i.e., $\theta < 1$, or the number of project loans is insufficient, i.e., $n < n^*$.*

From Proposition 3, whenever $k(\theta, n) > w^*$ is revealed, depositors will be better off making early withdrawals since $d_1 > d_2(\theta, n, w)$ for $(\theta, n) \neq (1, n^*)$. In fact, recognizing that the bank has no incentive to deny loans to n^* projects with the i = g rating given full monitoring, the observation of $k(\theta, n) > w^*$ indicates inadequate monitoring.

In my model, the mechanism that reveals the number of rejected type-g applicants, $k(\theta, n)$, is their decision to make early deposit withdrawals. To the extent that such withdrawals are observable to other depositors, the deposit liquidity mechanism motivates full monitoring. Let $\Pr(\theta = 1|w)$ denote savers' and rejected type-b applicants' posterior beliefs about full monitoring given that the number of early withdrawals is w. Then, with

$$\Pr(\theta = 1|w) = \begin{cases} 1 & \text{if } w \leq w^*, \\ 0 & \text{if } w > w^*, \end{cases}$$

it is incentive compatible for the banker to carry out full monitoring and subsequently finance projects with the i = g rating. Indeed, given $n \leq n^*$, Proposition 3 shows that if the monitoring level is $\theta < 1$, the number of rejected type-g potential borrowers and hence that of their early withdrawals will be $k(\theta, n) > w^*$. Consequently, savers and rejected type-b applicants will infer inadequate monitoring and will also make early withdrawals. Since these withdrawals will result in the bank's liquidation and the loss of the banker's control rent, the threat of such withdrawals provides the monitoring incentive. Only with full monitoring will the number of rejected type-g applicants and that of their early withdrawals be $k(1, n^*) = w^*$ and will other depositors then maintain their deposits.

4. RESTRICTION ON LOAN QUANTITY

In Section 3, the number of project loans that the bank can make is limited to $n \leq n^*$. Clearly, this restriction is needed only when this number is unobservable. With an observable n, unless $n = n^*$, Propositions 1 and 2 imply that depositors would make early withdrawals. If n is unobservable, however, the number of rejected type-g potential borrowers being $k(\theta, n) = w^*$ need not indicate full monitoring. For any $\theta < 1$, the bank can reduce the number of rejected type-g applicants, $k(\theta, n)$, simply by increasing the number of project loans, n. Indeed, from (12), the bank ensures $k(\theta, n) = w^*$ even when $\theta < 1$ by choosing

$$n = \pi_g - (1 - p)(\pi_g - \pi_b)\theta > n^*.$$

One way to restore the information link between the number of rejected type-g applicants, $k(\theta, n)$, and the monitoring level θ is to mandate the disclosure of bank loan quantity. Such a disclosure requirement will prevent the bank from concealing inadequate monitoring by making an excessive number of loans with low credit standards. Evidently, it should be easier to verify bank loan quantity than quality.

An alternative way is to restrict the amount of bank loanable funds. In particular, to limit the number of loans to no more than n^*, the bank's loanable funds should be no more than $n^*(I - d_1)$. This can be achieved by imposing a storage (reserve) requirement. Let β denote the required reserve per unit deposit the bank holds. By setting

$$\beta = \frac{1 + \alpha - n^* I}{1 + \alpha - n^* d_1},$$

the amount of bank loanable funds is

$$F = (1 - \beta)(1 + \alpha - n d_1) = \frac{1 + \alpha - n d_1}{1 + \alpha - n^* d_1} n^*(I - d_1).$$

If the number of loans is $n \leq n^*$, $F \geq n^* \, (I - d_1) \geq n(I - d_1)$, and the amount of loanable funds is sufficient to finance these projects. But, if $n > n^*$, $F < n^* \, (I - d_1) < n(I - d_1)$, and the amount is no longer sufficient. Interestingly, the reserve requirement here is motivated by the need to limit the amount of bank loanable funds to prevent the banks from making excessive loans with low standards.

5. COMPENSATION CONTRACT AND MONITORING INCENTIVE

I have thus far assumed that the size of the banker's control rent is sufficiently large, so that there is no need for a direct compensation. This assumption is not essential. My analysis would remain unchanged if the banker were paid a fixed amount, independent of bank asset value at $t = 2$. A fixed compensation of course is restrictive itself. The important question is whether there exists some compensation contract that would motivate monitoring without utilizing deposit liquidity. In this section, I address this question.

Obviously, for a compensation contract to possibly motivate unobservable monitoring, the banker's payoff must be linked to bank asset value at $t = 2$. Given the level of monitoring, θ, the number of loans, n, and the number of interim withdrawals, w, the bank's $t = 2$ asset value is a random variable,

$$\tilde{Y} = 1 + \alpha - nI - wd_1 + \tilde{L}(\theta, n),$$

where $\tilde{L}(\theta, n)$ denotes the random payoff of bank loans, with its expected payoff being $L(\theta, n)$ given in Section 2. If the returns of both type-g and type-b projects are i.i.d. random variables, the law of large number implies perfect diversification. In this case, the bank's realized aggregate loan payoff, given θ and n, will equal its expected value $L(\theta, n)$. Then, if the banker is compensated with the $t = 2$ payoff function

$$\xi(Y) = \begin{cases} m & \text{if } Y \geq 1 + \alpha - n^*I - w^*d_1 + L(1, n^*), \\ 0 & \text{if } Y < 1 + \alpha - n^*I - w^*d_1 + L(1, n^*), \end{cases}$$

it is incentive compatible for her to engage in proper monitoring and financing, i.e., choosing $\theta = 1$ and $n = n^*$, without utilizing deposit liquidity.

In general, project returns are not perfectly diversifiable, and the aggregate loan payoff at $t = 2$ need not have a degenerate distribution. In the case where the probability density function $f(Y;\, \theta, n)$ of the bank's $t = 2$ asset value $\tilde{Y}$ is positive and continuous almost everywhere, I will show that there is no competitive compensation contract that would induce full monitoring by the banker. To see this, suppose instead that such a contract $\xi(Y)$ exists. Given competitive monitoring services, the banker expects to be compensated only for her monitoring cost m. That is,

$$\int_0^\infty \xi(Y) f(Y; 1, n^*)\, dY = m.$$

This implies that $\xi(Y) > 0$ for at least some $(a, b) \subset [0, \infty)$. By choosing full monitoring, i.e., $\theta = 1$, and $n = n^*$, the banker's net payoff would be

$$\int_0^\infty \xi(Y) f(Y; 1, n^*)\, dY - m = 0.$$

But, with no monitoring, i.e., $\theta = 0$, and $n = n^*$, her net payoff would instead be

$$\int_0^\infty \xi(Y)f(Y;0,n^*)dY > 0,$$

where the last inequality follows since $\xi(Y) > 0$ in (a, b) and the function $f(Y;\ \theta, n)$ is positive and continuous almost everywhere. In this case, the deposit liquidity mechanism is therefore needed to ensure proper monitoring. Although the above analysis does not consider the possibility of paying the banker more than the competitive rate, such a contract is less desirable to bank depositors and borrowers because it forces them to share with the banker the gains of financial intermediation.

6. CONCLUSION AND IMPLICATIONS

A theory of financial intermediation is developed in this paper that links deposit liquidity to bank monitoring. The use of liquid deposits is shown to motivate greater monitoring of credit applicants despite the lack of observability of bank monitoring. Monitoring reduces the number of good loan applicants who are erroneously denied loans. Since rejected good applicants will make interim deposit withdrawals, a large number of such withdrawals will convey negative information about the bank and can cause other depositors to do the same. While deposit liquidity makes early withdrawals not costly to the withdrawing depositors, these withdrawals are costly to the bank by threatening its very existence. Although my analysis focuses on the extreme threat of bank failure, it would be applicable if the cost of early withdrawals were in the form of a loss of bank reputation capital and/or a reduction of bank deposit base, as long as this was sufficiently costly.

In my model, it is the threat of bank runs that motivates bank monitoring. A similar threat mitigates the bank's moral hazard incentives in Calomiris and Kahn (1991) and in Flannery (1994). In all these three models, the information-based bank runs never materialize because the mere threat is sufficient to eliminate the incentive problems.

As in Diamond and Dybvig (1983), however, the bank here is vulnerable to "sunspot" runs even when its underlying assets are sound. Indeed, even with adequate bank monitoring, depositors who believe that others will make early withdrawals will have the incentive to do the same since $d_1 = d_2(1, n^*, w^*) > 1$ implies that the bank will run out of resource before all depositors can be paid off. Moreover, as in Chari and Jagannathan (1988), if the number of interim withdrawals is imperfectly observed, there is the possibility of "mistaken" bank runs. These types of bank runs are socially costly because promising projects that have been privately identified by the bank might be unable to secure alternative funding and might have to be liquidated.

A desirable policy should prevent sunspot and mistaken runs while preserving the incentive link between the threat of bank runs and bank monitoring. In view of this, full governmental deposit insurance is suboptimal despite its effectiveness in maintaining bank stability, as this would eliminate depositors' concerns for bank safety.[11] A better policy would be the creation of a regulatory agency that is empowered to suspend deposit withdrawals (convertibility) when they are abnormally large. Subsequent to the suspension, the agency may be able to investigate and determine the reasons for withdrawals—whether the soundness of the bank was an issue—and thereby stop sunspot and mistaken runs without compromising the incentive link between deposit liquidity and bank monitoring. Only when abnormal withdrawals resulted from a lack of bank monitoring and hence poor loan quality should the agency enforce bank closure and impose on it the liquidation cost. The bank obviously should not have the power to suspend the convertibility of deposits, which would allow it to short-circuit the incentive link and make deposits illiquid de facto.

My analysis has implications for the narrow banking proposal that calls for separation of banks' liability services (deposit-taking and payment system functions) from asset services (information production and loan origination). Underlying this proposal is the belief that there is no apparent reason for the bank to finance illiquid loans with liquid deposits. Thus, requiring that deposit-taking and check-clearing intermediaries hold only safe and liquid claims, such as Treasury securities, while lending intermediaries rely on less liquid sources of funds would eliminate depositors' concerns for the former intermediaries' asset quality and ensure the stability of the payment system. The problem is that if there is a link between the liquidity of intermediaries' sources of funds and their monitoring incentive, as suggested in this paper, inability to use liquid deposits as a funding source would adversely affect the lending intermediaries' motivation for information production. A consequence of this could be that these intermediaries would lower their loan quality, or that they would not originate more information-intensive loans, leaving projects with considerable private information unfunded.

An empirical implication of my analysis is that it suggests differences between loans funded by commercial banks that engage in both liquidity provision and loan origination and those funded by institutions such as finance companies that rely on less liquid sources of funds. If the liquidity of institutions' liabilities motivates greater monitoring, commercial banks would make less liquid loans which necessitate considerable information production while finance companies would focus on more homogeneous, more easily marketable loans. While banks originate a large quantity of

[11] Evidently, deposits are insured for up to $100,000 in the U.S.

commercial and industrial (C&I) loans, finance companies appear to concentrate relatively more on credit card loans, auto loans, and equipment lease financing.[12] The much greater extent to which the latter credit contracts have been securitized, compared to bank C&I loans, is not inconsistent with the general argument linking the liquidity of the intermediary's liabilities and its lending activity.

APPENDIX

Proof of Proposition 1. First, note that if $w \geq (1 + \alpha)/d_1$, then $d_2(\theta, n, w) = 0 < d_2(1, n^*, w^*)$. Thus, it suffices to focus on the alternative case: $w < (1 + \alpha)/d_1$. Recall that

$$d_1 = d_2(1, n^*, w^*) = \frac{1 + \alpha - n^*I + L(1, n^*)}{1 + \alpha - n^*}.$$

By substituting d_1 and $L(1, \theta)$,

$$\begin{aligned} &d_2(1, n^*, w^*) - d_2(\theta, n, w) \\ &= \frac{(1 + \alpha - n)(1 + \alpha - n^*I + L(1, n^*)) - (1 + \alpha - n^*)(1 + \alpha - nI + L(\theta, n))}{(1 + \alpha - n^*)(1 + \alpha - n - w)}. \end{aligned}$$

Thus, showing that $d_2(1, n^*, w^*) \geq d_2(\theta, n, w)$ and in particular $d_2(1, n^*, w^*) > d_2(\theta, n, w)$ for $(\theta, n) \neq (1, n^*)$ is equivalent to showing that for $(\theta, n) \neq (1, n^*)$,

$$\frac{1 + \alpha - nI + L(\theta, n)}{1 + \alpha - n} < \frac{1 + \alpha - n^*I + L(1, n^*)}{1 + \alpha - n^*}.$$

To do so, define a function

$$h(\theta, n) = \frac{1 + \alpha - nI + L(\theta, n)}{1 + \alpha - n}.$$

[12] Carey *et al.* (1996) found that C&I loans were more significant for banks than for finance companies. They also reported that based on observable risk attributes, borrowers of bank C&I loans appeared to be safer than those of finance company C&I loans. This evidence, however, is not necessarily inconsistent with the implications of my model because information production by intermediaries is to uncover quality attributes that are not readily observable.

Consider first the case with $n \leq \theta n^*$; here, $L(\theta, n)$ is given in (7) and

$$h(\theta, n) = \frac{1 + \alpha - nI + n(n^*)^{-1}(p\pi_g\phi_g + (1-p)\pi_b\phi_b)(I - d_1)r}{1 + \alpha - n}.$$

The partial derivative of $h(\theta, n)$ with respect to n is

$$h_2(\theta, n) = \frac{(1+\alpha)((n^*)^{-1}(p\pi_g\phi_g + (1-p)\pi_b\phi_b)(I - d_1)r - I + 1)}{(1 + a - n)^2}.$$

Then, if

$$(I - d_1)r > \frac{n^*(I-1)}{p\pi_g\phi_g + (1-p)\pi_b\phi_b},$$

$h_2(\theta, n) > 0$. Thus, given θ, $h(\theta, n)$ is increasing in n for $n \leq \theta n^*$, and $n = \theta n^*$ maximizes the function $h(\theta, n)$ for $n < \theta n^*$.

In the alternative case with $n \geq \theta n^*$, $L(\theta, n)$ is given in (8). Then,

$$h(\theta, n) = \frac{\begin{array}{c}1 + \alpha - nI + (\theta(p\pi_g\phi_g + (1-p)\pi_b\phi_b)\\ + (n - \theta n^*)(p\phi_g + (1-p)\phi_b))(I - d_1)r\end{array}}{1 + \alpha - n},$$

$$h_2(\theta, n) = -\frac{(1+\alpha)(I-1)}{(1+\alpha-n)^2} + \frac{\begin{array}{c}(\theta(p\pi_g\phi_g + (1-p)\pi_b\phi_b)\\ + (1 + \alpha - \theta n^*)(p\phi_g + (1-p)\phi_b))(I - d_1)r\end{array}}{(1+\alpha-n)^2}.$$

Thus, $h_2(\theta, n) < 0$ if

$$\begin{aligned}(I - d_1)r &< \frac{(1+\alpha)(I-1)}{(1+\alpha)(p\pi_g\phi_g + (1-p)\pi_b\phi_b) + p(1-p)(\pi_g - \pi_b)(\phi_g - \phi_b)}\\ &\leq \frac{(1+\alpha)(I-1)}{(1+\alpha)(p\pi_g\phi_g + (1-p)\pi_b\phi_b) + \theta p(1-p)(\pi_g - \pi_b)(\phi_g - \phi_b)}.\end{aligned}$$

That is, given θ, $h(\theta, n)$ is decreasing in $n \geq \theta n^*$, and $n = \theta n^*$ also maximizes the function $h(\theta, n)$ for $n > \theta n^*$.

Since $h(\theta, \theta n^*) > h(\theta, n)$ for $n \neq n^*$, to ensure that $h(1, n^*) > h(\theta, n)$ for $(\theta, n) \neq (1, n^*)$, what remains is to show $h(1, n^*) > h(\theta, \theta n^*)$ for $\theta < 1$. Consider the derivative of $h(\theta, \theta n^*)$ with respect to θ,

$$h'(\theta, \theta n^*) = \frac{(1+\alpha)((p\pi_g\phi_g + (1-p)\pi_b\phi_b)(I-d_1)r - n^*(I-1))}{(1+\alpha-\theta n^*)^2} > 0,$$

where the last inequality follows since

$$(I-d_1)r > \frac{n^*(I-1)}{p\pi_g\phi_g + (1-p)\pi_b\phi_b}.$$

Thus, $\theta = 1$ maximises $h(\theta, \theta n^*)$.

Therefore, given the restriction on $(I - d_1)r$, $h(1, n^*) > h(\theta, n)$ for $(\theta, n) \neq (1, n^*)$. It follows that $d_2(1, n^*, w^*) \geq d_2(\theta, n, w)$ for $(\theta, n, w) \neq (1, n^*, w^*)$, with the inequality being strict when $(\theta, n) \neq (1, n^*)$. To see that the set where d_1 and r satisfy the restriction on $(I - d_1)r$ is nonempty, note that

$$\frac{(1+\alpha)(I-1)}{(1+\alpha)(p\pi_g\phi_g + (1-p)\pi_b\phi_b) + p(1-p)(\pi_g-\pi_b)(\phi_g-\phi_b)} - \frac{n^*(I-1)}{p\pi_g\phi_g + (1-p)\pi_b\phi_b} > 0.$$

The last inequality follows since

$$\begin{aligned}&(1+\alpha)(p\pi_g\phi_g\pi_g + (1-p)\pi_b\phi_b)\\ &\quad - n^*((1+\alpha)(p\phi_g + (1-p)\phi_b) + p(1-p)(\pi_g-\pi_b)(\phi_g-\phi_b))\\ &= (1+\alpha-n^*)p(1-p)(\pi_g-\pi_g)(\phi_g-\phi_b) > 0.\end{aligned}$$

■

Proof of Proposition 2. Note that $1 + \alpha > I > n^*I + w^*$; thus, the bank has sufficient funds to finance all projects that are rated $i = \mathrm{g}$. Given $\varepsilon > 0$, $d_1 > 1$, and for $\varepsilon \to 0$, $(I - d_1)r$ satisfies condition (10). By substituting d_1 and $(I - d_1)r$ into $d_2(\theta, n, w)$, $d_2(1, n^*, w^*) = 1 + \varepsilon = d_1 > 1$. Thus, conditions (3), (4), and (5) are satisfied. To see that (6) is also satisfied, note that for $\overline{R}_g$ sufficiently large and ε sufficiently small,

$$\begin{aligned}\overline{R}_g I &> \max\left\{\phi_g S_g + \varepsilon + \phi_g(I-d_1)r, 1 + \varepsilon + \frac{\phi_g(n^*(I-1) + (1+\alpha-n^*)\varepsilon)}{p\pi_g\phi_g + (1-p)\pi_b\phi_b}\right\}\\ &= \max\{\phi_g S_g + \varepsilon + \phi_g(I-d_1)r, 1 + \varepsilon + \phi_g(I-d_1)r\},\end{aligned}$$

which implies (6) since $d_1 = d_2$ and

$$\phi_g(R_g I - (I - d_1)r) > \max\{\phi_g S_g + \varepsilon, 1 + \varepsilon\} = d_1 + \phi_g S_g - 1.$$

Therefore, the banking mechanism is viable. ■

Proof of Proposition 3. Consider first $n \leq \theta n^*$. From (11),

$$k(\theta, n) - w^* = p\pi_g\left(1 - \frac{n}{n^*}\right).$$

If $n < n^*$, clearly, $k(\theta, n) > w^*$. If $\theta < 1$, however, $n \leq \theta n^* < n^*$; thus, $k(\theta, n) > w^*$. Conversely, if $k(\theta, n) > w^*$, then $n < n^*$. For $n^* \geq n > \theta n^*$, from (12),

$$k(\theta, n) - w^* = p(\pi_g(1 - \theta) - (n - \theta n^*)).$$

If $\theta < 1$, then $\pi_g(1 - \theta) > n^*(1 - \theta) \geq n - \theta n^* > 0$; thus, $k(\theta, n) > w^*$. If $n < n^*$, it must be $\theta < 1$ to ensure $n > \theta n^*$; again, $k(\theta, n) > w^*$. Conversely, if $k(\theta, n) > w^*$, then $\pi_g(1 - \theta) > n - \theta n^* > 0$, which implies $\theta < 1$. ■

REFERENCES

Besanko, D., and Kanatas, G. (1993). Credit market equilibrium with bank monitoring and moral hazard, *Rev. Finan. Stud.* **6,** 213–232.

Bhattacharya, S., and Thakor, A. V. (1993). Contemporary banking theory, *J. Finan. Intermediation* **3,** 2–50.

Boyd, J. H., and Prescott, E. C. (1986). Financial intermediary coalitions, *J. Econ. Theory* **38,** 211–232.

Bryant, J. (1980). A model of reserves, bank runs, and deposit insurance, *J. Bank. Finan.* **4,** 335–344.

Calomiris, C. W., and Kahn, C. M. (1991). The role of demandable debt in structuring optimal banking arrangement, *Am. Econ. Rev.* **81,** 497–513.

Calomiris, C. W., Kahn, C. M., and Krasa, S. (1991). Optimal contingent bank liquidation under moral hazard, mimeo, University of Illinois.

Campbell, T. S., and Kracaw, W. A. (1980). Information production, market signalling, and the theory of financial intermediation, *J. Finance* **54,** 863–882.

Carey, M., Post, M., and Sharpe, S. (1996). Does lending by banks and finance companies differ? mimeo, Board of Governors of the Federal Reserve System.

Chari, V. V., and Jagannathan, R. (1988). Banking panics, information, and rational expectations equilibrium, *J. Finance* **43,** 749–761.

Diamond, D. W. (1984). Financial intermediation and delegated monitoring, *Rev. Econ. Stud.* **51,** 393–414.

Diamond, D. W., and Dybvig, P. H. (1983). Bank runs, deposit insurance, and liquidity, *J. Polit. Econ.* **91,** 401–419.

Fama, E. F. (1985). What's different about banks? *J. Monet. Econ.* **15,** 29–39.

Flannery, M. J. (1994). Debt maturity and the deadweight cost of leverage: Optimally financing banking firms, *Am. Econ. Rev.* **84,** 320–331.

Gorton, G. (1985). Bank suspension of convertibility, *J. Monet. Econ.* **15,** 177–194.

Gorton, G., and Pennacchi, G. (1990). Financial intermediaries and liquidity creation, *J. Finance* **45,** 49–71.

Jacklin, C. J. (1987). Demand deposits, trading restrictions, and risk sharing, *in* "Contractual Arrangements for Intertemporal Trade" (E. Prescott and N. Wallace, Eds.), pp. 26–47, Univ. of Minnesota Press, Minneapolis.

Jacklin, C. J., and Bhattacharya, S. (1988). Distinguishing panics and information-based bank runs, *J. Polit. Econ.* **96,** 568–592.

James, C. M. (1987). Some evidence on the uniqueness of bank loans, *J. Finan. Econ.* **19,** 217–235.

Nakamura, L. I. (1993). Commercial bank information: Implications for the structure of banking, *in* "Structural Change in Banking" (M. Klausner and L. J. White, Eds.), pp. 131–160, Business One Irwin, New York.

Niehans, J. (1978). "The Theory of Money," Johns Hopkins Univ. Press, Baltimore.

Qi, J. (1994). Bank liquidity and stability in an overlapping generations model, *Rev. Finan. Stud.* **7,** 389–417.

Qi, J. (1996). Efficient investment and financial intermediation, *J. Bank. Finan.* **20,** 891–900.

Rajan, R. G., and Winton, A. (1995). Covenants and collateral as incentives to monitor, *J. Finance* **50,** 1113–1146.

Ramakrishnan, R. T. S., and Thakor, A. V. (1984). Information reliability and a theory of financial intermediation, *Rev. Econ. Stud.* **51,** 415–432.

Wallace, N. (1988). Another attempt to explain an illiquid banking system: The Diamond and Dybvig model with sequential service taken seriously, *Fed. Res. Bank Minn. Quart. Rev.* Fall, 3–16.

Wallace, N. (1990). A banking model in which partial suspension is best, *Fed. Res. Bank Minn. Quart. Rev.* Fall, 11–23.

Winton, A. (1995). Delegated monitoring and bank structure in a finite economy, *J. Finan. Intermediation* **4,** 158–187.