

Liquidity Risk, Liquidity Creation, and Financial Fragility: A Theory of Banking

Douglas W. Diamond and Raghuram G. Rajan

University of Chicago and National Bureau of Economic Research

Loans are illiquid when a lender needs relationship-specific skills to collect them. Consequently, if the relationship lender needs funds before the loan matures, she may demand to liquidate early, or require a return premium, when she lends directly. Borrowers also risk losing funding. The costs of illiquidity are avoided if the relationship lender is a bank with a fragile capital structure, subject to runs. Fragility commits banks to creating liquidity, enabling depositors to withdraw when needed, while buffering borrowers from depositors' liquidity needs. Stabilization policies, such as capital requirements, narrow banking, and suspension of convertibility, may reduce liquidity creation.

Banks perform valuable activities on either side of their balance sheets. On the asset side, they make loans to difficult, illiquid borrowers, thus enhancing the flow of credit in the economy. On the liability side, they provide liquidity on demand to depositors. We know from Diamond and Dybvig (1983) that banks can transform illiquid assets into more liquid demand deposits. But there seems to be a fundamental incompatibility between the two activities: the demands for liquidity by depositors may arrive at an inconvenient time and force the fire sale liquidation of illiquid assets. Furthermore, because depositors are served in sequence, the prospect of fire sales may precipitate self-fulfilling runs

We acknowledge helpful comments from an anonymous referee, Patrick Bolton, Michael Brennan, John Cochrane, Phil Dybvig, Ron Giammarino, Gary Gorton, Lars Hansen, Bengt Holmström, Oliver Hart, Raphael Repullo, and Bruce Smith. We thank Heitor Almeida for valuable research assistance. We are grateful for financial support from the Center for Research in Security Prices and the National Science Foundation.

[Journal of Political Economy, 2001, vol. 109, no. 2]

© 2001 by The University of Chicago. All rights reserved. 0022-3808/2001/10902-0003\$02.50

that further jeopardize bank activities. In country after country, an army of regulators supervise banks to protect them from their own fragility. And after every banking crisis, economists point to how risky the combination of bank activities is and how it makes sense to legislate the separation of banking activities. But is there logic, hitherto unnoticed, for the bank's choice of activities? Is financial fragility a desirable characteristic of banks?

Our paper suggests yes. Assets are illiquid in our framework because the best users of the asset cannot commit to employing their specialized human capital on behalf of others. So, for example, an entrepreneur with a project can threaten to quit at an interim stage. This gives him bargaining power over the surplus generated, no matter what contract he signs at the outset with financiers; hence he cannot promise to pay out surplus fully. Thus real assets are illiquid. They cannot be sold or borrowed against for the full value they generate because of this limited ability to commit.

Of course, lenders can be given control rights such as the right to seize assets and put them to alternative use (i.e., liquidate). This will give them some bargaining power over the surplus that the entrepreneur generates. Also, as a lender sees the entrepreneur's *modus operandi* over the course of the financing relationship, she will understand the entrepreneur's business better and thus will be able to liquidate assets for more. So is illiquidity no longer an issue if a lender develops specific skills over the course of the relationship that give her a sufficiently strong liquidation threat to extract full repayment?

The answer is no if the "relationship" lender is likely to face a need for liquidity. After having made the loan, the lender may need money for a new business or for consumption, and she may not be able to raise money elsewhere to finance this need. She will have to sell the loan, or use it as collateral, to raise money to meet either of these liquidity needs. The amount raised will be unavoidably low if her specific ability to collect future loan payments from the entrepreneur is lost when she undertakes the opportunity. Even if her loan collection skills persist after she undertakes the new business or consumption opportunity, there is a potential problem. Because she typically cannot commit to using her skills to extract repayment from the entrepreneur on behalf of others, her loan to the entrepreneur is illiquid: she cannot sell it or raise money against it to the full value of the repayment she expects to extract from the entrepreneur.

The lender's inability to sell, or borrow against, a loan for the full present value of what she could extract in the future, when faced with a need for liquidity, affects the initial terms of her loan and her interactions with the entrepreneur. Even though the loan is riskless when held to maturity, its low sale price makes it risky for one with potential

liquidity needs. Since the loan does not repay much when she has the highest use for money, the lender may demand a premium from the entrepreneur and may incorporate contractual terms that allow her to liquidate the entrepreneur's project when she is in need of liquidity. Thus the lender's need for liquidity creates liquidity risk for the entrepreneur, and the lender may even refuse to lend if her likelihood of having a liquidity need is high enough.

The adverse consequences of illiquidity could be avoided if a relationship lender with persistent loan collection skills could borrow the full value of the loan to the entrepreneur when she faces a need for liquidity. She can do this only if she can commit to deploying her extraction skills for free in the future on behalf of the new lender(s). One way to commit is for the relationship lender to borrow using demand deposits: a fragile capital structure that is subject to a "run." If the relationship lender threatens to withdraw her specific collection skills as a ploy to get more rents, she will precipitate a run by depositors, which will drive her rents to zero. Fearing this outcome, she will not attempt to renegotiate any precommitted payments and will be able to pass through to depositors all that she extracts from the entrepreneur. So the fragility of her capital structure enables the relationship lender to borrow against the full value of the illiquid loan she holds. This then enables her to lend up-front without demanding an expected return premium for illiquidity and without liquidating the entrepreneur when faced with a liquidity shock. Financial fragility allows liquidity creation.

More generally, we can extend this idea to explain financial intermediaries such as banks. Suppose that no single lender has the wealth to fully finance an entrepreneur's project. If lenders have to club their money together, it makes sense for one of them to become an intermediary, whom we shall call the banker. The banker can issue demand deposits to the other lenders. The fragile deposit structure allows persistent relationship building to be delegated to the banker (much as in Diamond [1984], where the intermediary monitors on behalf of investors) because the banker can commit to paying depositors what she can extract from the entrepreneur on the basis of her specific loan collection skills. If initial depositors have liquidity needs at an interim stage, the banker can refinance by issuing fresh demand deposits and can thus meet their needs. New depositors will be willing to replace old depositors who withdraw since new depositors will be confident that the bank will repay. As a result, the bank's deposits are a desirable asset for investors who have need for liquidity and are liquid even though loans made by the bank are illiquid. Moreover, the bank shields entrepreneurs from the liquidity needs of depositors, thus creating liquidity on both sides of the balance sheet.

The rest of the paper is organized as follows. We present the frame-

work in Section I and derive the nature of the optimal financing contract in Section II. In Section III we explain how a bank can achieve the second-best outcome. Our model has a number of implications, which we explore in Section IV. We relate the paper to the literature in Section V, especially Diamond and Dybvig (1983) and Calomiris and Kahn (1991). We conclude in Section VI with a description of some of the many extensions to this model that are possible.

I. The Framework

A. Projects

Consider an economy with entrepreneurs and potential financiers. The economy lasts for two periods and three dates—date 0 to date 2. All agents have linear utility of consumption. Each entrepreneur has a project, which lasts for two periods. The project requires an investment of up to \$1 at date 0. If an entrepreneur works on his project, it produces a riskless cash flow of C_t at date t (all amounts are per dollar invested). Investment in the project is observable and contractible: the entrepreneur cannot divert the funds to another use. There is also a storage opportunity that returns \$1 for every dollar invested.

B. Financing

Entrepreneurs do not have money to finance their projects. There are many potential financiers with an endowment of one unit at date 0 and arbitrarily many other financiers with smaller endowments at each date. The exact distribution of endowment is not critical.

The entrepreneur can raise money by issuing contracts (which for convenience only we shall call loans). We assume very little about the form of the contract other than that there is a required payment on a particular date or dates and the lender gets control rights over the asset if the entrepreneur defaults. This specification subsumes a contract in which the lender always has control rights since that is obtained by setting the required payment to infinity. So a contract specifies only repayments P_t that the borrower is required to make at date t (with repayments possibly contingent on the liquidity shock to the lender, which we shall describe shortly).

C. Relationship Lending

The date 0 lender to a project, whom we shall call the relationship lender, develops specific skills in identifying the liquidation value of the assets: she has been in a relationship with the entrepreneur at an early

enough stage to know how the business was built, knows which markets personnel were hired from and where assets were bought, and knows what alternative strategies were considered. So she can identify the second-best use(r) of the asset more precisely than anyone else. Formally, before date t , she can take the asset away from the entrepreneur and put it in an alternative use to generate a present value X_t . This is the "liquidation value" of the assets. After date 2, the liquidation value collapses to zero. There is symmetric information about cash flows and liquidation values.

Because other lenders who come in later do not have her specific skills in finding the next-best alternative use, they can generate only βX_t , where $0 \leq \beta < 1$, from the asset. Since educating the initial lender takes time and effort, we assume that an entrepreneur can borrow from only one such lender.¹ We shall discuss various other interpretations of the relationship lender's skills later.

D. Limited Commitment

There are two limitations on the willingness of financiers to lend. First, at any date an agent can commit to working on the specific venture only for that date (as in Hart and Moore [1994] and Hart [1995]): the law prevents him from irrevocably selling himself into bondage. This implies that after borrowing and investing at date 0, the entrepreneur could threaten to quit before cash flows are due to be produced at date 1 unless the terms of financing are renegotiated. He can do this again before date 2. Similarly, the relationship lender cannot commit to others that she will use her specific skills on their behalf at any future date.

This implies that loans can be renegotiated. For simplicity, we assume that the entrepreneur has all the bargaining power; if the entrepreneur defaults on a scheduled payment, he can make a take-it-or-leave-it offer with a revised menu of payments and can commit not to work for that period if an impasse is reached. We show later that our qualitative results hold when the lender has some additional bargaining power. If the lender accepts the revised schedule, the entrepreneur produces that date's cash flow, makes the spot payment required by the revised schedule, and continues in possession of the asset. If the lender rejects the revised schedule, the cash flow is not produced that period; the lender takes possession of the asset and does as she chooses with it (see figs. 1 and 2).

¹ The banker's specific skills in our model resemble the relationship-specific information bankers get in Greenbaum, Kanatas, and Venezia (1989), Sharpe (1990), and Rajan (1992).

Entrepreneur offers an alternative current payment, $P_2' < P_2$.

He will supply current human capital and make the alternative current payment if and only if agreement is reached.

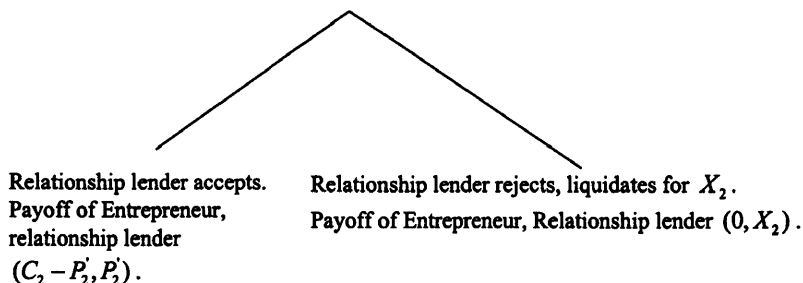


FIG. 1.—Bargaining at date 2 between a relationship lender and an entrepreneur

E. Liquidity Shock

The second limitation on the willingness to lend up-front is that with probability θ at date 1, the relationship lender could get a liquidity shock—a highly valued investment or consumption opportunity—which makes her impatient (we denote this type of lender by superscript I). The shock increases her personal rate of time preference, making one unit of date 1 goods worth R units of date 2 goods to her. We refer to a lender who does not get a liquidity shock as patient or $\sim I$. We assume that the realization of the liquidity shock is the relationship lender's own private information: outsiders have no way of finding out how strong her desire to consume is or how good the investment opportunity really is. This specification of liquidity shocks is similar to that in Bryant (1980) and Diamond and Dybvig (1983), but without the introduction of risk aversion.

Apart from those who get the liquidity shock described above, no one discounts future consumption: the discount rate is zero. We assume that endowments are sufficiently large relative to projects and that there are enough investors who do not get a liquidity shock that storage is always in use at the margin at each date. So there is no aggregate shortage of liquidity of the type in Diamond (1997), and at any date a claim on one unit of consumption at date $t + 1$ sells in the market for one unit at date t .

We make the following assumptions.

ASSUMPTION 1. $\min [C_1 + C_2, (C_1 + C_2)/X_1] > R > 1$.

ASSUMPTION 2. $C_2 > X_2$.

ASSUMPTION 3. $\max [X_1, X_2] \geq 1$.

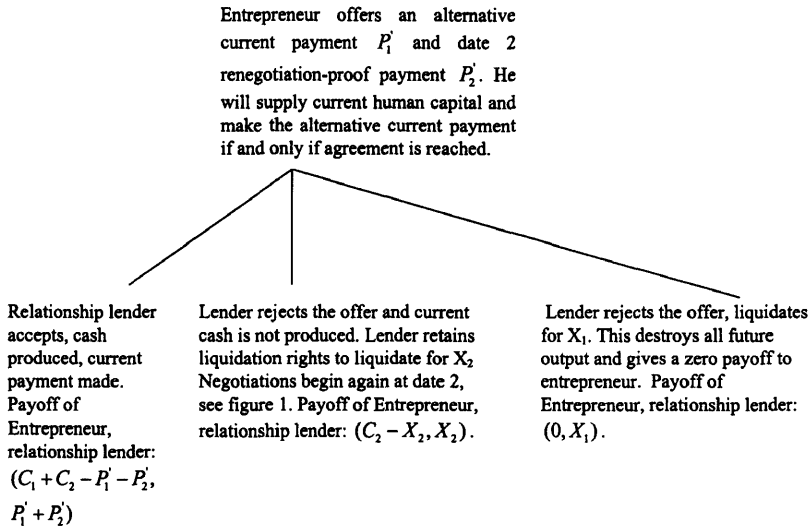


FIG. 2.—Bargaining at date 1 between a patient relationship lender and an entrepreneur

Assumption 1 indicates that the project produces greater returns, $C_1 + C_2$, viewed from both the date 0 investment and the date 1 opportunity cost of X_1 , than the relationship lender's discount rate, even when hit with a liquidity shock. This ensures that it is always efficient for the entrepreneur to continue his project. Assumption 2 indicates that the project is worth continuing at date 2. Assumption 3 ensures that the project can be financed if the lender does not suffer liquidity shocks, because it will turn out that the entrepreneur can commit to paying the relationship lender an amount equal to X_1 at date 1 or X_2 at date 2, the liquidation value of his assets.

The time line of events thus far is seen in table 1.

II. Optimal Contracts

The amount invested in each project can be anywhere between zero and one. It turns out that in this world of certainty, the entrepreneur will always optimally invest everything he raises at date 0 and not store any cash.² If the lender finds the terms acceptable for any specific

² The contract space is rich enough that the entrepreneur does not need to store cash to protect himself against unnecessary liquidation by the lender. Moreover, the entrepreneur's incentive to protect himself against necessary liquidation by holding cash can be dealt with by setting the date 1 payment high enough.

TABLE 1
TIME LINE

Date 0	Date 1	Date 2
1. Entrepreneur offers loan terms to the lender	1. Lender suffers (or does not suffer) a private liquidity shock	1. Entrepreneur makes any required payment or offers alternative schedule
2. Loan is made if the lender accepts	2. Entrepreneur makes any required payment or offers alternative schedule	2. If alternative schedule is offered, lender accepts or takes possession of asset
3. Amount is invested	3. If alternative schedule is offered, lender accepts or takes possession of asset	
4. Lender acquires specific expertise in liquidation through the relationship	4. Loan may be sold to another lender	

amount raised, linearity implies that she will find them acceptable (suitably scaled) for any other amount. Taken together, this implies that we can focus, without loss of generality, on situations in which the entrepreneur borrows \$1 and invests it entirely in the project.

A. *The Entrepreneur's Optimization Problem*

In offering to borrow using a particular loan contract up-front, the entrepreneur wants to maximize his expected payoffs. If it is to be rational for the relationship lender to make the loan at date 0, she should expect to get at least $(1 - \theta) + \theta R$, which is her expected utility, given the probability of a liquidity shock, using her alternative of storage. Because there is no shortage of funds at date 0, we assume that the entrepreneur can borrow by matching the expected utility from this outside option of storage.

a) Because the returns from continuation exceed R , the entrepreneur would like to ensure that his project does not get liquidated unless this is absolutely necessary for the lender to make the loan—unless liquidation is the only way to satisfy the lender's individual rationality constraint.

b) Since the entrepreneur values cash flows the same in either state and at either date (because his discount factor is one), he would like to commit up-front to giving the relationship lender as much as possible at date 1 when she is impatient and needs liquidity. Define the illiquidity premium as the increase in expected payments the entrepreneur has to make over and above the expected payments the lender would receive

if she had invested in storage at date 0 (which pays \$1 for sure at date 1 regardless of the lender's type). Then the illiquidity premium is³

$$\theta(R - 1)(1 - \text{date 1 cash flow to impatient lender}), \quad (1)$$

where θ is the probability that the lender gets the liquidity shock, and the cash flow to the impatient lender includes any payments by the entrepreneur and any proceeds from the loan sale or project liquidation. We call an asset with a zero illiquidity premium "liquid." An asset with a positive illiquidity premium is one in which more is paid in expectation over time to the holder because it does not pay as much when she is impatient at date 1 as storage. Thus a relationship lender is willing to pay less at date 0 for such an asset than the present value of its future repayments discounted at the gross market interest rate (of one). It is "illiquid." Moreover, we shall see that the reason for the illiquidity premium is that the impatient relationship lender will realize less from the loan at date 1 than the present value of payoffs if she held the loan to maturity and discounted at the market interest rate. Thus a loan is illiquid because it has poor state-contingent payoffs, and it has poor state-contingent payoffs because it cannot fetch as much in times of need as the present value of what the holder could realize if she did not have the need.

Finally, note that if the amount that an impatient lender can realize at date 1 exceeds \$1 (and consequently the total payment to a patient lender is less than \$1), the loan provides the lender *liquidity insurance*. In this case, the expected return premium required is negative (i.e., the loan has a lower expected return than a liquid asset). We shall see that because the need for liquidity is private information, no such "more than liquid" asset exists.

When principles *a* and *b* are taken together, the entrepreneur's first priority in the contract he offers is to minimize the probability of liquidation, following which he will focus on reducing the illiquidity premium by ensuring that the contract pays the maximum possible to the lender if she turns out to be impatient at date 1. We shall show in a wide variety of circumstances that the loan from the relationship lender to the entrepreneur will be illiquid because of their inability to commit their specific skills to others.

³ Let V be the expected payment in all other states and V_1' be the date 1 cash flow to the impatient lender. The illiquidity premium is $V + \theta V_1' - 1$. For the lender to lend, $V + \theta R V_1' = (1 - \theta) + \theta R$. Solving for V , substituting in the illiquidity premium, and rearranging, we get the simplified expression.

B. Contract Renegotiation

There are limitations on how much the entrepreneur can commit to paying because he can always threaten to quit and thus bargain his payments down. Let us describe what happens when a contract is renegotiated. Suppose that at date 2 the loan has not been previously sold. The entrepreneur has to renegotiate with the relationship lender. The entrepreneur may refuse to make the prespecified payment P_2 and, instead, may make an offer of a lower payment. In response, the relationship lender can accept the offer or reject it and liquidate the assets to obtain X_2 (see fig. 1). Thus if P_2 exceeds X_2 , the entrepreneur will renegotiate knowing that the lender will be satisfied with an offer of X_2 . Thus at date 2 the entrepreneur will pay $\min\{P_2, X_2\}$.

Now consider what happens at date 1 if the relationship lender is patient (i.e., has not received a liquidity shock). If the entrepreneur initiates renegotiation, the lender can accept the entrepreneur's offer. Alternatively, she can liquidate and get X_1 or hold on to the asset and get X_2 at date 2 (see fig. 2). Since the relationship lender has the best extraction skills, selling when patient is a weakly dominated option. So she will accept any offer that makes payments amounting to $\max\{X_1, X_2\}$ over dates 1 and 2; any payment left for date 2 should be enforceable, that is, should be less than X_2 . If the promised payments $P_1 + P_2$ exceed $\max\{X_1, X_2\}$, they will be renegotiated down to this level (note that if a state-contingent contract had been written, P_1 and P_2 would be the payments given that the relationship lender has no need for liquidity).

Suppose, instead, that the lender does receive a liquidity shock at date 1 and becomes impatient. If the borrower attempts to renegotiate, the lender can liquidate the asset and get X_1 or get a present value of X_2/R by waiting till date 2.

There is another option. A number of potential lenders at date 1 have endowments but have not received a liquidity shock. So they can substitute for the relationship lender, albeit imperfectly because they do not have her specific skills. So a third option for the relationship lender is to sell the loan contract at date 1 to an unskilled lender who has not suffered a shock. All claims and rights of the initial lender pass on to the loan buyer. Let S be the maximum that the loan buyer can collect on the loan at date 2. For now, we take S as exogenous. We discuss its determination in Section IID below. Since the buyer will be from the large pool of lenders who have not suffered a liquidity shock and since the prevailing interest rate is zero, he will be willing to buy the claim at date 1 for what he expects to recover from the loan at date 2, that is, S . Since at most X_2 can be recovered from the entrepreneur at date 2, $S \leq X_2$.

In summary, the maximum date 1 present value the impatient relationship lender can extract (i.e., the largest renegotiation-proof amount) is $E^I = \max\{X_1, S, X_2/R\}$, and the maximum amount she can extract if patient is $E^{\sim I} = \max\{X_1, X_2\}$.

C. Optimal Financial Arrangements and the Consequences of Illiquid Loans

The next question is, How much can actually be paid to the relationship lender at date 1; that is, what is the cash or liquidity available for her to consume or invest elsewhere? If the lender does not liquidate, she can get the maximum at date 1 if she collects all the date 1 cash flow generated by the entrepreneur and sells date 2 promised payments for the maximum they could fetch. In this case, she can get up to $C_1 + S$. She will get, at maximum, C_1 if the loan is not sold and the project is not liquidated. If she liquidates the project, she can be paid X_1 .

These limits on extractability and liquidity bound the maximum possible payments by the entrepreneur. If it is possible for him to offer a contract to the relationship lender with terms contingent on the realization of the lender's date 1 liquidity need, then these limits can be achieved. Even so, as lemma 1 indicates, the entrepreneur may be forced to pay liquidity premia or be liquidated, solely to meet the needs of the lender.

LEMMA 1. If the initial loan contract between the entrepreneur and the relationship lender can be directly contingent on the relationship lender's type (i.e., contingent on whether she receives a liquidity shock or not), then the financing available to the entrepreneur is characterized as follows:

- i. The entrepreneur will be financed at date 0 and will provide the lender liquidity insurance, with the loan carrying a negative illiquidity premium, if $\min\{C_1 + S, E^I\} > 1$.
- ii. The entrepreneur will be financed at date 0 and the loan will be a liquid asset with a zero illiquidity premium if $\min\{C_1 + S, E^I\} = 1$.
- iii. If $\min\{C_1 + S, E^I\} < 1$, the entrepreneur will be financed at date 0 but the loan will be illiquid and he will have to pay a positive illiquidity premium if either

$$C_1 + S < E^I, \quad E^{\sim I} \geq 1 + \frac{\theta}{1 - \theta} R[1 - (C_1 + S)], \quad (2)$$

$$C_1 + S \geq E^I, \quad E^{\sim I} \geq 1 + \frac{\theta}{1 - \theta} R(1 - E^I), \quad (3)$$

or

$$E^{\sim I} \geq 1 + \frac{\theta}{1 - \theta} [R(1 - \min \{C_1, E^I\}) - \min \{\max \{(E^I - C_1)R, 0\}, X_2\}]. \quad (4)$$

- iv. If none of inequalities (2), (3), or (4) hold, the entrepreneur will be financed at date 0 but only with the asset being liquidated when the lender is impatient if

$$E^{\sim I} > 1 + \frac{\theta}{1 - \theta} R(1 - X_1). \quad (5)$$

- v. The entrepreneur will not be financed at all at date 0 otherwise.

Proof. Consider the mechanism design problem the entrepreneur faces, keeping in mind that the contract can be renegotiated at date 1 after the lender's liquidity needs are determined (her type is revealed). We characterize the contract in terms of the payments it draws forth from the entrepreneur. Define V_t^j as the cash paid by an entrepreneur when the relationship lender is of type j (where $j \in \{I, \sim I\}$) at date t . We can restrict attention to renegotiation-proof contracts, where the entrepreneur can and will make the payment V_t^j . The entrepreneur's goal is to maximize Φ with respect to V_t^s , where

$$\Phi \equiv \theta(C_1 + C_2 - V_1^I - V_2^I) + (1 - \theta)(C_1 + C_2 - V_1^{\sim I} + V_2^{\sim I})$$

$$\text{if } V_1^I, V_1^{\sim I} \leq C_1 \text{ (no liquidation),}$$

$$\Phi \equiv \theta(X_1 - V_1^I) + (1 - \theta)(X_1 - V_1^{\sim I})$$

$$\text{if } V_1^I, V_1^{\sim I} > C_1 \text{ (liquidation in all states),}$$

$$\Phi \equiv \theta(X_1 - V_1^I) + (1 - \theta)(C_1 + C_2 - V_1^{\sim I} - V_2^{\sim I})$$

$$\text{if } V_1^I > C_1, V_1^{\sim I} \leq C_1 \text{ (liquidation when lender is type } I),$$

and

$$\Phi \equiv \theta(C_1 + C_2 - V_1^I - V_2^I) + (1 - \theta)(X_1 - V_1^{\sim I})$$

$$\text{if } V_1^I \leq C_1, V_1^{\sim I} > C_1 \text{ (liquidation when lender is type } \sim I)$$

subject to the following conditions: Payments are renegotiation-proof and individually rational:

$$V_1' \leq \max [X_1, C_1]$$

(maximum feasible date 1 payment) and

$$V_2^{\sim I} \leq X_2$$

(maximum date 2 payment enforceable by type $\sim I$).

If the date 2 portion of the loan, V_2' , is sold at date 1 by type I (impatient), then

$$V_2' \leq S$$

(maximum date 2 payment enforceable by buyer),

$$V_1' + V_2' \leq \max \left[X_1, S, \frac{X_2}{R} \right]$$

(maximum total payment enforceable if sold), and

$$\theta[(V_1' + V_2')R] + (1 - \theta)(V_1^{\sim I} + V_2^{\sim I}) \geq \theta R + (1 - \theta)$$

(lender's individual rationality when V_2' is sold).

If the date 2 portion of the loan, V_2' , is kept at date 1 by type I (impatient), then

$$V_1'R + V_2' \leq \max [X_1R, X_2]$$

(maximum total payment enforceable),

$$V_2' \leq X_2$$

(maximum date 2 payment enforceable by original lender), and

$$\theta(V_1'R + V_2') + (1 - \theta)(V_1^{\sim I} + V_2^{\sim I}) \geq \theta R + (1 - \theta)$$

(lender's individual rationality when V_2' is kept).

Contingent payments other than at date 1 to the impatient lender enter the objective function and the individual rationality constraints in the same way, and all combinations that satisfy the individual rationality constraint are equally good as long as there is no liquidation. The conditions in the lemma follow simply by substituting the maximum feasible date 1 payment to the impatient lender into the individual rationality condition and solving for the required payment to the patient lender. As long as that payment can be extracted, it will be made: since the patient lender is indifferent between payments at either date, the timing of cash payments is not relevant.

The optimal contract will have the impatient lender always selling the loan at date 1 if the individual rationality constraint can be satisfied when she does so, for this minimizes the illiquidity premium. However, the required payments to the patient lender may be so high that they cannot be extracted. There are two possibilities left. Since the relation-

ship lender can extract more than the loan buyer, it may be necessary for her to keep the loan even when impatient instead of selling it and extract enough to meet her individual rationality constraint. Of course, she has to extract enough to compensate for the higher illiquidity premium she will now need. Condition (4) ensures that this is feasible. If not, the entrepreneur may be able to borrow by allowing himself to be liquidated if the lender suffers a shock (see part iv of lemma 1). This can reduce the required payment to the patient lender to an extractable level. If not, it is not individually rational for the lender to make the loan in the first place (part v of lemma 1). Q.E.D.

Lemma 1 shows that the loan may have an illiquidity premium and liquidation may arise even when the lender's liquidity need can be directly contracted on.⁴ However, in contrast to what is assumed in lemma 1, the lender's liquidity need is private information. Therefore, the best the entrepreneur can do is to offer a contract that has a menu of possibilities at date 1 and allow the lender to select her preferred option on the basis of her type. This will introduce additional self-selection restrictions on what is possible in lemma 1.

COROLLARY 1. When the relationship lender's need for liquidity is private information, (i) the loan will be liquid (with a zero illiquidity premium) under the conditions of parts i and ii of lemma 1; and (ii) the loan will be illiquid, the illiquidity premium will be weakly higher, and the loan will be sold in weakly fewer circumstances than under the conditions of part ii of lemma 1. There will be no increase in liquidation relative to the case in which the lender's type can be directly contracted on.

Proof. Since the lender's type is not observable, we have to add incentive-compatibility constraints to the mechanism design problem. Payments are incentive compatible:

$$V_1^I + V_2^I \leq V_1^{\sim I} + V_2^{\sim I} \quad (\text{IC1})$$

(incentive compatibility of type $j = \sim I$),

$$V_1^I R + V_2^I \geq V_1^{\sim I} R + V_2^{\sim I} \quad (\text{IC2})$$

⁴ If liquidity needs were verifiable and were not an aggregate risk and if agents other than the entrepreneur have no limit on the commitment to paying, both storage and the loan payments could be transformed into type-contingent insurance contracts. For example, the lender could store and sell off the payment of \$1 in states in which she does not get a liquidity shock in return for additional payments when she does get a liquidity shock. In this case, the lender's outside option is $V_1^I = 1/\theta$ and $V_1^{\sim I} = V_2^{\sim I} = V_2^I = 0$. Unless the loan offers at least the same possibilities, i.e., unless it offers at least a payment of \$1 that either can be consumed at date 1 or can be assigned to others in states in which the lender is not in need of liquidity, it will require a premium expected return. In any case, the condition for the loan to be illiquid (and the entrepreneur not liquidated) remains $\min\{C_1 + S, E\} < 1$.

(incentive compatibility of type $j = I$ if type I keeps the loan), and

$$(V_1' + V_2')R \geq V_1'R + V_2' \quad (\text{IC3})$$

(incentive compatibility of type $j = I$ if type I sells the loan).

Condition IC1 indicates that the patient type must get (weakly) more in market value if she claims to be patient. Therefore, an incentive-compatible contract can, at best, pay both patient and impatient types the same total market value. Liquidity insurance is thus ruled out, for that would require the impatient type to be paid more than \$1 at date 1 and the patient type less than \$1 in total. However, fully liquid contracts are incentive compatible since they pay both types \$1 at date 1. Hence this proves part i of corollary 1.

Self-selection for the impatient type (IC2 and IC3) can typically be achieved by back-loading payments to the patient type and front-loading payments to the impatient type. A straightforward check of the various scenarios shows that IC2 (incentive compatibility for the impatient lender when she keeps the loan) imposes no additional constraints on the optimal type-contingent payments but IC3 can in one situation. If the cash available to the impatient relationship lender after a loan sale, $C_1 + S$, is small, the amount the patient type has to be paid has to be disproportionately large to meet the individual rationality constraint. The higher the probability of the liquidity shock, the higher this amount. But then even if this amount is back-loaded, the impatient lender may prefer it to receiving $C_1 + S$. Thus incentive compatibility may bind. In this case, we get payments to be incentive compatible for the impatient type only by having her retain, rather than sell, the loan at date 1. The set of parameters for which loans are designed to be sold is smaller than in the full-information case, and the illiquidity premium will be higher, because the impatient type does not sell the loan at date 1. However, it is easily shown that the lender can always be paid the additional premium, so liquidation does not increase. Q.E.D.

Finally, let us pinpoint the source of the illiquidity.

COROLLARY 2. If the maximum loan sale price, S , equals X_2 , the optimal contract will always be liquid.

Proof. If $S = X_2$,

$$E^I = \max \left\{ X_1, S, \frac{X_2}{R} \right\} = \max \{ X_1, X_2 \}.$$

Thus the maximum that can be paid at date 1 to the impatient lender in the absence of self-selection constraints is $\min \{ \max \{ X_1, X_2 \}, C_1 + X_2 \}$. But this is also the maximum a patient lender can extract, so it cannot be less than one for lending to occur. Self-selection constraints

(corollary 1) will limit the amount paid to \$1; hence the optimal contract will be liquid but will not provide insurance. Q.E.D.

D. The Sources of Illiquidity

Lemma 1 indicates that illiquidity premia and liquidation arise even with optimal contracts under perfect information because the entrepreneur cannot commit his human capital to the project. Corollary 1 indicates that, in general, when the lender's liquidity shock is private information, the required illiquidity premium will increase. The optimal contract under private information can, at best, provide full liquidity (zero illiquidity premium) to the lender but cannot provide liquidity insurance (negative illiquidity premium).

As corollary 2 indicates, the source of the illiquidity premium—which leads to liquidation and even denial of credit—is that the maximum loan sale price, S , is lower than the present value of the amount obtainable by the patient relationship lender, X_2 . To see why this may be so, we consider two cases.

Case 1.—On selling the loan, the relationship lender loses her liquidation skills and is not available to collect at date 2.

The assumption here is that by selling the loan, the relationship lender loses contact with the entrepreneur, and the business changes fast enough that she needs contact to keep her specific liquidation skills honed. Alternatively, if the liquidity shock is a profitable opportunity that takes the lender away from her existing ventures, it need not be the sale itself but rather the refocusing on the new opportunity that destroys collection abilities. For example, consider a firm offering trade credit. It can do so, in part, because it can repossess the goods it supplies and sell them elsewhere. If the firm starts a new business and drops the old business, it will lose the leverage it had with all those it had offered credit to.⁵ With the relationship lender having lost her skills, the loan buyer can extract a payment from the entrepreneur of, at maximum, βX_2 at date 2, where $\beta < 1$. So $S = \beta X_2$. There is nothing the relationship lender can do to increase the sale price above βX_2 , and lemma 1 then describes the best available financing outcomes.

Case 2.—The relationship lender retains liquidation skills even after selling the loan and can collect on the buyer's behalf at date 2.

⁵ Note that according to this interpretation, the lender who chooses to retain the project loan and stay in the old business does not have the time to fund the profitable new opportunity. This interpretation changes the individual rationality constraint if the loan is not sold. It is now

$$\theta(V_1' + V_2') + (1 - \theta)(V_1'' + V_2'') \geq \theta R(1 - \theta).$$

Qualitatively, the results are unchanged.

Case 2 is particularly appropriate if we think of the project's characteristics as relatively static and the relationship lender as one whose primary business is lending (e.g., a financial institution) and who acquires permanent liquidation skills relevant to the entrepreneur's business (such as knowledge of secondary markets for the borrower's assets) to extract payment. This is in contrast to case 1, where the liquidation skills are related to the lender's business.

What would a buyer pay now for the loan at date 1, given that he can hire the relationship lender to collect on his behalf at date 2? To determine this, suppose that at date 2 the entrepreneur tries to negotiate down the amount to be paid. The sequence of all such negotiations will be as follows: The entrepreneur will make an opening offer to the current lender, in this case, the unskilled date 1 loan buyer. The buyer can accept the offer (in which case the bargaining ends) or reject it (in which case the offer is off the table). At this point, the buyer can negotiate with the relationship lender about who will hear the last offer from the entrepreneur and exercise control rights. After these negotiations, the entrepreneur makes his last and final offer to whoever it is decided will respond, and the offer is either accepted or rejected and the project liquidated by the responder.

The negotiations between the loan buyer and the relationship lender take a simple form. The relationship lender offers to take over the loan collection for a fee. If the offer is rejected, it is the loan buyer who responds to the entrepreneur's last offer; if accepted, the relationship lender responds.

Even though the relationship lender still possesses specific skills at date 2, it turns out that all the loan buyer will get at date 2 is βX_2 , so he will pay only $S = \beta X_2$ at date 1. The reason is instructive. With backward induction from date 2, when the entrepreneur makes a final offer to the buyer, he will offer to pay βX_2 and the buyer cannot do better than accept. In anticipation of the final offer the borrower would receive from the entrepreneur, the relationship lender will offer to pay the buyer βX_2 from the X_2 she would collect from the entrepreneur if chosen to respond to his last offer. Again, anticipating that he cannot do better by rejecting the offer, the buyer will accept. Finally, in anticipation of these low offers, the entrepreneur's opening offer to the buyer will be βX_2 and the buyer will accept. The relationship lender's rents will be zero despite her retaining her skills.⁶

The relationship lender's skills do not get used despite her retaining them and being better at collecting. Note that this is not a socially

⁶ The precise split of the surplus between the loan buyer and the relationship lender when they negotiate does not alter this result, except if the loan buyer has all the bargaining power. In that case, the loan buyer gets all the surplus from bargaining, and it is as though we are back in a world with full commitment.

inefficient outcome, *ex post*, because all that her skills are used for at date 2 is in forcing a transfer from the entrepreneur. The reason she does not get a slice of the pie is that the buyer owns the loan and any control rights emanating from it, so the buyer can, and does, conclude a deal with the entrepreneur without reference to the relationship lender.

We thus come to the precise reason for illiquidity: If the loan buyer had the same skills as the relationship lender, then $\beta = 1$ and $S = X_2$ in both cases. So a necessary condition for both the illiquidity of the real asset (the project) and the financial asset (the loan) is specific skills. In the case of the project, it is the entrepreneur's greater ability to run it relative to a second-best operator (as, e.g., in Shleifer and Vishny [1992]); in the case of the loan, it is the relationship lender's better ability to recover payments relative to someone who buys the loan.

Case 2 highlights a further requirement: that the relationship lender not be able to commit at date 1 to using her specific skills at date 2 on behalf of the loan buyer. In other words, the loan is illiquid not just because the relationship lender's human capital is specific but also because she cannot explicitly commit to deploying it on behalf of others in the future.

E. An Example

Let $X_1 = 0.9$, $X_2 = 1.1$, $C_1 = 0$, $C_2 = 1.5$, $R = 1.4$, and $S = 0.8$. Note that the date 1 cash flow, C_1 , is zero in this example. Because $\max[X_1, X_2] = 1.1$, the entrepreneur cannot commit to paying more than \$1.1 even though he generates \$1.5 from the project. Moreover, when hit by a liquidity shock, the relationship lender gets more in present value by liquidating ($X_1 = 0.9$) than by selling the loan ($S = 0.8$) or holding it to maturity ($X_2/R = 0.79$).

When the probability of the liquidity shock, θ , is low, the relationship lender will sell the loan when hit by the shock in preference to retaining it. We have $V_1' = V_1'^I = 0$, $V_2' = S = 0.8$, and $V_2'^I = P_2$, where P_2 is set to satisfy the lender's individual rationality constraint and will rise from 1 to 1.1 as θ increases from 0 to 0.26. Since the entrepreneur cannot commit to paying the relationship lender more than $P_2 = 1.1$, when θ increases beyond 0.26, the only way the entrepreneur can satisfy the lender's rationality condition is to allow her to liquidate at date 1 and get 0.9 if she suffers the shock. So if $\theta > 0.26$, the entrepreneur will offer $V_1' = V_1'^I = 0$, $V_2' = 0.9$, and $V_2'^I = P_2$, where P_2 is again set to satisfy the lender's individual rationality constraint. Since liquidation generates more than a loan sale, the relationship lender will again find it rational to lend for the range $0.26 < \theta < 0.42$. But when $\theta > 0.42$, P_2 exceeds 1.1

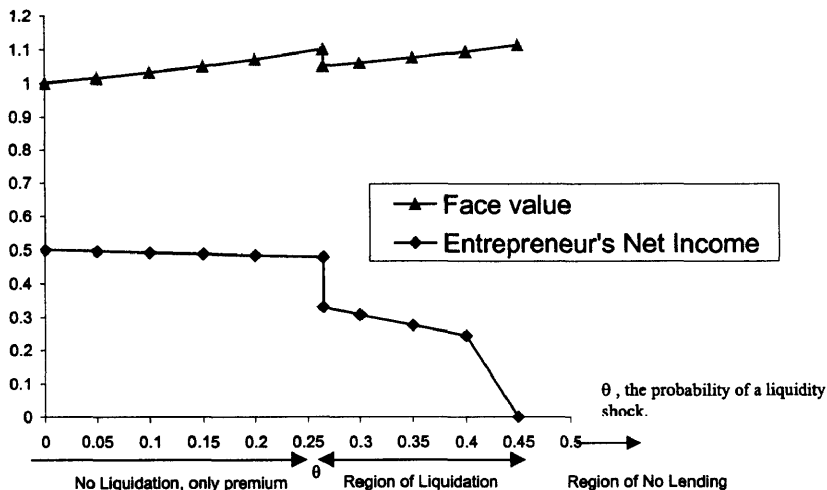


FIG. 3.—Date 2 face value when patient and entrepreneur's net income. The face value is V_2^l (see the proof of lemma 2), and the entrepreneur's net income is given by the objective function in the entrepreneur's maximization problem (see the proof of lemma 2). Both values are in units of output.

and is again not collectible. At this point, the probability of a liquidity shock for the lender is so high that lending is not individually rational.

We plot in figure 3 V_2^l (the date 2 payment required to be made by the entrepreneur) and the expected net income for the entrepreneur and in figure 4 the illiquidity premium. Both the date 2 payment and the illiquidity premium rise with θ , fall discontinuously once the project is liquidated, and rise again with θ until lending is infeasible. The entrepreneur's expected net income falls initially with θ because he pays a higher illiquidity premium, falls discontinuously once the project is liquidated conditional on a shock, and then falls again with θ because of not only the rising illiquidity premium but also the increasing liquidation.

Finally, what is the interpretation of the contracts that implement the outcomes described above? When $\theta < 0.26$, the contract is a standard long-term debt contract, and the lender will sell the date 2 payment if liquidity is needed. When $0.26 < \theta < 0.42$, the contract is a callable loan with face value V_2^l that has to be repaid on the demand of the lender; the lender does not get more than V_2^l from the collateral if the assets are liquidated. If the lender gets a liquidity shock, she demands immediate payment; since this cannot be made in full, she liquidates. Otherwise, she continues and accepts a payment of V_2^l at date 2.

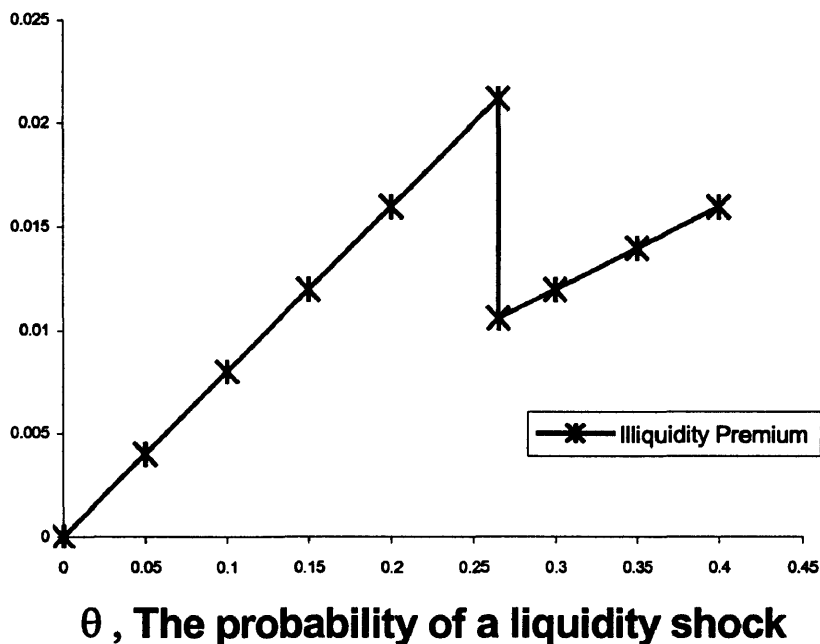


FIG. 4.—Illiquidity premium. The illiquidity premium is defined in the text and is in units of output (not percent).

F. Summary

If the relationship lender's loan collection skills do not persist beyond her liquidity shock, loans are unavoidably illiquid, with attendant consequences. Even if these skills do persist, however, the relationship lender cannot write explicit contracts committing to using her skills on behalf of buyers. This makes the loan illiquid if sold. But could she somehow devise a setting in which she can effectively commit to using her skills to recover payments from the entrepreneur and pass them on?

III. Financial Intermediation

We now argue that if the relationship lender "borrows against the loan" by setting up as a financial intermediary with a fragile capital structure (one subject to a run), she can commit to passing through everything she extracts from the entrepreneur. This allows her to raise up to X_2 at date 1 from investors, which is the same as having $S = X_2$. As a result, the intermediary can drive the illiquidity premium in the loans she

makes to zero. Entrepreneurs will not have to suffer liquidation or be unable to borrow simply because of the liquidity needs of the intermediary. This is what we now show.

A. *The Basic Argument*

Suppose that the loan made to the entrepreneur (henceforth the “project loan”) is in default at date 2. What we showed in Section II is that if the unskilled lender owns the project loan and all the control rights associated with it, he can reach a deal directly with the entrepreneur without the consent of the relationship lender. The entrepreneur will pay βX_2 , giving the relationship lender nothing. By contrast, if the relationship lender’s consent is necessary to any overall agreement, for example, if she owns the project loan, she can collect X_2 from the entrepreneur. So if the relationship lender wanted to borrow from an unskilled lender at date 1 (instead of selling the project loan), ideally she would retain ownership of the loan at date 2 and thus all rights to collect it as long as she makes a prespecified payment, say X_2 , to the unskilled lender.⁷ If a smaller payment were offered, the unskilled lender would have the right to seize the project loan, with the attendant loss of rents to the relationship lender.

Unfortunately, this will not work. Before dealing with the entrepreneur at date 2, the relationship lender can threaten to not collect the loan for the unskilled lender. The single unskilled lender will accept an offer from the relationship lender, who asks to retain loan collection rights in return for making a payment of βX_2 .

The role of demand deposits issued to multiple unskilled lenders (i.e., the fragile capital structure) is to deter unskilled lenders from accepting such an offer. The reason they refuse is that they have the unilateral right to demand immediate payment of their full claim, with depositors being paid in the order in which they show up for payment, until the relationship lender has nothing left. Thus a subset of depositors can be made whole if they run to demand payment, even when the collective cannot. Furthermore, in order to satisfy these depositors, the relationship lender will either have to sell the project loan to raise cash to pay them or give them the project loan (or fractional pieces thereof). In either case, the depositor run will transfer the ownership of the project loan to the unskilled, with the attendant loss of rents to the relationship lender. Anticipating the run and the loss of rents, the relationship lender will not attempt to renegotiate and can thus commit, by issuing demand

⁷ If weak preference is insufficient, then substitute $X_2 - \epsilon$, where ϵ is an arbitrarily small rent that goes to the relationship lender to make her strictly prefer honoring her commitments to having the loan seized.

deposits at date 1, to paying out up to X_2 at date 2. Thus the role of the first-come, first-served constraint in demand deposits is to create a collective action problem that forces a transfer of the ownership of the project loan whenever the relationship lender attempts to renegotiate. Let us now elaborate.

B. *The Demand Deposit Contract*

Suppose that at date 1 the relationship lender (henceforth called the "banker") borrows against the project loan by issuing demand deposits to a large number, n , of unskilled lenders, each depositing $1/n$ of the amount. Let P_2 be the amount owed at date 2 by the entrepreneur, where $P_2 \leq X_2$ without loss of generality. If the banker can commit to paying this out to depositors, the total promised to depositors will be d_2 , which equals P_2 , and the amount promised each depositor will be d_2/n . It will be convenient to think of values with the project loan as the unit, so each depositor holds $1/n$ units.

A depositor who did not previously withdraw and is not paid the promised amount on demand can withdraw if assets remain in the bank. To be concrete, we assume that withdrawal amounts to seizing financial assets (i.e., the project loan), with market value equal to the promised amount of the deposit. We assume that depositors seize assets, though this is identical in outcomes to depositors forcing the lender to sell assets and paying them the realized cash until they are made whole. Note that the market value of the project loan is $\min\{P_2, \beta X_2\}$ if the entrepreneur has not defaulted and βX_2 if he has. Once assets are seized or sold, the ownership of the project loan transfers away from the banker. This effectively "disintermediates" the bank.

C. *Negotiations with a Banker*

Suppose that the banker tries to renegotiate deposit payments down at date 2 and makes an offer (see fig. 5). The banker's threat to withdraw her human capital unless depositors accept lower payment is nontrivial because the entrepreneur will be faced with unskilled lenders if the banker does not provide her services, and he will pay less.

Each depositor must simultaneously choose whether to accept the banker's offer or to withdraw. By accepting the offer, the depositor forfeits his right to demand payment and must receive payments from bank assets that have not been seized by other depositors.⁸ If the depositor runs to withdraw, he will be paid as long as there are enough

⁸ This is not strictly necessary. It is sufficient that he also be forced to join the withdrawal line in order to obtain the lower payment he has agreed to.

assets in the bank to pay all those who demand payment. However, if the total promised payment to those who run exceeds the market value of bank assets, the bank's assets will have to be rationed.

Rationing takes the form of a first-come, first-served mechanism. Those who run are assigned a place in a line by a fair lottery. Each depositor in the line seizes a fraction $d_2/\beta X_2$ of the defaulted project loan per unit held until the bank runs out of assets. A depositor's right to his share of the bank's assets is ensured only when he physically withdraws it. The depositors who run will thus get paid in full or get nothing, except, perhaps, for one depositor who will get only a fraction of his claim paid before the bank runs out. The banker cannot unilaterally prohibit seizure, implying that she cannot suspend convertibility of deposits to cash.

After the banker and depositors have determined who holds the project loan (perhaps through a run), the entrepreneur negotiates with the holder and makes the negotiated payment. The sequence of events is as follows:

1. Banker makes offer to depositors.
2. Depositors either accept or withdraw. If they withdraw, they receive a random place in line if rationed. Those who withdraw seize assets with a market value equal to the face value of their deposit.
3. Entrepreneur negotiates with whoever holds the loan (the banker as in fig. 1 or, if disintermediated, the depositor or loan buyer as in fig. 3).

D. Response to the Banker's Offer to Negotiate

Unless the market value of the bank's assets is sufficient to pay all depositors, *any* offer that the banker makes that imposes losses on depositors will lead to a run in which all depositors queue to demand payment. To see this, suppose that the entrepreneur has attempted to renegotiate the loan and the banker offers depositors a net payment of $d'_2 \leq d_2$ per unit. As in the notation in Diamond and Dybvig (1983), let f denote the fraction of depositors who run to withdraw. Each depositor gets d_2 per unit held if sufficient assets remain in the bank. This requires a fraction $d_2/\beta X_2$ of the loan to be seized per unit held. If f , the fraction who run, exceeds $\hat{f} = \beta X_2/d_2$, all the bank's assets will be seized, and depositors must be rationed on the basis of their (random) place in line. Let f_j be the fraction of deposits withdrawn before depositor j . The realized payoff, per unit held, from running is

B offers D $d_2' < d_2$

D accepts, does not join a bank run, and receives payment from portion of loan that remains in the bank after completion of run by other depositors.

B negotiates the portion of the loan left after run with E .

D payoff per unit deposited is

$\text{Min}\{d_2', \lambda(f)\}$, where $\lambda(f)$ is given by:

$$\lambda(f) = \begin{cases} \frac{1-f \frac{d_2'}{\beta X_2}}{1-f} X_2 & \text{if } f \leq \frac{\beta X_2}{d_2} \\ 0 & \text{if } f > \frac{\beta X_2}{d_2} \end{cases}$$

D rejects, joins a bank run (simultaneously with other depositors) to seize specified amount of loans and negotiate directly with entrepreneur.

Expected depositor payoff from run (per unit deposited) is $\rho(f)$ given by:

$$\rho(f) = \begin{cases} d_2 & \text{if } f \leq \frac{\beta X_2}{d_2} \\ \frac{\beta X_2}{f} & \text{if } f > \frac{\beta X_2}{d_2} \end{cases} \quad (\text{see below})$$

where f is the fraction of depositors that run.

D rejects, does not join a bank run and enters into negotiation with B about who will negotiate loans that remain in the bank after completion of run by other depositors. D 's payoff per unit deposited is bounded above by $\beta \lambda(f)$ (where $\lambda(f)$ is defined under the payoff from accepting the offer).

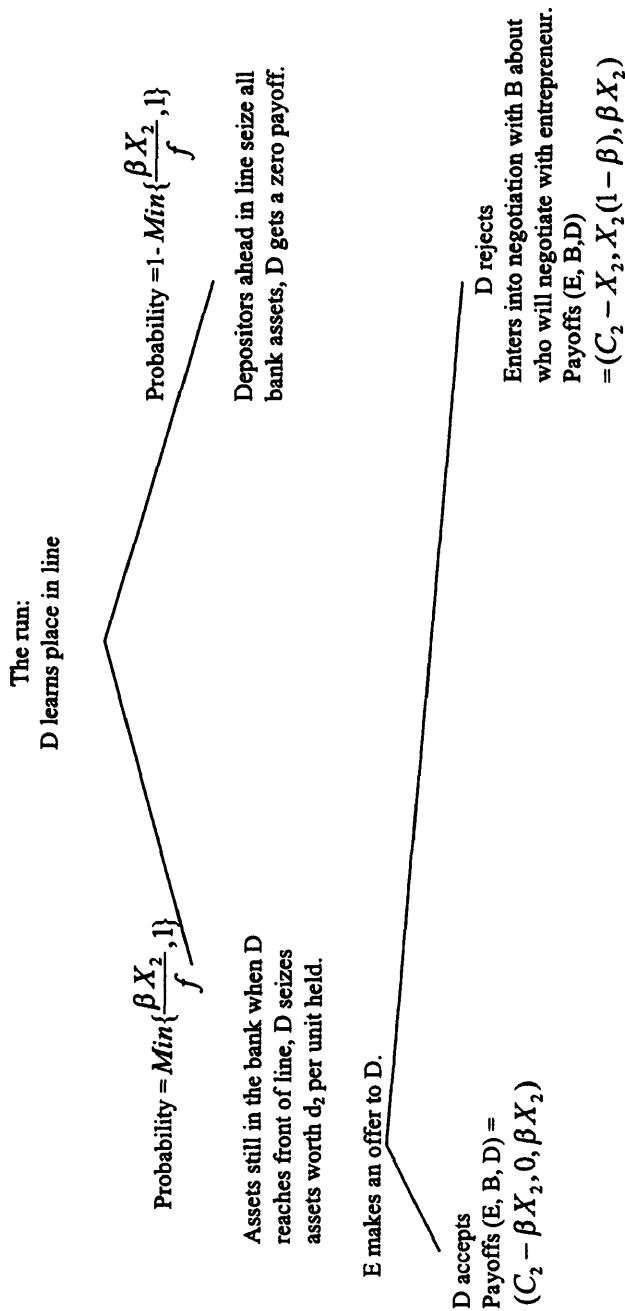


FIG. 5.—Depositor's (D) response to banker's (B) offer of $d_2 < d_2$

$$\rho(f_j) = \begin{cases} d_2 & \text{if } f_j \leq \frac{\beta X_2}{d_2} \\ 0 & \text{if } f_j > \frac{\beta X_2}{d_2} \end{cases}$$

When the expectation is taken over the place in line, the expected payoff, per unit held, from running is

$$\bar{\rho}(f) = \begin{cases} d_2 & \text{if } f \leq \frac{\beta X_2}{d_2} \\ \frac{\beta X_2}{f} & \text{if } f > \frac{\beta X_2}{d_2} \end{cases}$$

The payoff from accepting the banker's offer d'_2 must be paid from assets that remain after depositors who run are paid off. So the payment to the depositors who do not run is $\min\{d'_2, \lambda(f)\}$, where

$$\lambda(f) = \begin{cases} \frac{1 - f(d_2/\beta X_2)}{1 - f} X_2 & \text{if } f \leq \frac{\beta X_2}{d_2} \\ 0 & \text{if } f > \frac{\beta X_2}{d_2} \end{cases}$$

For any offer $d'_2 < d_2$ requiring concessions, either a depositor can make himself whole by running or $f > \beta X_2/d_2$. The payoff from running is positive and that from accepting the banker's low offer is zero. Therefore, running is a dominant strategy since at best the depositor will be at the head of the line and come out whole; at worst he will be at the end of the line and in the same position as if he had not run.

The payoff from running never exceeds d_2 for all values of withdrawals, f . Therefore, if at date 2 the banker commits to paying d_2 , there will be no run.⁹

In fact, we prove a stronger result in the lemma. Suppose, as is more realistic, that the first-come, first-served mechanism does not assign the project loan to depositors. Instead, it gives them a position in a real queue in front of the bank window, where they have the right to seize assets or waive their right when they reach the window, in response to fresh offers the banker may make. If a depositor waives his rights, more depositors farther back in line will be able to withdraw. Can the banker's original offer of less than full value (which triggered the run) be followed by a sequence of offers to those in line that persuades them not to seize assets? The answer is no. If there is sufficient market value of

⁹ Because the banker can commit to d_2 and those who currently store can replace depositors who withdraw their money, it is reasonable that there will be no panic-based runs.

assets in the bank to pay a given depositor's claim in full, the depositor will make no concession; he can simply withdraw and be made whole. As long as there are depositors behind him who will not get paid in full, the depositor cannot accept a promised payment out of future loan collections. A depositor farther back in line would seize the assets, leaving nothing for this depositor. As a result, any original offer that gives depositors less than full value results in total disintermediation.

Once the bank is disintermediated by a run, the project loan is in hands other than the banker's. As in the discussion in the previous section, the entrepreneur will make offers that the new holders of the loan will accept without hiring the banker to collect. As a result, the run drives the banker's payoff to zero. Anticipating this outcome, the banker will not try to renegotiate the original deposit claims. Consequently, the banker can commit to passing through the full amount that she can collect at date 2 from the entrepreneur to unskilled depositors and thus can raise $d_2 = P_2$ at date 1.

This argument is essentially a sketch of the proof that banks can create liquidity by borrowing more than the market value of their illiquid loans.

LEMMA 2. (i) If the number of depositors, n , is large enough, a banker gets a zero payoff from an attempt to offer depositors less than full value d_2/n per depositor at date 2. Even if n is small, the banker's payoff is bounded above by the rent that an intermediary could obtain in bilateral negotiation with a single lender over collecting loans equal in market value to d_2/n . (ii) A bank can commit to paying out all that it can collect at date 2 by issuing demand deposits to sufficiently many depositors that the total promised payment equals the amount that it can collect.

Proof. See the Appendix.

The lemma implies that if, at date 1, the entrepreneur offers to pay $P_2 = X_2$ at date 2, the banker can immediately borrow $S = X_2$ against this promise by issuing demand deposits with face value $d_2 = X_2$. By corollary 2, this implies that the banker can make a loan to the entrepreneur at date 0 that entails no illiquidity premium and requires no liquidation. Thus the banker, by committing her future collection skills using demand deposits, creates liquidity for the entrepreneur.

E. More General Discussion

Let us describe more generally what we have shown so far. Ownership of an asset is the transferable right to make decisions on the use of the asset. Contracts are mechanisms that assign ownership in a particular

way, perhaps contingent on payments.¹⁰ Each period, each contract can be renegotiated. Once there is an assignment of ownership in a period, the owner has the right to make a decision about the use of the asset. We assume that while contracts assign ownership, they cannot directly enforce decisions. In other words, after ownership is assigned, negotiations also take place over the decisions that will be made. During the negotiations, the entrepreneur and the relationship lender can commit to current payments and decisions, but not to future ones. They can also reassign ownership and write contracts over the future. Thus every decision is subject to negotiation.

When there is a single lender, negotiations over the assignment of ownership and negotiations over decisions on the use of the asset collapse into one. When there are multiple lenders, there are two sets of negotiations: one in which ownership is assigned (who gets to bargain with the entrepreneur) and one in which decisions are made (whether or not the asset is liquidated).

It is easy then to see that a fragile capital structure, by itself, is insufficient to increase the payment by the entrepreneur. Suppose that the entrepreneur directly issues demand deposits—where lenders establish ownership only by seizing the asset or a direct claim to it—without going through an intermediary.¹¹ If the entrepreneur attempts to renegotiate, depositors will run and some depositors will seize ownership of the asset. But once ownership is established over the asset, the entrepreneur can ask each depositor not to liquidate in return for a per unit payment of βX_2 , and they will accept. More generally, since ownership can be assigned only to unskilled lenders, we have the following lemma.

LEMMA 3. If the entrepreneur borrows from unskilled lenders at date 1, he will pay only up to βX_2 at date 2.

This implies that a fragile capital structure by itself does not enhance the entrepreneur's ability to commit to paying. Contrast this with a bank, which is a set of contracts in which if the entrepreneur commits to paying P_2 he owns the project. If the entrepreneur does not commit to that payment and the relationship lender commits to paying d_2 per unit to each of many unskilled lenders, then the relationship lender owns the project asset. If neither makes a commitment, each unskilled lender

¹⁰ More generally, contracts assign only temporary claims on ownership; i.e., they assign certain (and not all) rights of ownership for a specific period. For example, having paid off this period's interest, the entrepreneur has the right only to operate his asset over the period. He does not have the right to sell it if debt maturing in the future is secured by it. He also may lose ownership if he cannot pay future debt. None of this is germane to the discussion at date 2, but it is important to qualify the term "ownership" if we discuss earlier periods.

¹¹ Since we have assumed for simplicity that the physical asset is indivisible, depositors seize ownership certificates to portions of the asset.

has the right to seize d_2 per unit he holds. Even though contracts can be renegotiated so that the relationship lender can offer unskilled lenders a lower payment in return for leaving ownership in her hands, unskilled lenders crucially suffer from a collective action problem that deters them from accepting a low offer from the relationship lender. Therefore, a default by the relationship lender leads to a transfer of ownership to the unskilled lenders. It is the prospect of the actual transfer of ownership from the relationship lender to the unskilled lenders in a bank that disciplines the relationship lender and forces her to pay out up to X_2 .

In summary, the problem of illiquidity cannot be solved solely by introducing a collective action problem among the entrepreneur's creditors or solely by having a relationship lender. Instead, it requires both.

F. Liquidity Creation for Depositors over Two Periods

Thus far, we have explained how a relationship lender can meet a need for liquidity at date 1 by issuing demand deposits. But this suggests a more general point. The banker herself may or may not need liquidity, but she can specialize in acquiring relationship lending skills and can make potentially illiquid loans. She can meet the liquidity needs of her financiers by issuing fresh demand deposits to replace any amount they may want to withdraw. This then suggests an additional role for demand deposits. At date 0, the bank can rely exclusively on funding from unskilled investors who may have privately observed and uncertain needs for liquidity at date 1, decoupling the scale of the bank from the banker's own wealth. By issuing deposits to them, which they can withdraw whenever they have a liquidity need, a bank can offer totally liquid claims to investors despite funding loans that are illiquid. Thus deposit claims play a dual role in a bank. They offer an *ex ante* assurance of liquidity to investors who may have a demand for it, and the first-come, first-served mechanism offers a practical way to meet unverifiable needs. *Ex post*, the first-come, first-served mechanism enables the bank to actually provide this liquidity by allowing it to raise new deposits on the basis of the implied commitment to pass through the value it can extract. It is not inconceivable that the economic elegance of this mechanism explains its historical persistence.

Specifically, the bank offers demand deposits with the option to withdraw $d_1 = 1$ at date 1 or $d_2 = 1$ at date 2 per unit held.¹² This matches

¹² In the Jacklin (1987) analysis of the Diamond and Dybvig (1983) model (see also Bhattacharya and Gale 1987; Jacklin and Bhattacharya 1988), a result of this type would suggest that banks could not create liquidity. This is not the case in our model because the specific illiquidity of financial assets implies that a bank can enhance liquidity without offering depositors a return that differs from market rates of return.

the returns and liquidity of storage, and it satisfies the incentive-compatibility constraint imposed by the private information about the need for liquidity.¹³

PROPOSITION 1. A bank offering demand deposits to a sufficiently large number of depositors in which (i) the deposit contract promises equal amounts at both periods ($d_1 = d_2 = 1$) and (ii) the depositors can seize financial assets of market value equal to the promised payment on a first-come, first-served basis if the bank does not pay will commit the banker to collecting the loan on behalf of depositors; that is, all cash flows extracted by the bank from entrepreneurs can be paid out to depositors. The deposits are liquid, although they finance otherwise illiquid loans.

Proof. See the Appendix.

Thus when lemma 2 and proposition 1 are put together, the bank creates liquidity on both sides of the balance sheet, shielding entrepreneurs from the liquidity needs of depositors and offering depositors a liquid claim even though they finance illiquid assets.

G. *Alternative Assumptions on Bargaining*

Before we detail the implications, we should point out that the bargaining games we have assumed can be generalized without any qualitative change in the results. We have assumed that each lender has no bargaining power and receives just his or her outside option, the liquidation value. In principle, we could allow each lender some bargaining power and also allow for different ways by which bargaining leads to a division of surplus. All that is important for our results is that the amount a borrower can commit to paying is an increasing function of the value the lender can get from the collateral assets, which is true of both the commonly used outside option bargaining and Nash bargaining solutions. All our results will go through (except in the non-generic case in which lenders have all the bargaining power—and take all surplus—in which case the borrower's specific ability is not an issue) simply by replacing the payoff from rejecting an offer with the payoff from the alternative bargaining game.

Similarly, the result that a run on a bank will result in full disintermediation is independent of the bargaining game since it is the re-

¹³ The one difference between date 2 and date 1 is that if the banker can set up a bank again after a run, the entrepreneur may not be able to commit to paying depositors as much at date 1 as the banker can in the aftermath of a run. If the entrepreneur calls for renegotiation at date 1 after the bank seeks to negotiate with depositors, the entrepreneur would end up negotiating with the bank even if the bank is run and would (at best) gain nothing (see the proof of proposition 1). Therefore, if the banker threatened not to use his skills on behalf of the depositors at date 1, the entrepreneur would simply make the scheduled payment, and the banker's skills would not be needed.

quirement that property rights be perfected that makes depositors seize the loan or force its sale. Finally, the result that the banker's rents are driven down when she is disintermediated depends primarily on the banker's no longer having control rights over the loan rather than on the sequence or structure of the bargaining. As long as disintermediation changes the ownership of the loan, our results that runs can provide discipline continue to hold.

IV. Implications

Even though our model has looked at banks somewhat traditionally as making loans and offering demand deposits, its implications hold more generally in distinguishing various kinds of financial institutions today. Loans are a metaphor for positions (whether on or off the balance sheet) that are complex so that the bank's skills are needed to manage them to maturity. Demand deposits are a metaphor for guarantees of liquidity, whether they are loan commitments, irrevocable letters of credit, or demand deposits. Our paper then says that the bank has a comparative advantage in both holding complex positions *and* offering guarantees of liquidity because the guarantees precommit the bank to not absorbing too much of the rents created by its ability to manage complex positions. The complementarity is highest when the bank's value addition is largely already embedded in its positions and its skills are used mainly in effecting transfers rather than in creating new value. The bank's rents are then most threatened by a run, and it can thus enjoy cheap financing by issuing demandable claims. Let us now consider some implications from the model.

A. *Distinguishing Financial Institutions*

In relating the institutional structure of the bank to its function, our model also explains how other kinds of financial intermediaries are different. Unlike commercial banks, money market mutual funds do not create liquidity. In a money market mutual fund, each depositor has the right to seize a proportion of assets equal to his proportion of total deposits. In other words, depositor holdings are marked to market so that the mutual fund is run-proof. If indeed mutual fund assets required active intervention by the fund manager, depositors would not be able to discipline her, and the manager would capture a rent. Partly as a consequence, money market funds avoid illiquid assets that have to be managed actively and hold liquid assets passively. Money market mutual funds cannot create liquidity. Depositors get liquidity simply because of the underlying liquidity of the money market fund's holdings and not because the fund adds any liquidity of its own.

Also, unlike commercial banks that provide liquidity on demand, insurance firms provide large payments conditional on some observable and verifiable events such as death. Insurance companies can provide considerable amounts of insurance: their contracts correspond to those that emerge from the mechanism design problem in Section II when liquidity needs are observable. Banks, on the other hand, are limited to a milder form of insurance because the liquidity demand they service is inherently unobservable or unverifiable. Of course, some life insurance companies have partly demandable claims that allow withdrawal of a fixed amount even without the occurrence of the insurable event (and runs do sometimes occur). To the extent that they hold illiquid "relationship" assets, our ideas may be relevant to understanding such life insurance contracts.

Commercial banks are different from investment banks because the value of investment banks is largely in future transactions rather than embedded in current positions. Venture capitalists also play an extensive ongoing managerial role in the start-ups they finance. Because investment bankers and venture capitalists continue to provide valuable services after the initial financing, they cannot efficiently be cut out of the deal. A coalition without the investment banker or venture capitalist generates less total surplus. As a result, demand deposits are unlikely to provide discipline. This is a potential reason why the capital structure of investment banks and venture capitalists is usually different from that of commercial banks.

B. Narrow Banking, Bank Capital, and Deposit Insurance

Our paper suggests that financial fragility built in through demand deposits allows the bank to fund itself at low cost, disciplines bank rent extraction, and enables it to provide liquidity to both depositors and borrowers. Our view contrasts with the more traditional view that resurfaces every few years (e.g., Simons 1948; Bryan 1988) that financial fragility is unnecessary in banks and can be legislated away by requiring "narrow" banks, so that illiquid assets are funded with long-term liabilities and money raised from demand deposits is invested in liquid paper. Our model suggests that such legislation would kill bank liquidity creation and result in a lower credit availability to borrowers.

Of course, by examining only a world with no aggregate uncertainty, we have not investigated the adverse consequences of financial fragility. When we introduce uncertainty, the bank will be run and disintermediated in some circumstances when the project cash flows or project loan values are too low. In deciding her capital structure, the banker now has to trade off liquidity creation against the cost of bank runs. It may be optimal for the bank to partly finance itself with a softer claim

such as capital, which has the right to liquidate, but does not have a first-come, first-served right to cash flows (see Diamond and Rajan 2000). While this allows the banker to extract some rents, thus reducing her ability to create liquidity, it also buffers the bank better against shocks to asset values. These ideas are also relevant to the discussion of whether developing economies rely too heavily on short-term debt (see Diamond and Rajan 2001).

What if the government insures private bank deposits? Complete deposit insurance, in which all depositors are insured, will remove the commitment value of deposits. Insured depositors will not run, and there will be no threat of even partial disintermediation. Unless the deposit insurer has special negotiation or commitment skills, a fully insured bank does not create liquidity (implying that if banks raise deposits in excess of the market value of loans, the excess will be a subsidy provided by the deposit insurer). A fully insured bank has no more commitment ability than an “all-capital” bank: one in which there is no collective action problem among claim holders.

However, a less than fully insured system with some discretion in the ex post government guarantee of deposits may still provide commitment to bankers. In practice, most medium and large banks have many uninsured deposits or deposit-like liabilities, such as large deposits, commercial paper, and interbank borrowing. Since runs by large creditors can bring most banks down, the thrust of our results should hold. Moreover, if the insurer takes time coming to the bank’s aid or is committed to punishing an undercapitalized bank’s management (perhaps by closing or merging the bank), we could get effects very similar to those in the model. Seen in this light, legislative requirements that regulators take “prompt corrective action” when a bank gets into trouble could improve the efficiency with which banks intermediate credit, if they succeed in making regulators play more of a disciplinary role.

For the threat of disintermediation to provide full commitment, all depositors must have the option to withdraw (or at least enough depositors to fully disintermediate the bank), and not just those who need to withdraw for liquidity purposes. This makes it difficult to use suspension of convertibility of deposits into cash to stabilize a bank. If a bank can suspend convertibility of deposits into cash at will, it cannot pay investors any more than an all-capital bank. This gives a new perspective on suspension: it may be a transfer to the bank (effectively, a recapitalization) from depositors, who are now forced to negotiate. Historically, however, banks were allowed to suspend convertibility only when they agreed to do so as a collective because the borrowers in the region could not pay immediately or because aggregate liquidity was scarce, as a result of crop failure or panic. This requirement of regional (clearinghouse) agreement may serve as a way of dealing with aggregate

uncertainty (which is absent from our model but is present in Diamond and Rajan [2000*b*]) without eliminating the commitment role of deposits for individual banks. More work is needed to examine the implications of this view of suspension.

V. Relationship to the Literature

It is useful to contrast our model with the existing literature to get a better understanding of its contribution. Two aspects are particularly important: how the model relates to other papers explaining the structure of banks and how it relates to other papers in which short-term creditors play a disciplinary role.

A. *The Structure of Banks*

A number of papers suggest that a bank is an intermediary that structures its contracts to minimize the risk it shares with uninformed outside investors when it possesses private information about the value of its portfolio (see, e.g., Diamond 1984; Boyd and Prescott 1986; Gorton and Pennacchi 1990). We show that a short-term nonrenegotiable claim can be important even without an uncertain portfolio value, and we stress the significance of fragile deposits rather than just low risk liabilities.

Diamond and Dybvig (1983) focus on the liquidity the bank provides depositors, taking the illiquidity of real assets as given. They provide a reason to finance illiquid assets with demand deposits.¹⁴ We provide an alternative foundation, based on illiquid financial assets, for the assumed illiquidity in their paper. In addition, the bank does more than redistribute the returns from real assets across states and over time; it is also critical in enhancing the returns from real assets over and above what would be available if investors tried to manage them directly. Thus the institutional alternatives that Jacklin (1987) proposes to a bank would not be able to carry out the functions a bank performs in our model.

It is useful to contrast our notion of banks with the recent view put forward by Holmström and Tirole (1997, 1998). In their model, as in ours, firms may be liquidated or denied funding because too small a fraction of their future returns can be paid to outsiders. In their model, ex post unprofitable wealth transfers (provided by an explicit state-contingent contract) to these firms can help them survive. While individuals are assumed not to be able to commit to honoring promises to make the state-contingent payments, an intermediary can hold collateral ex

¹⁴ Other important research on synergies between deposits and loans is presented in Nakamura (1988), Flannery (1994), Myers and Rajan (1998), Qi (1998), and Kashyap, Rajan, and Stein (1999).

ante and thereby commit to making the payments. While the bank in our model also has a role in providing reliable funding (by buffering the firm from the liquidity demands of direct investors), its value to firms comes from the ability to make (or continue) loans that are unprofitable to others but privately profitable. In a sense, the bank enhances the collateral value of the entrepreneur's date 2 asset from βX_2 to X_2 and passes it through. The bank therefore does not merely utilize the economy's collateral better but also enhances it. It provides liquidity because of its ex post ability, and not because of its ex ante state-contingent pledges.

B. The Role of Multiple Creditors and Short-Term Debt

Bolton and Scharfstein (1996) argue that multiple creditors can reduce the entrepreneur's incentive to default strategically because once an entrepreneur has reached agreement with one creditor, he has a greater incentive to reach a deal with the second and pays out more (see also Gertner 1990). Bolton and Scharfstein assume a Shapley value outcome for bargaining, and complementarity, between assets to establish this result. Under our assumption that the borrower can make a take-it-or-leave-it offer to every creditor, he will not lose bargaining power when dealing with multiple creditors, even if assets are complementary, so their effects are not seen.

Berglof and von Thadden (1994) rely on the inaccessibility of later creditors to construct a staggered debt structure that will extract more from the entrepreneur. This will not work in our framework because everyone is present to negotiate. As lemma 3 indicates, no matter what a firm's capital structure, it cannot pay unskilled lenders more than βX_2 at date 2.

This then points to an essential difference between our work and its closest counterpart: Calomiris and Kahn (1991). The authors describe demand deposits with sequential service as a way to provide incentives for outside investors to monitor a borrower. Depositors who are the first to withdraw their capital get paid in full and will have an incentive to anticipate actions of the borrower that would reduce the value of the borrower's assets. Their rush to withdraw prevents a "crime in progress" from occurring and alerts passive outside investors that the borrower will act against their interest. Because demand debt is useful whenever, at a cost, one can predict value-reducing self-interested actions of management, it should be as effective in industrial firms as in banks. Our model, by contrast, emphasizes their special role in the creation of bank liquidity. In addition, we focus on the ex post disciplinary role of deposits in disintermediating the banker whenever some asset losses are revealed or anticipated. In this view, in contrast to that of Calomiris and Kahn,

depositors are not required to undertake detailed monitoring of the bank to anticipate future problems. They need only commit to a sufficiently harsh punishment once problems are revealed. While both roles may be important, ours is more relevant if depositors do not get much more than public information by monitoring or if the punishment provided to bankers by depositors also punishes depositors.

VI. Conclusion

When a lender makes loans that can be collected only with her specific collection skills, the loans are illiquid. The reason is that the lender's specialized human capital cannot be easily committed to collecting the loans; hence they will sell at a discount or will be poor collateral. Mechanisms that commit the lender to collecting the loans create liquidity. A bank is structured to be such a mechanism. If the lender finances using demand deposits, she cannot hold up depositors and, instead, has to pay them the promised amount. Intuitively, the sequential service constraint creates a collective action problem among depositors, which makes them run on the bank whenever they think that their claim is in danger. When the bank has the right quantity of deposits outstanding, any attempt by the banker to extort a rent from depositors by threatening to withhold her specific collection abilities will be met by a run, which disintermediates the banker and drives her rents to zero. Thus the banker will not attempt to extort rents and can commit her human capital to collecting the loans.

An important contribution of this paper is to tie the notion of illiquidity of financial assets to the specific skills a lender may have. These skills may have to do with the knowledge the lender acquires about a particular borrower or with expertise acquired by repeated interaction with a certain class of borrower or with a certain type of contract. If these skills are not widely available, the lender's financial assets are illiquid, with attendant consequences to the borrower. This notion of illiquidity could be useful in other contexts also.

We have largely focused on situations in which the lender's acquired lending skills persist even after a liquidity shock. The lender we have in mind in this situation is a specialized financial intermediary whose main line of business before and after the shock is likely to be credit. In fact, we also examine situations in which such a relationship lender need not get a shock of her own but creates liquidity for others who do.

In addition, however, we have explored situations reminiscent of trade credit, where a firm in a business has an advantage in lending to buyers because it can sell repossessed goods better. It will lose this advantage if it gets a shock that forces it to leave the business. One example of

such a shock is an adverse one such as bankruptcy. In these situations, we have two possibilities. Those who have obtained credit may stop repaying the firm because it has little ability to sell what it repossesses. The outstanding credit will not come down quickly for such a distressed firm, and we shall have the seemingly paradoxical situation of the distressed financing others. Alternatively, the firm may be forced to liquidate credits because it has little extraction ability in the future. In either case, realizations will be low when the firm is distressed, implying that the illiquidity premium associated with trade credit (and other business to business credit) should be high.

Our simple model can be enriched in a number of other directions to provide more policy implications. One obvious direction is to introduce uncertainty about asset values, which will create a role for bank capital (see Diamond and Rajan 2000). Another is to explore what would happen if there were an aggregate shortage of liquidity. This could introduce a role for bank reserves, which is absent in our model thus far. Preliminary work indicates that an investigation along these lines could help identify a set of policies to deal with financial crises that seem far more nuanced than the conventional wisdom. One could go on, but this should be enough to suggest that there is much to be done.

Appendix

Proof of Lemma 2

The argument in the text uses functions that implicitly require depositors to be small. Nothing depends on this, however, and if depositors decide simultaneously whether to run, then any offer from the banker seeking concessions will lead to a run (with full disintermediation and $f = 1$). We now generalize this. The banker can make an initial offer. If the offer is for less than d_2 per unit, a run will start (from the same logic as in the text), and each deposit will be assigned a place in line. We now allow the banker to make a second offer to each depositor as the depositor comes up to the bank to withdraw on the basis of his place in line. If the banker makes an offer that deters a given depositor from withdrawing (seizing or forcing the sale of loans), then the loan will remain in the bank for possible seizure or forced sale by a depositor behind him in line, who has not yet reached an agreement.

The banker must offer each depositor his outside option given the number of loans remaining in the bank that have not yet been seized or sold. The outside option is the smaller of the full value of the deposit and the market value of the remaining loans. If a fraction f_j of deposits have withdrawn before depositor j arrives to withdraw d_2/n , the outside option per unit of deposits is

$$\rho(f_j) = \begin{cases} d_2 & \text{if } f_j \leq \frac{\beta X_2}{d_2} - \frac{1}{n} \\ \beta X_2(1 - d_2 f_{j-1}) & \text{if } f_j \in \left(\frac{\beta X_2}{d_2} - \frac{1}{n}, \frac{\beta X_2}{d_2} \right) \\ 0 & \text{if } f_j > \frac{\beta X_2}{d_2} \end{cases}$$

Any offer that the banker makes must be payable using assets that remain in the bank after all depositors in line have had the chance to withdraw. Consequently, the outside option is a nonincreasing function of place in line. If any depositor is to make a concession, the last in line must have an outside option below d_2/n . Anticipating this, all depositors with $f_j \leq (\beta X_2/d_2) - (1/n)$ will withdraw since their outside option to take assets is worth d_2/n . Let m be the number of depositors who can withdraw in full. The total units of loans seized or sold to meet the needs of these depositors will be $md_2/n\beta X_2$. What remains should be less than what is needed to fully satisfy one depositor, so

$$1 - \frac{md_2}{n\beta X_2} < \frac{d_2}{n\beta X_2}.$$

This implies that $m = \lfloor (n\beta X_2/d_2) - 1 \rfloor$, where $\lfloor \cdot \rfloor$ is an operator that gives the smallest integer greater than the argument.

Once m depositors have withdrawn, the banker can ask the next depositor, for whom there are not enough assets to be made whole, to make a concession. But in order for this depositor to leave the remaining assets in the bank, the bank should be able to convince depositors who follow him not to withdraw. All these depositors should be paid at least the market value of the collateral that is left (since that is their outside option). The total that can be paid to all nonwithdrawing depositors, even with the banker's superior collection skills, is $1/\beta$ times the market value of the assets remaining at the end of the line. So at most $1/\beta$ depositors can be given offers that deter them from withdrawing. Let $\lceil 1/\beta \rceil$ denote the greatest integer less than or equal to $1/\beta$. Then for the banker to convince the remaining depositors not to withdraw, it must be that $n - m \leq \lceil 1/\beta \rceil$. So a run does not fully disintermediate the bank if

$$n - \left\lfloor \frac{n\beta X_2}{d_2} - 1 \right\rfloor \leq \left\lceil \frac{1}{\beta} \right\rceil.$$

Since the left-hand side is increasing in n , there is an n^* such that the bank is fully disintermediated if $n > n^*$. Even if the bank is not fully disintermediated, the banker's rent is small because the bank's remaining assets have a market value of less than the amount of one depositor's deposit. The rent is at most the amount of added value that the banker can collect from one depositor: $1 - \beta$ times the full value of the assets remaining in the bank after the run (and less when there is more than one depositor at the end of the line who does not withdraw and must be paid by the banker). Q.E.D.

Proof of Proposition 1

The entrepreneur offers the bank the lowest-cost loan contract that will induce the bank to make the loan, and we prove below that the bank can pass through all that it collects to depositors. As a result, the bank will break even on its

lending and collect a total market value of payments from the entrepreneur of \$1, offering depositors the chance to withdraw for \$1 at date 1 or 2 ($d_1 = d_2 = 1$). No matter what the fraction, θ , of its depositors who will withdraw for liquidity purposes at date 1, the bank can meet the withdrawal because there is no aggregate liquidity shortage: there is storage in use from date 1 to date 2 after all profitable projects have been funded. Either the bank can borrow $1 + \theta$ units at date 0 and store θ in cash until date 1, or it can borrow one unit at date 0 and issue new deposits of θ at date 1 to finance withdrawals. It is also clear that if the bank does not meet the depositor demand, or is anticipated not to meet demand, it will precipitate a run on the assets. All that remains to be shown is that the bank prefers to make precommitted payments rather than trying to renegotiate and precipitating a run.

We know from lemma 2 that the bank is disintermediated and gets zero rents if it tries to renegotiate at date 2. So it will prefer to make precommitted payments and can commit to passing through all date 2 collections to depositors.

Similarly, if the banker should seek concessions from depositors at date 1, a run will start. This follows because $d_1 = d_2$, and even a depositor who otherwise planned to hold until date 2 would seek to seize assets at date 1 by running. As with the logic used in proving lemma 2, all depositors will run. The consequences of a run at date 1 are only a little more complicated than at date 2. Assume first that the banker cannot reestablish a liability structure with demand deposits immediately after a run. In the aftermath of a run at date 1, if the depositors and the entrepreneur have a first chance to reach a deal (before depositors negotiate with the bank at date 1), the depositors have the option to hire the bank at date 2. The entrepreneur can make current payments up to C_1 and can commit to making date 2 payments equal to what depositors can enforce on their own. Since this is also equal to the greatest amount that the bank can commit to paying them without reestablishing the bank, the depositor will, in fact, deal directly with the entrepreneur, and the banker will be disintermediated after a date 1 run and will extract a zero rent. Thus the banker will not want to precipitate a run by demanding to renegotiate what she owes depositors.

If instead the banker could reestablish the demand deposit structure after a date 1 run, she could commit to paying depositors up to the amount the bank itself can collect at date 2. This is more than the entrepreneur can unilaterally commit to paying at date 2. If so, it is possible that the entrepreneur may not be able to commit to paying depositors as much at date 1 as the banker can in the aftermath of a run. If the entrepreneur calls for renegotiation at date 1 after the bank seeks to negotiate with depositors, the entrepreneur would end up negotiating with the bank and would (at best) gain nothing.¹⁵ Therefore, if the banker threatened not to use her skills on behalf of the depositors at date 1, the entrepreneur would simply make the scheduled payment, and the banker's skills would not be needed. As a result, the banker's threat to not negotiate at date 1 would not cause the value of bank assets to fall, and no run would occur.

So depending on our assumption, the banker's threat to withdraw her human capital at date 1 unless depositors make a concession either precipitates a run and she is disintermediated or does not precipitate a run, but the threat has no bite since the entrepreneur makes scheduled payments. Q.E.D.

¹⁵ This assumes that the loan contract need not be renegotiated if the banker is there to negotiate. We know that there exists a renegotiation-proof loan contract between the entrepreneur and the bank, on the condition that the entrepreneur and the bank are the parties that negotiate, if the bank can commit to paying all collected funds to depositors: this is the contract from lemma 1 with $S = X_2$.

References

- Berglof, Erik, and von Thadden, Ernst-Ludwig. "Short-Term versus Long-Term Interests: Capital Structure with Multiple Investors." *Q.J.E.* 109 (November 1994): 1055–84.
- Bhattacharya, Sudipto, and Gale, Douglas M. "Preference Shocks, Liquidity, and Central Bank Policy." In *New Approaches to Monetary Economics*, edited by William A. Barnett and Kenneth J. Singleton. Cambridge: Cambridge Univ. Press, 1987.
- Bolton, Patrick, and Scharfstein, David S. "Optimal Debt Structure and the Number of Creditors." *J.P.E.* 104 (February 1996): 1–25.
- Boyd, John H., and Prescott, Edward C. "Financial Intermediary-Coalitions." *J. Econ. Theory* 38 (April 1986): 211–32.
- Bryan, Lowell L. *Breaking Up the Bank: Rethinking an Industry under Siege*. Homewood, Ill.: Dow Jones-Irwin, 1988.
- Bryant, John. "A Model of Reserves, Bank Runs, and Deposit Insurance." *J. Banking and Finance* 4 (December 1980): 335–44.
- Calomiris, Charles W., and Kahn, Charles M. "The Role of Demandable Debt in Structuring Optimal Banking Arrangements." *A.E.R.* 81 (June 1991): 497–513.
- Diamond, Douglas W. "Financial Intermediation and Delegated Monitoring." *Rev. Econ. Studies* 51 (July 1984): 393–414.
- . "Liquidity, Banks, and Markets." *J.P.E.* 105 (October 1997): 928–56.
- Diamond, Douglas W., and Dybvig, Philip H. "Bank Runs, Deposit Insurance, and Liquidity." *J.P.E.* 91 (June 1983): 401–19.
- Diamond, Douglas W., and Rajan, Raghuram G. "A Theory of Bank Capital." *J. Finance* 55 (December 2000): 2431–65.
- . "Banks, Short-Term Debt and Financial Crises: Theory, Policy Implications and Applications." *J. Monetary Econ.: Carnegie-Rochester Conf. Public Policy* 54 (Summer 2001), in press.
- Flannery, Mark J. "Debt Maturity and the Deadweight Cost of Leverage: Optimally Financing Banking Firms." *A.E.R.* 84 (March 1994): 320–31.
- Gertner, Robert. "Inefficiency in Three-Person Bargaining." Working paper. Chicago: Univ. Chicago, Grad. School Bus., 1990.
- Gorton, Gary, and Pennacchi, George. "Financial Intermediaries and Liquidity Creation." *J. Finance* 45 (March 1990): 49–71.
- Greenbaum, Stuart I.; Kanatas, George; and Venezia, Itzhak. "Equilibrium Loan Pricing under the Bank-Client Relationship." *J. Banking and Finance* 13 (May 1989): 221–35.
- Hart, Oliver. *Firms, Contracts, and Financial Structure*. Oxford: Clarendon, 1995.
- Hart, Oliver, and Moore, John. "A Theory of Debt Based on the Inalienability of Human Capital." *Q.J.E.* 109 (November 1994): 841–79.
- Holmström, Bengt, and Tirole, Jean. "Financial Intermediation, Loanable Funds, and the Real Sector." *Q.J.E.* 112 (August 1997): 663–91.
- . "Private and Public Supply of Liquidity." *J.P.E.* 106 (February 1998): 1–40.
- Jacklin, Charles J. "Demand Deposits, Trading Restrictions, and Risk Sharing." In *Contractual Arrangements for Intertemporal Trade*, edited by Edward C. Prescott and Neil Wallace. Minneapolis: Univ. Minnesota Press, 1987.
- Jacklin, Charles J., and Bhattacharya, Sudipto. "Distinguishing Panics and Information-Based Bank Runs: Welfare and Policy Implications." *J.P.E.* 96 (June 1988): 568–92.
- Kashyap, Anil K.; Rajan, Raghuram G.; and Stein, Jeremy C. "Banks as Liquidity

- Providers: An Explanation for the Co-existence of Lending and Deposit Taking." Manuscript. Chicago: Univ. Chicago, Grad. School Bus., 1999.
- Myers, Stewart C., and Rajan, Raghuram G. "The Paradox of Liquidity." *Q.J.E.* 113 (August 1998): 733–71.
- Nakamura, Leonard I. "Loan Workouts and Commercial Bank Information: Why Banks Are Special." Manuscript. Philadelphia: Fed. Reserve Bank, 1988.
- Qi, Jianping. "Deposit Liquidity and Bank Monitoring." *J. Financial Intermediation* 7 (April 1998): 198–218.
- Rajan, Raghuram G. "Insiders and Outsiders: The Choice between Informed and Arm's Length Debt." *J. Finance* 47 (September 1992): 1367–1400.
- Sharpe, Steven A. "Asymmetric Information, Bank Lending, and Implicit Contracts: A Stylized Model of Customer Relationships." *J. Finance* 45 (September 1990): 1069–87.
- Shleifer, Andrei, and Vishny, Robert W. "Liquidation Values and Debt Capacity: A Market Equilibrium Approach." *J. Finance* 47 (September 1992): 1343–66.
- Simons, Henry C. *Economic Policy for a Free Society*. Chicago: Univ. Chicago Press, 1948.