Liquidity risk, liquidity creation and financial fragility: A theory of banking

Douglas W. Diamond

Raghuram G. Rajan

University of Chicago

We acknowledge helpful comments from seminar participants at the Federal Reserve Bank of Richmond and thank Heitor Almeida for valuable research assistance. Diamond and Rajan are grateful for financial support from the Center for Research in Security Prices. Rajan also thanks the NSF (grant # SBR 9423645).
Both investors and borrowers are typically concerned about liquidity. Investors desire liquidity because they are uncertain about when they will want to eliminate their holding of a financial asset. Borrowers are concerned about liquidity either because they are uncertain about their ability to raise funds when needed unexpectedly, or because they are uncertain about their ability to continue to retain funding in the future. These different concerns about liquidity cannot be completely unrelated, if for no reason other than that commercial banks seem right at the center of meeting these needs for their customers. Depositors get better access to their funds than they would if they invested directly and earned the same expected return: this is liquidity creation. Borrowing firms too can find the bank to be a more reliable source of funding than another firm or individuals: banks insure borrowers against the liquidity risk that funding will be cut off prematurely. Research over the last fifteen years has separately identified these several demands for liquidity, and the role banks play in meeting each of them, yet the inter-relationship between these demands, and how the institutional structure of the bank deals with it, has been largely left unexplored. Our paper explores these issues, and in the process of integrating the existing results, obtains rather different implications.

We start by examining the borrowing decision of an operating firm. A borrower may have a difficult time obtaining, or retaining, outside funding, because not all future cash flows can be credibly pledged to outside investors. There are several plausible limits to an operating firm’s ability to pledge future payments to outside investors (see Diamond [1991, 1993a, 1993b], Hart [1995] and Hart and Moore [1994]). We follow the approach in Hart and Moore [1994], where the inability of the borrowing firm’s managers to commit to continue providing the value added from their human capital limits what outsiders can extract. This limit implies that the operating firm’s real assets are illiquid so that the firm cannot borrow against, or sell its assets for, the full value of the cash flows generated.  

---

1 When an operating firm’s assets are most valuable under the borrower’s control, the borrower can extract
A borrower who needs to raise a lot of money will need a way to commit to paying a large amount to outsiders. Such a commitment may be possible if he borrows from a lender who learns some details of the business because the lender will know more than uninformed outsiders about the second best use of the assets. For example, an incumbent lender may learn how to divide the physical assets into more saleable lots and offer each directly to a particular buyer that the lender knows assigns a high value to it. Since an incumbent lender can obtain more than if the assets were simply put up for auction, she can threaten the borrower with this greater specific ability to liquidate, and thus extract larger payments from him.

However, the specific cash extraction ability of an incumbent lender can cause liquidity problems if the lender faces an unexpected need for cash – if, for instance, the lender receives an investment opportunity of her own that needs immediate financing but does not yield pledgeable cash flows. The need for cash could force her to contemplate selling the loan, which may be worth little to outsiders who do not have her ability to extract. In the same way as the specific human capital of incumbent management makes a real asset illiquid, the specific ability to extract of an incumbent lender makes a financial asset illiquid; the financial asset can only be sold for less than the future cash flows the incumbent lender could extract, discounted at market rates of interest.

Alternatively, the lender might choose to call or reduce the loan rather than selling it for the low price offered by an outsider. Thus, the possibility that the lender will need cash (liquidity) can cause liquidity problems for both borrower and lender. The form of the financial contract will anticipate these possible ex-post liquidity problems, and the borrower’s cost of, and access to,

some fraction of the value that he adds in excess of the next best use of the assets. This will generally be true, since the borrower can always use his operational control to extract value from the lender when the borrower does not have a sufficient stake in the firm’s future. Thus, moral hazard combined with the limited wealth of a borrower, produce effects and implications that are close to this limit imposed by the inalienability of specific human capital.
capital will also be affected. We characterize the best direct financial contracts when the (loan) contracts themselves may become illiquid.

The lenders’ potential need for liquidity is reminiscent of the random need for liquidity in the banking models of Bryant [1980] and Diamond-Dybvig [1983]. In those models, the timing of the cash flows from physical assets makes them illiquid: they do not deliver adequate cash flows in the short run. Moreover, investors obtain only a low re-sale price for the physical asset, because there is a short-term aggregate shortage of liquidity. The role of banks is then to offer liabilities that have higher short-run returns than would be obtained by direct investment in the asset.

Our situation is, however, different; financial assets are illiquid because of the specific talents of the lender, rather than because of an aggregate shortage of liquidity. What is the role of banks here? The bank can make loans against illiquid real assets, develop specific talents, and yet avoid the costs of illiquidity of the financial assets (the loans) if it can raise cash through other means. Clearly, it obtains some cash from loan repayments, which can be used to meet some liquidity demands. It can also maintain sufficient liquid assets to fund the withdrawals. But most important, if it can commit to passing through all inflows on loans to depositors, it can attract new depositors who are willing to make a deposit that offsets the withdrawals. This allows the bank to maintain continuity and obviates the need to transfer illiquid financial assets to others.

The last solution is not straightforward. Since only the bank will have the specific skills needed to collect the maximal value on its loans, there will be relationship specific rents that, depending on how the deal is structured, could accrue to the banker, reducing the banker’s ability to borrow against loan repayments. That is, the banker has the potential to extract a fraction of the specific value that he creates for the depositors (through higher salaries and perks, for example). However, in our model where there is no aggregate default risk to loans, deposit contracts can be structured to eliminate this hold-up problem. A bank can commit to paying
depositors everything it collects from borrowers (in excess of pre-specified operating costs and normal compensation to bankers). ²

Banks do so by choosing a fragile capital structure – a capital structure that will trigger a financial crisis if the bank should attempt to get concessions from depositors, or if it is anticipated that the bank will not deliver on its promised payments. In other words, banks choose capital structures that could be subject to runs. Of course, in a model such as ours with complete information and no transactions costs, a bank run need not have any disciplinary effect since the banker is still the best person to renegotiate loans with the entrepreneur. However, the run has an effect because dis-intermediation changes the way bargaining is structured when loans are renegotiated. In particular, the banker loses the exclusive right to negotiate the loans, a right that now transfers to the depositors who seize the loans. This weakens the borrower’s incentive to repay loans, but also transfers some of the rent the banker might otherwise collect, away from the banker, to the borrower and the depositors. The threat of loss of this rent if a bank run occurs is what disciplines the banker. As a result, banks stick by their original deals and deliver their promised payments, so no runs actually occur. We do not analyze banking panics based on self-fulfilling fears that the bank will not stick by the deal (as in Diamond-Dybvig [1983]), but plan to do so in future work that generalizes our framework.

Importantly, our approach implies that, unlike banks, firms need not be able to commit to paying out more by adopting a fragile capital structure. A fragile capital structure provides the bank incentives because it threatens the bank’s rents from intermediation, but does not affect output. Hence the threat facing the bank does not involve ex post inefficient outcomes that can be renegotiated away. By contrast, it is always ex-post efficient for the entrepreneur to retain control of his assets, so a fragile capital structure can, at best, have no effect and, at worst, lead to inefficient liquidation. In fact, it is precisely because it wants to avert the risk of liquidation when

² Concern with holdups in banking usually goes the other way: a note passed to a teller threatening violence
an impatient lender wants his money that a firm chooses to borrow from a bank. As a result, the incentive effects of a financial crisis differ for an operating firm and for a bank.

In summary, the bank creates liquidity. The bank’s ability to pledge all of the value of the loans, including the portion that requires bank specific skills, means that it can promise depositors more than the market value of the loans it has made (the value that an unskilled loan buyer would pay for the loans). At the same time, the bank buffers borrowing firms from the liquidity demand of investors without introducing an additional layer of agency costs.

We are not the first to argue that a fragile bank capital structure may be intentional. Previous work has related what we call a fragile capital structure to providing liquidity (Diamond-Dybvig [1983]) and to a mechanism for ensuring outside investors monitor a borrower (Calomiris-Kahn [1991]). The ideas in those models, however, do not apply specifically to banks, but could apply more generally to any borrower, including an operating firm. We discuss the relationship of our paper to these models later.

The framework in this paper can be used to examine the costs of financial distress of banks. It is also useful for considering financial regulation and issues in deposit insurance and bank receivership. Some implications can only be developed in a framework with more uncertainty and with private information. Because these elements are not needed for there to be a liquidity creation role of banks or for the issue of liquidity risk to borrowers and to investors, we omit them. We plan to analyze these complications in future work.

I. Framework

1.1. Agents, Projects, and Endowments.

Consider an economy with entrepreneurs and merchants. The economy lasts for two periods and three dates -- date 0 to date 2. All agents are risk neutral and the discount rate is zero. Each entrepreneur has an idea for a project, where each project lasts for two periods. The project

unless cash is turned over.
has a maximum scale of 1, but can be funded with any investment \( i \) at date 0 where \( i \in (0,1] \). It returns a cash flow of \( C_1 \) at date 1 and \( C_2 \) at date 2 per dollar invested if the entrepreneur contributes his human capital. For every dollar invested, the assets created through the initial investment have value in best alternative use (without the entrepreneur’s human capital) of \( X_1 \) immediately after investment until cash flow \( C_1 \) is due to be produced, and \( X_2 \) from then till \( C_2 \) is due to be produced. At that point the value of the assets collapse to zero. Funds can also be invested at any date in a storage technology that returns $1 at the next date for every dollar invested.

Entrepreneurs do not have money to finance their projects. There are a number of merchants, each with $1 of endowment at date 0. Merchants can lend their money to an entrepreneur. The problem, however, is that a fraction \( \theta \) of merchants get an unverifiable “early” trading opportunity at date 1 which offers a private gross return of \( R \) at date 2 for any amount invested at date 1. \( R \) cannot be pledged or shared with outsiders.\(^3\) At date 0, merchant do not know if they will have an early opportunity.

There are \( M \) merchants endowed with a unit of good, and \( N \) entrepreneurs with projects. To be sure that the investment projects are infra-marginal and that there is no aggregate shortage of liquidity, we assume that \( M > \frac{N}{1 - \theta} \). This implies that there is sufficient endowment to fund all projects at date 0 as well as let fraction \( \theta \) of merchants invest 1 at date 1. It also implies that storage is in use at each date, and the present value of a claim on one unit of consumption at date \( t \) is one at date \( t-1 \).

We assume

\(^3\)More generally, the investment in the trading opportunity could be thought of as the inside equity the merchant has to put up in order to convince others to invest in it. For obvious reasons, this cannot be borrowed or pledged.
\[
Min\{C_1 + C_2, \frac{C_1 + C_2}{X_1}\} > R > 1 > C_1
\]  

so that the entrepreneur’s initial 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 merchant’s trading opportunity. All these investment opportunities, however, dominate storage. The last inequality indicates the entrepreneur cannot pay off debt with date-1 cash inflows. We will assume, for simplicity, that the project generates sufficient cash flow in the long run, i.e., \(C_2 > X_2\) so that illiquidity is never a problem at date 2.

1.2. Specific Learning about the Entrepreneur’s Projects.

The initial lender sees the business as it is put together and learns what it takes to run it. A lender who comes in later never gets to learn as much about the business since much of the trouble-shooting has already taken place. So a lender who lends over the initial period has a better sense of the alternative uses/users for the firm’s assets. Specifically, the initial lender can sell assets for \(X_t\) per dollar invested, while outsiders or later lenders can generate only \(\beta^o X_t\), where:

\[
1 > \beta^o > \frac{1}{R}
\]

Since educating the initial lender takes time and effort, we assume that each entrepreneur can borrow from only one such lender.

1.3. Contracting.

The entrepreneur needs to borrow money. We assume that contracts with investors can specify payments and can make the transfer of ownership of the assets contingent on these payments. Furthermore, we assume the existence of accounting systems that can track cash flows once they are produced. However, an agent can only commit to contributing his human capital to a specific venture in the spot market. Human capital cannot be bought or sold. This implies that agents will
bargain over the surplus that is created when they contribute their human capital, as in Hart and Moore (1994). Ex ante contracts over payments and ownership will constrain this bargaining.

In order to raise money, an agent has to give the investor some, possibly contingent, control rights. We consider contracts which specify the that agent owns the asset and has to make a payment to the investor, failing which the investor will get the right to put the asset to an alternative use/user and get some pre-contracted fraction of the proceeds (possibly one) from doing so. So a contract specifies repayments \( P_t \) the agent is required to make at date \( t \), as well as fractions \( f_t \) of value in alternative use that the investor gets if the agent defaults and the investor exercises her control rights.

Given the sparse contractual possibilities, when there is a single agent and a single borrower this contingent-right contract (which we will term a debt contract) is without loss of generality; the outcomes and allocations cannot be dominated by any other contract when the liquidity needs of the investor are not verifiable. A contract requiring infinite payment is one in which the investor decides the use of the asset unconditionally, while a contract that gives unconditional ownership to the agent gives the investor no means of extracting repayment, and will be dominated.

1.4. Bargaining.

The borrower can attempt to re-contract the payment he has agreed to in the past. We assume bargaining takes the following form; the borrower offers an alternative sequence of \( P_t \) and \( f_t \) from the one contracted in the past. He can also commit to making a current payment if his offer is agreed to, as well as commit his human capital this period. The lender can reject the offer. If so, the offer is no longer on the table. Now, the lender can liquidate the asset. Alternatively, the lender can enter into negotiation with the borrower. The act of entering negotiation implies that the lender foregoes finding an alternative user this period but reserves the right to do so next period. Negotiation takes the following form. Each party gets to make a take-it-or-leave-it offer.
with equal probability. If the offer is accepted, any current payments are made, the borrower contributes his human capital, and assumes control of the assets until the next default. If the offer is rejected, current cash flows are not produced and the lender chooses her preferred use next period. This framework is an adaptation of that in Hart and Moore (1994) and the sequence is summarized in figure 1.

We will also explore situations when multiple agents participate in the bargaining. While obviously differing in details, the essential structure of the bargaining will be as above.

II. The Simple Model of Direct Lending.

Consider an entrepreneur who approaches a merchant for finance. Will the merchant lend? To answer this, we have to examine the bargaining at subsequent dates. Bargaining will differ depending on whether the merchant gets an opportunity. First consider a merchant who receives no trading opportunity.

2.1. Direct lending: Merchant has no opportunity.

If lending is individually rational, linearity ensures the merchant will lend her entire endowment of $1. The entrepreneur invests $i$ of this and keeps $s_0 = (1 - i)$ as cash at date 0. At date 2, the entrepreneur may refuse to make the contracted payment and, instead, initiate bargaining with an offer of a lower payment. In response, the merchant can reject the offer and liquidate the assets to obtain $f_1^2(iX_2 + s_1)$ where $f_t^k$ is the fraction of liquidation value the lender gets at date $t$ in the (potentially renegotiated) contract in force from date $k$ and $s_1$ is the entrepreneur’s cash holdings at date 1. If she does not, the entrepreneur and the merchant make take-it-or-leave-it offers. The entrepreneur will demand the entire cash flow of $iC_2$ in return for contributing his human capital, if he gets to make the offer. The merchant has no choice but to accept, since otherwise the cash flow will not be produced and the assets (though not the cash reserve) will be worthless. Similarly, when she gets to make the offer, the merchant will demand the entire cash flow of $iC_2$
in return for allowing access to the asset. The entrepreneur cannot do better than accept. Since
each agent gets to make the offer with probability $\frac{1}{2}$, the merchant expects $\frac{1}{2}iC_2$ from
bargaining. Since she can also liquidate, the merchant can extract at most
$$\text{Min}[P_2^1, \text{Max}[f_2^1(iX_2 + s_1), \frac{1}{2}iC_2 + f_2^1s_1]]$$
where $P_k^t$ is the date $t$ payment that is in the
(potentially renegotiated) contract in force from date $k$.  

Now consider what happens at date 1. If the entrepreneur initiates re-negotiation, the
merchant can liquidate and get $f_1^0(iX_1 + s_0)$. Alternatively, she can negotiate. If he gets to make
the offer, the entrepreneur offers his human capital and a payment $P_2^1 = \infty$ in return for the cash
$iC_1$ generated at date 1. The merchant can do no better than accept since refusal means that she
only gets the asset and cash reserves (and the right to extract payment for its use at date 2) but the
date-1 cash flow is not generated. This is precisely what the offer gives her since $P_2^1 = \infty$ implies
the entrepreneur will default for sure at date 2 and the merchant will get to determine use at that
date. When the merchant gets to make the offer, she will ask for $iC_1$ and a date-2 payment
$P_2^1 = \infty$. Since the merchant makes the offer with probability $\frac{1}{2}$, she expects to get
$$\frac{1}{2}iC_1 + \text{Min}[\infty, \text{Max}[f_2^0(iX_2 + s_0), \frac{1}{2}iC_2 + f_2^0s_0]]$$
from bargaining. She also has the right to
liquidate. Given the initial contracted payments and liquidation fractions, the merchant will
accept an opening offer at date 1 from the entrepreneur that promises date 1 and date 2 payments
totaling $\text{Max}[f_1^0(iX_1 + s_0), \frac{1}{2}iC_1 + \text{Max}[f_2^0(iX_2 + s_0), \frac{1}{2}iC_2 + f_2^0s_0]]$. Since the maximum

\[4\] Unlike physical assets, any cash reserves at date 2 will not lose value even if the borrower and lender do
not reach agreement in the alternating offer game. Since the lender can seize a fraction of the cash reserve
regardless of who gets to make the offer in the alternating offer game, the value of this fractional cash
reserves is her status quo point in that game.
the entrepreneur can commit to pay at date 2 is obtained by setting $f_2 = 1$, a maximum payment of $Max[iX_2 + s_1, \frac{1}{2}iC_2 + s_1]$ can be left for date 2. In what follows, we will assume that $Max[X_1, \frac{1}{2}C_1 + Max[X_2, \frac{1}{2}C_2]] > 1$ so that the entrepreneur can borrow absent shocks.

Note that cash reserves carried into a particular date by the entrepreneur can give the lender a greater ability to liquidate, and hence enable the entrepreneur to commit to pay more.

**Direct lending: The Merchant Receives an Opportunity.**

What happens if, instead, the merchant gets an investment opportunity at date 1 which requires funding? Since she has no other wealth, and the opportunity cannot be borrowed against, the merchant has to squeeze as much cash as possible out of her loan to the entrepreneur. Suppose the entrepreneur initiates re-negotiation at date 1 when the merchant has an opportunity. The merchant has three options. First, she can liquidate the assets. Second, she can enter into negotiation in which case she foregoes the first option for this period. Third, she can sell the loan. Since the incumbent lender’s opportunity value of cash is $R$, she will choose the permissible action that generates the most in present value terms at this discount rate.

**2.2.1 Present value of Lender’s Options at Date 1.**

The present value of immediately liquidating assets is $f_1^0(iX_1 + s_0)$. If the merchant does not liquidate the entrepreneur’s assets or sell the loan, and enters into negotiations, she cannot do better than get $\frac{1}{2}iC_1 + \frac{Max[f_2^0(iX_2 + s_0), \frac{1}{2}iC_2 + f_2^0s_0]}{R}$. If the loan is sold, the merchant can obtain $\frac{1}{2}iC_1 + Max[f_2^0(\beta^0X_2 + s_0), \frac{1}{2}iC_2 + f_2^0s_0]$ where the amount the buyer pays depends on her lower ability to extract, $\beta^0$. So the most the merchant can extract depends on the maximum of these terms. But whether the merchant actually liquidates when the first term
dominates depends on whether the entrepreneur can offer the merchant commensurate value through other means. The entrepreneur will be willing to commit to paying the merchant all his date-1 cash and reserves, \( iC_1 + s_0 \), so as to avoid liquidation. He will also be willing to make up to the maximum date-2 payments, and this claim can be sold by the incumbent lender at a discounted value for cash (by (2) the merchant prefers selling the loan after collecting date-1 payments rather than retaining it). So if the entrepreneur asks for re-negotiation at date 1, the outcomes can be described as follows:

**Lemma 1:** (a) The entrepreneur’s assets will be liquidated at date 1 if \( P_1^0 > iC_1 + s_0 \) and

\[
 f_1^0(iX_1 + s_0) > iC_1 + s_0 + \text{Max}[\beta^0 iX_2, \frac{1}{2} iC_2].
\]

(b) The entrepreneur will renegotiate debt if either \( P_1^0 > iC_1 + s_0 \) or

\[
 P_1^0 + P_2^0 > \text{Min}[\overline{P}, iC_1 + s_0] + \text{Max}[\overline{P} - iC_1 - s_0)R,0], \text{ and}
\]

\[
 f_1^0(iX_1 + s_0) \leq iC_1 + s_0 + \text{Max}[\beta^0 iX_2, \frac{1}{2} iC_2] \text{where}
\]

\[
 \overline{P} = \text{Max}\{ f_1^0(iX_1 + s_0), \frac{1}{2} iC_1 + \text{Max}\{ \frac{1}{2} iC_2 + f_2^0 s_0, \beta^0 f_2^0(iX_2 + s_0) \}\}. \text{The renegotiated schedule will be that the entrepreneur pays} \ P_1^1 = \text{Min}[\overline{P}, iC_1 + s_0] \text{ and}
\]

\[
 P_2^1 = \text{Max}\{\overline{P} - iC_1 - s_0)R,0\}.
\]

---

5 The most present value the merchant can get with the remaining stub of loan, when she has full control in the future conditional on default (\( f_2 \) is set at its maximum level of 1) is by retaining it if

\[
 \frac{iX_1}{R} \geq \text{max}[\beta^0 iX_2, \frac{1}{2} iC_2]
\]

and selling it otherwise. Since \( \beta^0 > \frac{1}{R} \), the merchant prefers selling.
(c) The entrepreneur will make the prior contracted payment if $P_1^0 \leq iC_1 + s_0$ and, either

$$f_1^0(iX_1 + s_0) > iC_1 + s_0 + Max[\beta^0 iX_2, \frac{1}{2}iC_2] or$$

$$P_1^0 + P_2^0 \leq Min[P_iC_i + s_0] + Max[(P - iC_i - s_0)R, 0].$$

Condition (a) says the entrepreneur will be liquidated if he cannot get enough money to make the required date-1 payment and if liquidation is more attractive to the merchant than any combination of date-1 and date-2 payments the entrepreneur can make, or can generate through a loan sale. Condition (b) says that the entrepreneur will be forced to renegotiate debt if he cannot make the date-1 payment and is not liquidated, or he will want to renegotiate if he can bring down the total payments he has to make. Since the merchant has a higher discount rate, the optimal renegotiated payment will be one where the entrepreneur pays all the cash reserves he brings to date 1, together with the cash he generates at date 1, and thereby makes the least possible payment that satisfies the lender’s threat point at date 2. Finally, (c) says that if the entrepreneur knows that he will be liquidated if he defaults, or that the renegotiated payment schedule will not be more favorable than the current schedule, he will make the required payments.

Condition (a) points to two different sources of illiquidity here that cause the entrepreneur to be liquidated: First, the project may not spew out enough cash in the short run to enable the entrepreneur to pay off the needy merchant. The real asset is illiquid, and this directly affects the entrepreneur’s ability to pay. Second, absent the specific talents of the initial merchant, outsiders may not have much liquidation ability, and will not pay a high price if the loan is sold. Thus the financial asset can be illiquid. Liquidity tends to buffer one agent from another, illiquidity, as we shall see, links their fates.

**Optimal contract at date 0.**

Having determined the outcome of date 1 negotiation, we can examine the optimal contract at date 0. Instead of determining the optimal contract directly, we first determine optimal
payments consistent with the lender’s ability to extract and the borrower’s ability to pay. It is then easily checked that a contract can achieve the outcome. Let us refer to the state in which the merchant gets an opportunity as state \( \theta \) and the other state as state \( \sim \theta \).

Define \( V'_t \) as the cash actually paid by an entrepreneur when the merchant is in state \( s \) at date \( t \). The initial merchant will negotiate the loan at date 1, but if the loan is sold, a new merchant will negotiate at date 2. Let the maximum date-2 payment that the original merchant can extract in state \( s \) be \( VO'_2 = \text{Max}[iX_2 + s_1, \frac{1}{2}iC_2 + s_1] \). Let the maximum amount that a new merchant can extract at date 2 be \( VN'_2 = \text{Max}[\beta_0iX_2 + s_1, \frac{1}{2}iC_2 + s_1] \). From assumption 2, we know that \( VN^0_2 > VO^0_2 \) and \( VN^0_2 \leq VO^0_2 \). We will use this in deriving the set of constraints below.

The entrepreneur’s goal is to maximize \( \Phi \) w.r.t. \( V'_t \), \( i \), \( s_1 \) where

\[
\begin{align*}
\Phi &= \theta \left[ i(i(C_1 + C_2) + (1 - i) - (V^0_1 + V^0_2)) + (1 - \theta)(i(C_1 + C_2) + (1 - i) - (V^0_1 + V^0_2)) \right] \quad \text{if} \quad V^0_1, V^0_2 \leq iC_1 + (1 - i) \\
\text{Subject to:} \\
\theta \left[ iX_1 + (1 - i) - V^0_1 \right] + (1 - \theta)(iX_1 + (1 - i) - V^0_1) &\geq \theta R + (1 - \theta) \\
\theta \left[ i(C_1 + C_2) + (1 - i) - (V^0_1 + V^0_2) \right] + (1 - \theta)(iX_1 + (1 - i) - V^0_1) &\geq iC_1 + (1 - i) \\
\theta \left[ iX_1 + (1 - i) - V^0_1 \right] + (1 - \theta)(i(C_1 + C_2) + (1 - i) - (V^0_1 + V^0_2)) &\geq iC_1 + (1 - i) \\
\theta \left[ i(C_1 + C_2) + (1 - i) - (V^0_1 + V^0_2) \right] + (1 - \theta)(iX_1 + (1 - i) - V^0_1) &\geq \theta R + (1 - \theta) \\
\end{align*}
\]

\[
\theta \left[ i(C_1 + C_2) + (1 - i) - (V^0_1 + V^0_2) \right] + (1 - \theta)(iX_1 + (1 - i) - V^0_1) \leq \theta R + (1 - \theta)
\]

\[
V^0_2 \leq VN^0_2
\]

(maximum the entrepreneur will pay at date 2 in state \( \theta \), after loan sale)

\[
V^0_2 \leq VO^0_2
\]

(maximum entrepreneur will pay at date 2, given no sale in state \( \sim \theta \))

\[
V^0_1 \leq \text{Max}[iX_1 + (1 - i), C_1 + (1 - i)]
\]

(maximum payment possible by entrepreneur at date 1 in state \( \theta \) given liquidity constraints)

\[
V^0_1 + V^0_2 \leq \text{Max}[iX_1 + (1 - i), \frac{1}{2}iC_1 + \max(\beta_0iX_2 + (1 - i), \frac{1}{2}iC_2 + (1 - i))] = E^0
\]

(maximum enforceable payment in state \( \theta \))

\[
V^{-0}_1 \leq \text{Max}[iX_1 + (1 - i), C_1 + (1 - i)]
\]

(maximum payment possible at date 1 in state \( \sim \theta \) given liquidity constraints)
\[ V_1^{\theta} + V_2^{\theta} \leq \max \{ iX_1 + (1 - i)C_1 + \max \{ iX_2 + (1 - i)C_2 \} \} = E^{\theta} \]  
(maximum enforceable payment in state ~\( \theta \))

\[ V_1^{\theta} + V_2^{\theta} \geq V_1^\theta + V_2^\theta \]  
(payments made by borrower greater in state ~\( \theta \))

The constraints on the entrepreneur’s maximization problem are self-explanatory, except perhaps for the last one. This simply states that since loan contracts cannot be explicitly state contingent, the entrepreneur pays weakly less in state \( \theta \). This is because while the initial lender enjoys the same control rights, she is willing to compromise in return for immediacy. This is the only cross-state restriction that we have.

2.3.1. Solution.

While the above program lends itself to standard solution methods, the optimal payments can be seen more directly. The entrepreneur will ensure the initial merchant’s IR constraint is met with equality. Therefore,

\[ V_1^{\theta} + V_2^{\theta} = 1 + \frac{\theta}{(1 - \theta)} R[1 - (V_1^\theta + V_2^\theta)] \]  
(3)

An asset is illiquid in some states if the present value of the asset payoffs when the asset is held to maturity and discounted at market interest rates, is higher than the cash that can be realized against it immediately in those states. The loan’s illiquidity premium is the second term on the right hand side of (3), because it is the increased expected return a borrower must provide in excess of the return provided by an asset whose value to the lender is independent of the state. So if \( V_1^{\theta} + V_2^{\theta} = V_1^{\theta} + V_2^{\theta} \), the asset is perfectly liquid since \( V_1^{\theta} + V_2^{\theta} \) reflects the present value if the loan is held to maturity. The last constraint of the program above indicates that the asset can

---

\(^6\) Implicit in this formulation is that the initial lender will either sell the loan or liquidate the borrower rather than holding it to maturity. Therefore, the initial lender gets cash at date 1 in state \( \theta \) of \( V_1^{\theta} + V_2^{\theta} \).
never have a negative illiquidity premium. Given (3), two intuitive principles govern the entrepreneur’s choice of contract;

(a) Since the entrepreneur values cash flows the same in either state and at either date (because he has no investment opportunity), he would like to commit up front to pay the merchant as much as possible at date 1 in state $\theta$ where the merchant has a higher valuation for cash, i.e., the entrepreneur wants to minimize the illiquidity premium.

(b) The entrepreneur would like to structure the contract such that he does not get liquidated unless this is absolutely necessary to satisfy the lender’s IR constraint.

We can now show that

**Lemma 2:** The entrepreneur cannot add value, and may subtract value by storing at either date.

Proof: Omitted.

The entrepreneur will store at date 0 only if the loan contract cannot protect him against (ex ante) suboptimal liquidation. Since there is no uncertainty about the project, the loan contract will result in date-1 liquidation only if required by the lender ex ante to satisfy her IR constraint. As a result, storage by the entrepreneur at date 0 is undesirable – at best it averts a desired outcome (liquidation), at worst the returns from the project are not realized. Since the merchant can lend $1 at most, and the entrepreneur can invest this entire amount in the project, we can focus without loss of generality at situations where the entrepreneur does not store at date 0.

Storage at date 1 also does not improve matters. It does not increase the merchant’s ability to extract (which is determined by the value of assets coming in to date 1), nor the entrepreneur’s ability to pay (while storage could commit the entrepreneur to paying more at date 2, he could equally well pay the cash out at date 1 and reduce the payment required at date 2). Furthermore,

---

7 It may turn out that for certain contracts that lead to liquidation in state $\theta$, the entrepreneur may want to store after borrowing in order to stave off liquidation even though this is sub-optimal prior to contracting. Contract design can take care of this problem. Thus when the lender wants to liquidate in state $\theta$, she is always better off setting date 1 payments so high that the entrepreneur cannot make them even by storing.
the merchant may have a higher value for the cash at date 1. Since the cash inflow $C_1$ is not
enough to fully pay off the debt, welfare cannot be reduced, and may be increased, if the
entrepreneur does not store between date 1 and date 2.

This lemma allows us to set $s_0=(1-i)=0=s_1$ in the above program. It is then easy to show that

**Lemma 3:**

(i) Let $L = C_1 + \max\{\beta X_2, \frac{1}{2} C_2\}$. The entrepreneur will be financed at date 0 and will
not have to pay a illiquidity premium if $\min[E^{\theta}, L] \geq 1$.

If $\min[E^{\theta}, L] < 1$ and $L \geq E^{\theta}$,

(ii) The entrepreneur will be financed at date 0 but will have to pay a illiquidity premium if

$$E^{-\theta} \geq 1 + \frac{\theta}{(1-\theta)} R(1-E^{\theta})$$  (4)

(iii) The entrepreneur will not be financed at date 0 if

$$E^{-\theta} < 1 + \frac{\theta}{(1-\theta)} R(1-E^{\theta})$$  (5)

If $\min[E^{\theta}, L] < 1$ and $L < E^{\theta}$,

(iv) The entrepreneur will be financed at date 0 but will have to pay a illiquidity premium if

$$E^{-\theta} \geq 1 + \frac{\theta}{(1-\theta)} R(1-L)$$  (6)

(v) The entrepreneur will be financed at date 0 but only with the asset being liquidated in
state $\theta$ if

$$1 + \frac{\theta}{(1-\theta)} R(1-L) \geq E^{-\theta} > 1 + \frac{\theta}{(1-\theta)} R\left(1 - \max(L, X_1)\right)$$  (7)

Note that if date 1 cashflow is uncertain, the entrepreneur may beneficially store cash in order to make
liquidation more state contingent. This is not an issue here.
(vi) The entrepreneur will not be financed at all at date 0 if

\[ 1 + \frac{\theta}{1-\theta} R \left\{ 1 - \text{Max}(L, X_1) \right\} > E^{-\theta} \]  

The intuition is straightforward. \( V_1^\theta \) can be as high as \( C_1 \) without the entrepreneur being liquidated, while \( V_2^\theta \) cannot be more than \( \text{Max}[\beta^0, X_2, \frac{1}{2} C_2] \). Therefore the maximum liquidity \( L \) available at date 1 to satisfy the initial lender without her resorting to liquidation is \( C_1 + \text{Max}[\beta^0, X_2, \frac{1}{2} C_2] \). If the minimum of this and \( E^\theta \), which is the maximum amount that can be extracted in state \( \theta \) (see program), is not less than 1, the loan is liquid and the liquidity premium is zero (see (i)). If the minimum is less than 1, the loan will have an illiquidity premium (see (ii) and (iv)). The premium may rise so high that it cannot be extracted in state \( \sim\theta \). In this case (see (iii)), the entrepreneur may not be financed even though he has enough liquidity to pay at date 1, and even though the merchant can extract full repayment in normal times. The reason for credit rationing is that the lender cannot make up for his inability to extract in state \( \theta \) by extracting in state \( \sim\theta \).

Credit rationing may also arise because the entrepreneur cannot raise enough liquidity to pay at date 1 in state \( \theta \), even though the merchant can extract payment. If the date-1 liquidation value of assets is high enough, the entrepreneur may still be able to borrow by allowing himself to be liquidated in state \( \theta \) (see (v)). This can reduce the premium in state \( \sim\theta \) to an extractable level. However, if even liquidation does not bring down the premium sufficiently, no loan will be made (see (vi)).
**Example:** Let $C_1 + C_2 = 1.7$, $X_1 = 0.57(C_1 + C_2)$, $X_2 = 0.8C_2$, $\beta^0 = 0.7$, $R = 1.6$. The outcomes with the optimal contracts are plotted for different values of $\theta$ and $C_1$ in figure 2. Credit rationing and liquidation are more likely if $\theta$ is high and $C_1$ low.

The initial contract will use four variables, the pre-specified payments $P_1^0, P_2^0$ and the fractions the lender gets on liquidation $f_1^0, f_2^0$ to achieve the payment outcomes. Note that the pre-specified payments provide an element of protection to the borrower since the lender cannot exercise control rights unless the borrower defaults on them. The fractions provide control rights to the lender in case the borrower defaults. By setting payments and control rights appropriately, the contract can achieve a degree of state contingency even if the states cannot be written in to the contracts. We show in the appendix that the optimal outcomes are achievable through contracts, so these are indeed optimal contracts.

To summarize: with direct lending, an illiquidity premium can arise either because the entrepreneur cannot raise enough cash to pay the merchant when the latter needs liquidity, or because the merchant’s ability to extract cash lies in the future and is of little use when she faces the need for liquidity. It is easily checked that the higher is the chance that the merchant will need liquidity, the worse off the entrepreneur is.  

2.4. Borrowing against the loan.

---

8 We have assumed that the lender has the unconditional right to sell the loan to whomsoever she pleases. In practice, she may have to obtain the borrower’s permission to sell the loan (so that the borrower can prevent the lender from selling the loan to someone undesirable such as a lender with greater bargaining power, or to a competitor). Interestingly, such a seemingly innocuous clause can immediately result in an illiquidity premium – not having the unconditional right to sell the loan makes it less “liquid”. The reason quite simply is that the borrower will withhold the right to sell in state $\theta$ until the lender agrees to concessions. The lender cannot sell the loan if bargaining breaks down, and instead has to hold it till date 2. Thus the lender’s ability to extract in state $\theta$ is lower and becomes

$$E_0 = \text{Max}\{ X_1, \frac{1}{2} C_1 + \frac{1}{R} \text{Max}\{ X_2, \frac{1}{2} C_2 \} \}.$$ 

Interestingly, who controls the right to sell has no effect on the illiquidity premium when the lender extracts in state $\theta$ by threatening to liquidate, since the lender’s ability to sell the financial claim obviously has no effect on her bargaining position. In a sense, liquidation of the real asset makes the right to sell the financial claim irrelevant.
Before we conclude our discussion of direct lending, there is one more possibility we have to examine. Thus far, we have assumed that the merchant sells the loan when she needs money, and devotes her time entirely to her investment opportunity. This forces her to lose a substantial amount on the loan sale. Two other possibilities arise if she can continue maintaining her knowledge about the entrepreneur even when conducting her new business. First, she can borrow from a merchant using her loan to the entrepreneur as collateral. Second, she can sell the loan with the promise that she will act as an agent for the buyer. Could the merchant eliminate the illiquidity of the financial asset through one of these means?

The answer is no. It turns out that the merchant’s inability to pledge her future human capital will result in an illiquidity premium even with these means. Since this will demonstrate some of the issues that financial intermediation raises, we will explore these means in some detail.

First, let us examine borrowing against the loan – the initial lender (say merchant A) can promise to pay a fixed amount out of the repayment extracted from the entrepreneur at date 2 in return for a loan at date 1 from a merchant without an opportunity (say C). To make the problem interesting, we assume that \( X_2 > \frac{1}{2} C_2 \) so that the lender’s liquidation skills are valuable. In addition to specifying the amount to be repaid, we must specify C’s rights if A should attempt to renegotiate. Merchant A has no pledgeable collateral other than the loan to the entrepreneur. This implies that if merchant A attempts to renegotiate the repayment due to C at date 2, the lender’s cash flow and control rights specified in the loan contract the entrepreneur has signed simply transfer to merchant C if she so desires. In other words, C can seize the loan.

### 2.4.1. Illiquidity when merchant borrows against the loan.

To determine the payments at date 2, we have to be more precise about the sequence of moves when a loan and a loan against the loan are being negotiated (see figure 3). We assume
that the loans are negotiated separately on a bilateral basis. Similar to the entrepreneur, merchant A can also threaten to withdraw her specific (liquidation) skills at date 2, and thus reduce the amount that can be extracted from the entrepreneur. Obviously, A will make this threat only to obtain concessions from merchant C. The one privilege we assume merchant A, as intermediary, has is of going first, i.e., choosing whom she will negotiate with first. Finally, for consistency, we assume each agent can negotiate only once with any other agent (where by negotiate we imply the initial offer, any potential seizure or liquidation, and the alternating offer game) and each debtor will enter the alternating offer game at most once with a creditor or her agent.\footnote{This last assumption ensures we do not build additional bargaining power in with an intermediary because the debtor can be negotiated with multiple times. Also, note that when the real asset is liquidated, there is no sense in playing out the alternate offer game. When the financial asset is seized by merchant C, she may still want to have the right to engage in the alternate offer game with merchant A so as to extract concessions from the entrepreneur. So long as she has not entered the alternate offer game with A, she still has this option.}

Clearly, merchant A will choose to negotiate C’s claim down first. Else, once negotiations with the entrepreneur are complete, the entrepreneur will have contracted to provide his human capital in the spot market and pay the prescribed amount. If merchant A subsequently refuses to pay merchant C, the latter can simply seize the cash that the entrepreneur pays to merchant A as loan repayment (since cash inflows are entered in the accounts and cannot be stolen).

So merchant A will first turn to merchant C and threaten to not contribute her human capital (i.e., exercise her liquidation skills in negotiating with the entrepreneur) unless C accepts a lower amount. In response, C can seize the loan and negotiate directly with the entrepreneur. Or she can enter into negotiations with A to persuade the latter to use her skills in negotiating with the entrepreneur. If negotiations (over who will negotiate with the entrepreneur) break down, the default is that C negotiates with the entrepreneur.\footnote{There is a difference between the disagreement outcome when a creditor (merchant C) is bargaining with the intermediary (merchant A) and a creditor is bargaining with the entrepreneur. In the former case, the intermediary does not contribute her human capital to negotiations with the entrepreneur, and the creditor} Since they share the additional surplus that A
recovers from the entrepreneur equally, merchant C expects to get $\frac{1+\beta^o}{2} X_2$ from the alternating offer game with A (assuming that the promised payments from the entrepreneur to A, and from A to C, are large).\textsuperscript{11} This is indeed the opening offer of payment that merchant A will make to merchant C, and it will be accepted. Folding back to date 1, it is the maximum merchant A can borrow against her loan to the entrepreneur.

The loan to the entrepreneur is thus illiquid, where the discount is $\frac{1-\beta^o}{2} X_2$ -- the difference between what merchant A can expect to realize from the loan and what she can raise against it today. The discount is not the traditional discount for asymmetric information. Instead, it is a discount imposed because the more skilled lender A cannot commit to put her human capital at the service of merchant C when it is needed at date 2. Note that merchant A will get the discount back at date 2 as a rent for her services, so that she effectively only loses the opportunity to invest this rent at rate R at date 1. So if merchant A has the time to be both a lender and an investor, we have to substitute in our previous analysis $\frac{1+\beta^o}{2} + \frac{1-\beta^o}{2R}$ for $\beta^o$ in the original merchant’s receipts and 1 for $\delta^o$ in the entrepreneur’s payment. With this transformation, the results are unchanged.

Of course, merchant A is unlikely to be as careful about maintaining her liquidation skills if she borrows against the loan, since she knows she has to share part of the value extracted from

\textsuperscript{11} Merchant C can do no better by seizing the loan; once she seizes, the best she can do if she rejects the entrepreneur’s opening offer is to go back and negotiate with merchant A to represent her. She will have to offer the merchant the same rents that she offers above. Knowing this, the entrepreneur will open with an offer of $\frac{1+\beta^o}{2} X_2$, which is no better than the offer merchant A initially makes.
the entrepreneur with C. So this is most likely an upper bound and a reasonable measure of the fraction of value a buyer of the financial claim (or a lender against it) can capture is somewhere between $\beta^o$ and $\frac{1+\beta^o}{2} + \frac{1-\beta^o}{2R}$.

2.4.2. To borrow or sell?

The rent merchant A gets at date 2 stems from a pure transfer, since she can choose to hold merchant C up before extracting the money from the entrepreneur. Therefore, her ability as an intermediary to structure the sequence of negotiations between parties is crucial to her capturing the rent. To see this, consider what would happen if A sold the loan to C at date 1 with the suggestion that A would negotiate on C’s behalf at date 2. Since A cannot commit her human capital, merchant C will have to persuade her with a fee of $\frac{1-\beta^o}{2}$. So the loan can be sold by merchant A at date 1 for $\frac{1+\beta^o}{2} X_2$ -- exactly the amount she can borrow against it. So from the perspective of date 1 flows, merchant A is indifferent between borrowing against the loan or selling it.

However, merchant A will no longer be an intermediary at date 2. Since the entrepreneur is in direct contact with merchant C, at date 2 he will offer to pay $\frac{1+\beta^o}{2} X_2$ to merchant C if she will forget about hiring A to negotiate on her behalf. Since C does not expect to get any more by hiring A, she will agree. So the net effect of a loan sale at date 1 is that A puts the entrepreneur and C in direct contact, and thus gives up the intermediation rents that she would otherwise get. Merchant C does not also benefit from the disintermediation. Instead, the entrepreneur benefits because he has to pay less. So from A’s perspective, it is always better to borrow against the loan to the entrepreneur at date 1, and retain a central position in the subsequent negotiations at date 2, rather than sell the loan.
To summarize the discussion so far: Lemma 3 indicates that direct lending creates a liquidity problem – the entrepreneur is exposed to the merchant’s need for liquidity. This manifests itself ex post in two ways – the merchant may charge an illiquidity premium and may liquidate too often. Ex ante, the merchant may not want to lend if the loan is too illiquid. What we have shown is that simple intermediation by lenders with expertise does not solve the problem of illiquidity. Is there a better way to organize the intermediated borrowing? The next section examines the possible advantages of setting up a bank to provide better access to liquidity.

### III. Commercial Banking.

Consider now a merchant (henceforth called the “banker”) who takes deposits at date 0 from a large number of merchants. Let us normalize the total number (to ensure ease of comparison with the work in previous sections) to 1. Let the banker lend amount $i$ at date 0 and store $s_0 = (1 - i)$ as reserve. Furthermore, let him promise depositors $d_1$ if they come for their money at date 1 or $d_2$ if they come at date 2. We will later show that no contracts are better than these demand deposits. If the bank does not deliver the promised payment $d_t$ and attempts to renegotiate the deal, the depositors are permitted to take the bank’s cash and loans, sell or hold them, and acquire all of the rights and obligations associated with the loans. Specifically, a depositor who did not previously withdraw can take bank assets (i.e., cash and loans) with a market value equal to the promised amount $d_t$, where the market value of a financial claim is the amount that an arms-length investor would pay for the financial claim as determined in section 2. If the amount promised to depositors who claim collateral exceeds the market value of the assets, it will be rationed on a first-come first-served basis. Of course, if the promised payment

---

12 So if $\beta^O X_2 > C_2 / 2$, the market value of claim on the entrepreneur with a promised payment at date 2 of infinity is $\beta^O X_2$.

13 Effectively, the first-come first-served mechanism is a lottery where investors receive a random place in
on deposits does not exceed the market value of assets, there is no need for rationing.

For now, it will be convenient to assume that the banker lends to one large entrepreneur. Later, we can discuss the nuances when he deals with multiple borrowers. We now have to describe how the banker negotiates with borrowers and depositors.

3.2. Renegotiation.

As in the previous section, the intermediary can threaten to not exercise her superior liquidation skills in collecting from the entrepreneur unless depositors make concessions. As before, we allow a borrower to negotiate once with a lender where by “negotiate” we mean the sequence of initial offer, acceptance or rejection, followed by seizure of collateral and/or the bilateral bargaining game. The difference here is that the banker has multiple depositors rather than a single lender. Their incentive to establish claims on the bank’s assets if the bank renegotiates makes all the difference to outcomes.

3.2.1. The deposit contract and re-negotiation.

The intermediary can choose whom to negotiate with first. As before, the banker will negotiate first with depositors before concluding any deal with the entrepreneur. As long as the payments scheduled to be made by the entrepreneur can be enforced only with the banker’s special skills, the banker’s threat to withdraw his human capital is non-trivial; if the banker withdraws, the entrepreneur will be faced with unskilled lenders and will insist on renegotiating down his payments.

If at date 1 the banker does not call for renegotiation, only those depositors who prefer receiving \( d_1 \) to what they expect to get for their claim promising \( d_2 \) at date 2 will seek payment. At \( t=2 \), all remaining depositors naturally come for their money. Alternatively, the banker can announce that he intends to renegotiate and make an initial offer.
Depositors can either accept the new terms and indicate they want to withdraw cash if they want to be paid now rather than in the future, or join a line (with positions allocated randomly) to seize the bank’s assets of loans and cash (a run). All depositors make this decision simultaneously. At the end of this stage, either the bank or the depositor will be in possession of the loan to the entrepreneur.

An alternative interpretation of allowing the depositors to seize assets with market value equal to the amount owed is that the bank is forced to sell loans in the open market to meet withdrawal requests from depositors. All we need is that after a run, the claim on the entrepreneur is directly held by an agent who is different from the bank.

If depositors seized the loan, the entrepreneur can initiate negotiations by making an offer. The subsequent steps follow the sequence that we have already documented. For example, if the depositor seized the loan in the first stage, he can reject the offer and enter into the alternating offer game with the banker to get him to represent the depositor in negotiations with the entrepreneur. Alternatively, the depositor can enter the alternating offer game directly with the entrepreneur. Or as yet another alternative, the depositor can liquidate the entrepreneur’s assets underlying the loan.

### 3.2.2. Feasible Deposit Contracts.

Some results on the nature of the deposit contract are immediate.

**Lemma 4:** Any deposit contract with \( d_2 \neq d_1 \) will be renegotiated.

**Proof:** (Sketch) Depositors can trade, and deposit contract requires no special bargaining ability to enforce, so payouts cannot be different. If \( d_2 < d_1 \), everyone would withdraw at date 1 and those who did not have an opportunity till date 2 would store. If \( d_2 > d_1 \), those who need to withdraw at date 1 will sell their claim to others for \( d_2 \) rather than withdrawing. Given that there is no aggregate liquidity shortage at date 1, this is feasible.
More generally, the lemma says that if market discount rates are different from 1, the present value of withdrawing in the future cannot be different from value of withdrawing today without admitting arbitrage possibilities.\textsuperscript{14}

**Lemma 5:** Demand deposit contracts that allow depositors to seize financial assets with a market value equal to the amount they are owed are renegotiation-proof.

**Proof ; (sketch)** Clearly, depositors cannot be offered less than \( d \) (where the previous lemma allows us to omit the time subscript) else it is a weakly dominant strategy for an individual depositor to seize assets rather than accept the lower offer (though collectively, depositors may be better off accepting the offer). But they also do not have the capacity to get more than \( d \) by seizing financial assets, since the market value of the financial claim indicates what they will get net of any rents paid to the banker.

It now follows that

**Proposition 1:** In a world with no aggregate uncertainty, a bank offering demand deposits that have the following features:

(i) The deposit contract promises equal amounts at both periods

(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

can be structured to be a complete pass through; i.e., all cash flows extracted by the bank from entrepreneurs can be paid out to depositors.

**Proof (sketch):** Whatever the anticipated liquidity profile of depositors \( \theta \), the bank can hold cash reserves such that repayments from entrepreneurs plus reserves are just enough to meet the demands of depositors at both periods. It is also clear that if the bank does not meet the depositor

\textsuperscript{14} In the Jacklin\[1987\] analysis of the Diamond-Dybvig \[1983\] model, 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.
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 never prefers precipitating a run to making payments.

If the run is at date 2, this is straightforward. At date 2, the entrepreneur can make any payment up to $C_2$. After a run, the depositors hold all the claims on the entrepreneur (since the bank was structured so that all repayments were assigned to depositors, and the market value of these repayments is (weakly) lower than depositor claims). And if the entrepreneur does not conclude a deal directly with the depositors, the bank will be able to force the entrepreneur to pay an amount based on its higher liquidation threat. The entrepreneur who is in direct touch now with depositors will offer each depositor what the depositor expects to get if she employs the banker to negotiate on her behalf (i.e., the amount the banker extracts net of the banker’s rent). The depositor will accept as a result of which the banker will get nothing if a run takes place. Hence, at date 2, the banker weakly prefers not precipitating a run – because disintermediation destroys his rents -- and makes promised payments. The bank structure can commit the banker to pass through all date 2 collections to depositors.

The consequences of a run at date 1 are only a little more complicated. Assume first that the banker cannot re-establish a liability structure with demand deposits immediately after a run. In the aftermath of a run at date 1, if the depositors deal directly with the entrepreneur (and do not deal with the bank at date 1), they have the option to hire the bank at date 2. The entrepreneur can make current payments up to $C_1$ and can commit to make date 2 payments equal to what depositors can enforce on their own plus one-half of the excess that the bank could extract (because the threat to hire the banker at date 2 is credible). Since this is also equal to the greatest amount that the bank can commit to pay them, the depositors will, in fact, deal directly with the entrepreneur, the banker will be dis-intermediated after a date 1 run, and will extract a zero rent. Thus the banker will not precipitate a run by demanding to renegotiate what he owes depositors.
If instead the banker could reestablish the demand deposit structure after a date 1 run, he could commit to pay depositors up to the amount the bank itself can collect at date 2. This is more than the entrepreneur can unilaterally commit to pay at date 2. If so, it is possible that the entrepreneur may not be able to commit to pay 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.\footnote{This assumes that the initial contract between the firm and the bank was set up as renegotiation proof, conditional on the firm and the bank being the parties who negotiate.} 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. 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 his human capital at date 1 unless depositors make a concession either precipitates a run and he is disintermediated, or does not precipitate a run, but the threat has no bite since the entrepreneur makes payments.

The role of demand deposits is to ensure that depositors do not make a concession to the bank if the banker threatens not to collect the loans. Any concession is effectively made to the entrepreneur. Thus a bank run dis-intermediates the bank and makes the banker forego any rents from his human capital. Importantly, this is not because a bank run somehow destroys the banker’s skills, but because it destroys his position of centrality.\footnote{The entrepreneur gains from the run since he can save on the rent that would otherwise go to the bank. This leads to the somewhat counter-intuitive result that entrepreneurs like runs. Of course, in a more detailed model where the entrepreneur has re-investment options (see later), the entrepreneur wants his banker to be healthy since he wants to raise new funds and the banker is in the best position to provide them. Also, even if an entrepreneur wants his bank to fail, he may have no way of ensuring it does unless} Finally note that mechanisms which assign collateral as a function of the actual number of depositors who demand their money cannot do better than demand deposit contracts. The demand deposit contract is effective in deterring the banker from demanding concessions, which
is the best any mechanism can do. In contrast an alternative mechanism like a mutual fund, where each depositor has the right to seize the proportion of assets equal to his proportion of total deposits, would not prevent the banker from achieving concessions from depositors. Thus mutual funds cannot provide liquidity.

3.2.3. Liquidity Provision.

In addition to acting as a complete pass-through, the above results also suggest that the bank cannot provide an illiquidity premium to depositors, hence it has to provide them perfect liquidity. The bank wants to pay the maximum possible to merchants in state $\theta$ (who will withdraw at date 1). It is also constrained to have (discounted) deposit payout to be time invariant. So the only possible deposit contract is one that requires a payment of $1 on demand in return for a deposit of $1.

It remains to check whether the bank can eliminate the liquidity premium the entrepreneur pays, and whether it can meet the depositor demands. Unless the bank itself is threatened by a run if the entrepreneur defaults, it can always extract exactly what a merchant can in state $\sim\theta$, $E^\sim\theta$.$^{17}$ Since this has to be greater than one for lending to be at all possible, it is clear that the bank eliminates one reason for illiquidity under direct lending – that the lender’s ability to extract in one state is lower because of her impatience (see lemma 3).

As lemma 3 indicates, there is a second reason for the illiquidity premium under direct lending – the financial and real asset are illiquid so the lender cannot be paid off when in need. Again, the bank eliminates this rationale for the premium; First, the bank aggregates the liquidity demands of all depositors. As a result, it has to come up with $\theta$ at date 1 for every dollar deposited at date 0 rather than $1 with probability $\theta$. Since the entrepreneur is indifferent between a cash outflow at date 1 and date 2, he will be willing to pay $C_1$ at date 1 and the rest at date 2.

---

$^{17}$ So long as the bank has some borrowers whom it can liquidate if necessary, and some unpledged inflows...
So if \( C_1 \geq \theta \), the bank will lend all the money collected from depositors to entrepreneurs at date 0, and this money will all be invested in projects. Furthermore, the real liquidity generated by the projects will be enough to meet the demands of depositors. Even if \( C_1 < \theta \), so long as there is no aggregate liquidity shortage the bank can take in fresh deposits against future cash flows at date 1 and pay withdrawers. Since the bank is a pass-through, such borrowing is possible. Alternatively, the bank can store \( \frac{\theta - C_1}{1 - C_1} \) of every dollar deposited to meet withdrawals at date 1. So long as \( X_1 < 1 \) and storage is in use in the aggregate, this is the most efficient way of meeting depositor demands. We can state the following.

**Proposition 2:** In a world with no aggregate uncertainty, a bank offering demand deposits does not pay its depositors an illiquidity premium nor need it demand an illiquidity premium from entrepreneurs it lends to. Furthermore, the bank eliminates excess liquidation.

**IV. Interpretation and Robustness.**

Clearly, the model we have is a very stripped down version of what a bank does. The most general interpretation of our work is as follows; a commercial bank uses scale and borrower specific human capital to enhance the amount that can be extracted from borrowers. This enables the bank to lend more. The question we then ask is why cannot the banker extract, through higher salaries and perks, the incremental value he contributes. The answer is that once a loan is made in a full information world with no bargaining costs, the banker’s human capital only affects distribution, but not efficiency. Since the banker’s human capital is redundant ex post, the institutional structure of banking can re-allocate property rights and bargaining power in such a way that the banker can credibly be cut out of the rent sharing process if he attempts to renegotiate.

---

over the two periods, it cannot be threatened by any single borrower.
Of course, our analysis would go through even if the banker appropriated some of the rents to his human capital. Furthermore, the explicit, and somewhat stark bargaining process we have detailed should be thought of as a metaphor for a more implicit process by which the division of surplus between borrowers, banker and depositors is determined. It will, however, be useful to briefly examine extensions of the model that can add some richness to the basic framework. Detailed investigation is left for future work.

4.1. Reinvestment Opportunities for the Entrepreneur.

Thus far, we have assumed that the entrepreneur has no reinvestment opportunities. What if, with some probability, he has a reinvestment opportunity at date 1 that returns $R_1 > R$? The problem becomes interesting if the reinvestment opportunity is only partially pledgeable, so that the entrepreneur can commit to repaying the merchant between $1$ per dollar invested at date 1 and $R$.

Consider direct lending. If only the entrepreneur has an opportunity at date 1, the merchant will clearly agree to forego repayment at date 1 in return for a greater payment at date 2. Date 1 cash will be put to its most effective use, i.e., invested in the entrepreneur's opportunity. The precise division of the ex post surplus is a matter of detail that need not concern us. Alternatively, when only the merchant has an opportunity, we have seen earlier that it makes sense for the merchant to be paid as much as possible at date 1. The interesting case is when both have opportunities at the same time. Since the entrepreneur cannot promise to pay the merchant $R$ for every dollar lent at date 1, any repayment the merchant has the power to extract at date 1 will be invested in her opportunity rather than the entrepreneur’s socially optimal one. The more the merchant can extract at date 1, the more funds will be inefficiently invested.

It is then clear that our results remain substantially unchanged if either the entrepreneur’s date 1 project is pledgeable so that he can reschedule payments into the future when socially beneficial, or if the entrepreneur and the merchant get opportunities at different times. In these
cases, the ex ante contract will attempt to give as much as possible to the merchant in state $\theta$, the consequences of which we have examined earlier. However, if the merchant’s investment shock is strongly correlated with the entrepreneur’s, and the entrepreneur’s date-1 project is only partially pledgeable, the ex ante optimal loan contract emerges from a trade off. The illiquidity premium will increase if the merchant is paid less and has fewer liquidation rights at date 1, but investment efficiency will be higher.

Clearly, the bank can still be effective since it can allocate funds at date 1 in a better way than does direct lending. What is potentially different now is that the bank can contribute additional value at date 1 because of its superior ability to lend to the entrepreneur’s new opportunity. Note that the bank will now provide liquidity on both sides of the balance sheet, a point emphasized by Kashyap, Rajan and Stein (1998).

Future lending opportunities that can only be exploited with the help of the banker’s human capital, can be a source of rents that cannot be passed through to depositors. Therefore, the bank may no longer be a complete pass through. Exploring this is left for future work.

4.2. Strategic Behavior by Borrowers Against Bank.

We have assumed a single borrower thus far, and we have not considered the possibility that the borrower will strategically manipulate payments to the bank so as to cause it to fail and, thereby, secure a better deal from depositors. As discussed above, this may not be in the borrower’s interest if only the bank is in a position to finance his date-1 investment. More important, the single large borrower was assumed only for convenience, and in practice, borrowers would be small relative to the bank. But are there still equilibria where mature borrowers, who do not need fresh investment, could collectively cause the bank to fail.

Such strategic behavior is dominated in a number of situations, while it can be avoided in other situations by the appropriate choice of contracts. First, if the cumulative payment each entrepreneur owes the bank is not greater than the date-1 liquidation value of his assets, he cannot
cause the bank to fail. Any attempt to renege on payments will simply be met by liquidation. Therefore even a belief that others will renege will not prompt the individual entrepreneur to renege. Similarly, if the date-1 payment each entrepreneur is required to make is not greater than $\frac{1}{2} C_1$, no entrepreneur will have the incentive to renege. The intuition is that the bank can enter the alternating offer game with all those who renege and, by the Law of Large Numbers, come away with an average of $\frac{1}{2} C_1$. Thus the only way entrepreneurs can collectively cause the bank to fail is if the payment required from each entrepreneur at date 1 to meet the demands of depositors is between $\frac{1}{2} C_1$ and $C_1$.

Even in this case, there are equilibria where entrepreneurs conjecture that few other entrepreneurs will renege, the bank will not fail, and it is sensible to make the scheduled payment. But there are also equilibria where entrepreneurs conjecture that all other entrepreneurs will renege, and it is better to join them, force the bank to fail, and get a better deal with the depositors. If these equilibria are likely, the bank may be forced to lend less ex ante and maintain more reserves, so that all it needs from each borrower to meet the demands of date 1 withdrawers is $\frac{1}{2} C_1$.  

V. Relationship to the Literature.

Many of the issues we raise have been touched upon in other models, so we discuss these to acknowledge our debt, as well as pinpoint our contribution. Diamond [1984] presents a model where there are two possible financial contracts; a contingent contract where an exogenous default penalty can be invoked on a contingent basis by a lender who has monitored the borrower's ability to pay, and simple debt, which is the best contract that can be written without information on the firms' ex-post ability to repay.

18 Interestingly, these conditions resemble those in the depth of the Great Depression when a large number of borrowers sought a moratorium on debt payments, and asset liquidation values were extremely low. Economists argue banks maintained excess reserves, even multiples of what was required to service depositor demands. One explanation is that such a level of reserves was necessary to keep even the healthy borrowers from stopping payments and thus causing the bank to fail.
In his world, firms borrow from banks who monitor firms’ ability to repay, and thus enforce contingent contracts. This avoids the larger default costs that would arise with simple debt contracts. On the liability side, banks can write uncontingent simple debt contracts with depositors. Because the bank is diversified across a large number of borrowers, the probability of default and hence the expected default costs are low. Taken together, the resulting optimal contracts resemble a bank: the bank enforces contingent contracts and issues simple debt. Similar to the model in this paper, deposits bond the bank to pay depositors and to deliver its monitoring services, because it is costly to the banker to default on deposits. However, Diamond [1984] does not model bargaining between banks and borrowers, relying on contingent contracts instead. In addition, he does not model the source of financial distress costs (assuming that exogenous deadweight bankruptcy costs can be imposed as a contingency in contracts). Finally, liquidity does not play a role because it is a one period model.

In Diamond-Dybvig [1983] there is no explicit borrower other than the bank, and the illiquid physical asset of the bank is best interpreted as the bank loan and operating firm aggregated together. That is, Diamond-Dybvig [1983] does not examine any interaction between banks and firms, and focuses solely on the interaction with depositors. In addition, much of the fragility in their model is due to the “sequential service constraint” that withdrawals be processed on a first come first served basis. Wallace [1988] argues that constraints on trade between investors are an important part of the environment that gives rise to endogenous sequential service, but points out that the reasons for these constraints are not motivated in the Diamond-Dybvig [1983] model. In contrast, in our framework, the fragility imposed by the sequential service constraint is required to limit the banker’s ability to hold-up depositors. In addition, the specific loan collection skill of the banker both makes the bank’s loans illiquid and provides an endogenous limit to asset trade. 19

19 See Jacklin [1987] and Diamond [1997] for discussion of the effects of asset trades in the Diamond-
Calomiris-Kahn [1991] describes demand deposits with sequential service as a way to provide incentives for outside investors to monitor a borrower. Those 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 alerts passive outside investors that the borrower will act against their interest. The point they make is important, but partly reflects the more general control property of short-term debt. More important, because demand debt will be used in their model when, at a cost, one can predict value reducing self-interested actions of management, it should apply at least as well to an operating firm as to a bank.

V. Conclusion.

Why do we need another theory of banking? We would like to suggest three reasons. First, with a few exceptions, theories of banking typically focus on specific functions that intermediaries perform, but do not relate these back to institutional characteristics of the commercial bank – the fact that banks make illiquid loans to “difficult” credits, offer demandable liabilities served in sequence, etc. Of course, the institutional form of the bank may be a neutral mutation with no efficiency properties, but the historical resilience of the form, and its popularity in very different economic systems, suggests otherwise. A theory of banking should be able to explain why the bank is better than a insurance company or a mutual fund at performing the specific function being analyzed. Unfortunately, the role played by banks in the literature is often

---

20 A number of recent papers attempt to understand this better. The first set of papers take the bank’s payment and liquidity function as primary, and derive the bank’s credit function as an optimal institutional development that is driven by the former. Myers and Rajan (1998) argue that the historical form of the bank’s illiquid lending business arose to offset the inherent liquidity of the assets it needed to hold to service deposit demand. Kashyap, Rajan and Stein (1998) suggest that the lending business took the form of offering overdrafts and lines of credit on demand in order to best utilize the store of idle liquid assets held to meet deposit demand. A second set of papers take the opposite tack. Calomiris and Kahn (1991), Flannery (1994) and Qi (1998) take the credit function of the bank as primary, and suggest the demand deposit is the optimal institutional consequence.
equally well played by other financial intermediaries, and sometimes, even better by governments, leaving an area of institutional indeterminacy.

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-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 be in terms of 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 pre-commit 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 embedded in its positions, and its skills are used mainly in effecting transfers rather than in creating new value. Commercial banks are different from mutual funds because the latter do not add liquidity to the market value of their holdings. Commercial banks are also different from investment banks because the latter’s value is largely embedded in future transactions rather than in current positions. 21

A second reason for our model is that a theory distinguishing banks from other intermediaries is not simply an exercise in intellectual elegance. An integrated analysis of the various functions banks perform can resolve puzzles posed by theories that have focused more narrowly. For instance, Diamond and Dybvig [1983] focus on the liquidity the bank provides depositors and aggregate together banks and borrowers. The model does not distinguish financial claims on real assets from the real assets themselves. Problems arise if one interprets this to mean that financial claims are exactly as liquid as the underlying real assets, because as Jacklin [1987]

21 This could explain why the rents absorbed by the investment banker are likely to be higher.
points out, if financial claims and real assets can be freely traded, the bank in Diamond and Dybvig [1983] will no longer be able to cross-subsidize depositors. Our model takes the view that banks do provide special services to borrowers so financial claims are also illiquid. As a result, although financial markets do impose constraints on bank operations, bank deposits offer depositors more liquidity than loans held directly.

A third reason for this model is that it does not focus solely on institutions but also on markets. This is important for policy because it provides a natural framework for analyzing changes in the stability of the financial system. The model can address the costs of bank failure and bank financial distress, as well as spillover effects on the rest of the financial system.

It may be useful to contrast our notion of banks and liquidity from the recent view put forward by Holmström-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 to these firms can help them survive. While individuals are assumed to not be able to commit to making future payments, an intermediary can hold collateral ex ante (though any source of ex post wealth would do) and thereby commit to making payments. The pooling of such asset holdings in an intermediary ensures that the collateral assets are used optimally and not left idle ex post (also see Kashyap, Rajan and Stein (1998)).

While the bank in our model also has a role in providing reliable funding (by providing continuity and buffering the firm from the demands of direct investors), its value to firms comes from the ability to make loans that are unprofitable to others but privately profitable. The bank then provides liquidity because of its ex post ability, and not because of its ex ante commitments. Another way of saying this is that in Holmström-Tirole [1998], a firm (they do not consider depositors) has to have contracted ex ante for liquidity because it cannot express its
demand ex post. The ensuing ex post wealth transfer is what creates a specific link between firm
and intermediary. So long as these promises are taken over by another intermediary (or the
government), there is no loss in allowing an intermediary to fail. By contrast, in our model the
incumbent lender’s ability to offer liquidity comes from her specific skills vis a vis the firm. A
transfer of obligations without a transfer of skills will not eliminate the cost of bank failure.

Importantly, the role Holmström-Tirole [1998] suggest for intermediaries could be played
by a mutual fund (as they point out) or, even better, by the government. The reason for this
institutional ambiguity is that they largely abstract from the liability side of institutional balance
sheets, allowing inside equity to bond intermediary promises. By taking the liability side
seriously, we reduce the institutional ambiguity. Banks have the wherewithal to provide more
liquidity in our model than a passive asset holder, but financial fragility is essential for them to
pass this additional liquidity on to depositors. Hence, non-deposit-taking intermediaries such as
mutual funds or finance companies cannot create liquidity. Nor can governments because they are
immune from private threats, unless we assume that government promises are more credible than
private promises for some exogenous reason.

The difference in approach is not inconsequential. Consider the role for capital in our
model. Shocks to a bank’s current cash and its future value will lead to immediate financial
problems, because the bank cannot choose a capital structure that is easily restructured and also
create liquidity (issue deposits in excess of the market value of its assets). Because we decided to
make the model in this paper as simple as possible, there is no aggregate uncertainty and such a
financial crisis will not arise in equilibrium. However, uncertainty in either aggregate liquidity
needs or the value that entrepreneurs can pay will require that the bank issue capital -- which is a
claim not subject to liquidity demands -- and hold its counterpart, liquid reserves.

This is consistent with the observation that bank lines of credit contain “material adverse change
clauses,” that permit the bank to renege on the line of credit if the borrower’s situation deteriorates.
This view of capital and reserves as buffers is, on the one hand, traditional, but on the other, quite different in its detailed implications from that in the recent literature on capital constrained firms and banks. For example, in Holmström-Tirole [1997] inside equity owned by a banker or a borrower is necessary to take up the slack between present value of what the borrower can credibly pay to outsiders and the amount that is needed for real investment. In their model, the limit to what a bank or borrower can pay is moral hazard. By contrast, in our model outside equity can serve as buffer capital. Moreover, a large amount of outside equity does not reduce the depositors’ ability to get liquidity, though it will create some “inside equity” because it will allow bank managers to capture some rents or to operate somewhat inefficiently. How to set the optimal level of capital is an interesting task for future research.
References


Kiotaki and Moore [1997], “Credit Chains,” LSE working paper.


Wallace, N. [1988], “Another Attempt to Explain an Illiquid Banking System: The Diamond and Dybvig model with sequential service taken seriously.” Federal Reserve Bank of Minneapolis Quarterly Review, Fall, 3-16.
Appendix I

Demonstration that outcomes in lemma 3 can be achieved through contracts

We describe the contracts that will achieve the outcomes for each of the six cases. Since these are the optimal outcomes given the contracting environment, the contracts are optimal.

(i) A contract with \( P_1 = 1, P_2 = 0, f_1 = f_2 = 1 \) should achieve the outcome. The entrepreneur can raise enough liquidity \( L \) to pay off the loan, and the lender has enough bargaining power to extract it from him.

(ii) The entrepreneur can raise more liquidity than the lender can extract. So there is no fear of liquidation. He will set \( P_1 = 1 + \frac{\theta}{(1-\theta)} R(1-E^\theta) \), \( P_2 = 0 \), \( f_1 = f_2 = 1 \). Lender will get renegotiated down to \( E^\theta \) in state \( \theta \), and the entrepreneur can raise enough liquidity to pay this.

(iii) No contract is individually rational for the lender.

(iv) \( L < E^\theta \) only when liquidation value is high. The entrepreneur should be forced to pay out all his liquidity, but should be protected from liquidation. Therefore, \( P_1 = C_1 \) and \( P_2 = X_2 \). In state \( \theta \), provided the lender’s ability to extract is high, the entrepreneur will pay \( C_1 \) and the loan will be sold for \( \beta^\theta X_2 \). Thus the lender will get \( L \) at date 1. She has to extract the remaining in state \(~\theta\).

Because the contracted payments are high, she will extract more than she requires for individual rationality to hold. Her power has to be limited by limiting liquidation values. For example, if \( X_1 \) is greater than the amount she has to be paid in state \(~\theta\), then her ability to extract is limited to the requisite amount by setting \( f_1 = \frac{1 + \frac{\theta}{(1-\theta)} R(1-L)}{X_1} \), \( f_2 = \beta^\theta \). Her liquidation threat at date 1 is limited, so her claims will be renegotiated down. If \( X_1 \) is less than the amount she has to be paid in state \(~\theta\), but \( \frac{1}{2} C_1 + X_2 \) is greater, then \( f_1 = 1, f_2 = \frac{1 + \frac{\theta}{(1-\theta)} R(1-L) - \frac{C_1}{2}}{X_2} \).
(v) The entrepreneur is liquidated in state $\theta$. So $P_1 = 1 + \frac{\theta}{1-\theta} R(1 - X_1)$, $P_2 = 0$.

$$f_1^0 = \frac{1 + \frac{\theta}{1-\theta} R(1 - X_1)}{X_1}, \quad f_2^0 = 0.$$ 

(vi) Not individually rational to lend.
Figure 1

Bargaining at date k

Borrower offers alternative current and future payments, $P_t^k$, and collateral, $f_t^k$, for all $t \geq k$. Borrower will not supply human capital this period (date k) if no agreement is reached, but will supply human capital and commit to make the alternative current payment if agreement is reached.

Merchant (original lender) can reject the offer and liquidate for $X_k$, and retain $f_{k-1}^k X_k$. This destroys all future output.

Merchant (original lender) can reject the offer and enter into negotiation

$\frac{1}{2}$

Merchant makes final offer of current and future payments $P_t^k$, and collateral, $f_t^k$, for all $t \geq k$.

Merchant (original lender) can sell loan to new lender who negotiates as above.

$\frac{1}{2}$

Borrower makes final offer of current and future payments, $P_t^k$, and collateral, $f_t^k$, for all $t \geq k$.

Borrower Accepts

Borrower Rejects

Merchants Accepts

Merchants Rejects

Current cash $C_k$ is not produced. Negotiations start again next period.

Current cash $C_k$ is not produced. Negotiations start again next period.
Figure 2: Outcomes with optimal direct lending contracts.
Figure 3
Bargaining with an intermediary.

Intermediary (merchant A) threatens to withdraw her human capital from the bargaining unless her creditor (merchant C) makes concessions. If agreement reached, merchant A will make agreed payments after negotiating with entrepreneur.

Merchant C can refuse, seize loan and negotiate directly with entrepreneur.

Entrepreneur makes an offer to C.

C can refuse and liquidate assets.

½

Merchant C makes final demand for Payment.


Cash not produced this period. All payments made. All payments made.

½

Entrepreneur makes final offer of payment.

Merchant C makes final offer of a fee to A to negotiate on her behalf.

Merchant C accepts. Merchant C rejects.

Merchant A negotiates with borrower as in figure 1. Merchant A negotiates with borrower as in figure 1.”

Merchant C accepts. Merchant C rejects.

Merchant C negotiates with borrower as in figure 1.

Merchant C negotiates with borrower as in figure 1.

* If this part of the tree is entered after C rejects the entrepreneur’s offer, C now has the option of liquidating or negotiating with the entrepreneur but not of accepting the offer.