Liquidity and the Structure of Intermediation

Douglas W. Diamond1  Yunzhi Hu  Raghuram G. Rajan

Chicago Booth and NBER  University of North Carolina  Chicago Booth and NBER

Abstract

We develop a theory of the value provided by intermediaries, how this relates to their capital structure, and further, how both intermediary services and capital vary with fluctuations in prospective corporate net worth or liquidity. Prospective liquidity tends to ease corporate access to funding. With little need for other supports to financing, the essential functions that financial intermediaries perform such as improving corporate governance are crowded out. This tends to make debt returns more skewed for there is little to support debt when liquidity diminishes.

1 Diamond and Rajan thank the Center for Research in Security Prices at Chicago Booth for research support. Rajan also thanks the Fama-Miller Center, the Stigler Center and IGM at Chicago Booth for research support. We are grateful to helpful comments from Jason Roderick Donaldson, Ye Li, Arvind Krishnamurthy, S. "Vish" Viswanathan, Ming Yang, and participants at the 2019 AEA, FIFI, Chicago Booth, Columbia, SFS Cavalcade, FIRS, Rochester, and IMF. A previous version was titled “Liquidity and Securitization.”
In the run up to the financial crisis, the essential functions financial intermediaries played seemed to become less important. Commercial and industrial loans, as well as residential mortgages, all quintessential banking products, were securitized and sold. At the same time, the “skin in the game” intermediaries held to support activities like lending or securitization diminished, while their leverage increased. Some have suggested these developments stemmed primarily from rising agency problems in the financial sector. Instead, we argue they could also stem from rising liquidity in real asset markets. This has important implications for policy.

Let us elaborate. Consider an economy where capable expert managers bid for a cash-flow producing asset, henceforth called the firm. Experts fund their bid with a loan against the firm. The winning bidder becomes the incumbent expert. The other agents in the model are investors and financial intermediaries. Investors are individuals with some personal funds to lend, but who do not have the inclination or ability to engage closely with borrowers. Financial intermediaries, such as banks or sponsors of securitization vehicles, intermediate between investors and experts by raising funds from investors and lending to experts. Neither investors nor financial intermediaries can run firms. We assume that there are plenty of investors around, so their ability to finance firms directly, or via intermediaries, is unlimited.

The size of the loan that an expert receives for their initial bid depends on the debt capacity the firm can support. Lenders have two sorts of control rights, which allow them to be repaid and are the basis for the firm’s debt capacity. First, they have the right to repossess and sell the firm’s assets if payments are missed. This right only requires the frictionless enforcement of property rights in the economy, which we assume. It has especial value when there are a large number of potential buyers in the future, willing to pay a high price for the firm’s assets. Greater future wealth amongst experts outside the firm (that is, prospective non-incumbent expert liquidity, more practically, prospective corporate net worth of other firms in the industry) leads to higher prices in asset resale markets, with less of a fire-sale discount. This increases the availability of this asset-sale-based financing upfront, as in Shleifer and Vishny (1992). Clearly, this kind of control right is exogenous to the firm and depends on economic conditions.

The second right is that lenders can obtain some of the cash flows generated by the asset directly. Unlike asset-based rights, which depend on property right enforcement, cash flow rights are more endogenous; they stem from actions improving firm governance that increase the pledgeability of the firm’s cash flows so that they are more directly appropriable by creditors. Raising pledgeability might entail, for example, improving accounting quality or setting up a stronger board so that the expert cannot divert project cash flows into their own coffers. Higher
pledgeability is a way for the incumbent to increase what can be borrowed directly against the firm’s assets.

A key feature in our model is that the two rights interact. In general, both the higher prospective wealth of non-incumbent experts outside the firm (that is, prospective liquidity) as well as the higher amount of the firm’s future cash flow that a non-incumbent expert can borrow against (that is, higher future pledgeability of the firm’s cash flows) will increase their future bids for the firm. Higher prospective bids will increase debt repayments, and thus the willingness of creditors to lend up front. So higher prospective liquidity and pledgeability increase up front debt capacity.

We assume enhancing pledgeability takes time to set up and is also semi-durable. So the expert incumbent chooses pledgeability one period in advance, and it lasts a period. Consider now the expert incumbent manager’s incentives while setting cash flow pledgeability for the next period, after buying the firm with borrowed money. For this to be an interesting decision, we assume she may have some likelihood of selling some or all of the firm next period – either because she loses ability and is no longer capable of running it, or because she needs to raise finance for new investments.

If the incumbent had no debt claims outstanding, she would undoubtedly want to increase pledgeability, especially if the direct costs of doing so are small – this would simply increase the amount she would obtain by selling the firm to non-incumbent experts if she lost ability. Unfortunately, when the incumbent expert has taken on debt, she will see enhancing cash flow pledgeability as a double-edged sword. The higher future bid from non-incumbent experts also enables the upfront lender to collect more payments if the incumbent stays in control because the lender has the right to seize assets and sell them when not paid in full. In such situations, the incumbent has to “buy” the firm from the lender, by outbidding experts (or repaying the initial loan fully). The higher the probability she will retain ability and stay in control and the higher the outstanding debt, the lower her incentive to raise pledgeability. This means that when high pledgeability is needed for debt enforcement, outstanding debt cannot be very high.

Another way of seeing this is that any incumbent manager always has mixed motives in improving corporate governance (that is, pledgeability) – it improves her access to new external

---

2 Improving accounting quality is not instantaneous because it requires adopting new systems and hiring reputable people. Equally, getting rid of a reputable accountant or changing accounting practices has to be done slowly, perhaps at the time the accountant’s term ends, if it is not to be noticed.
finance but also make existing outside claims stronger. So higher the existing claims relative to new financing needs, lower her incentives to improve governance.

Now consider the effect of prospective liquidity on pledgeability choice. When prospective liquidity is not high, higher pledgeability, as we have seen, helps increase outside expert bids. However, experts will never pay more for the firm in the future than its fundamental value. Therefore, when future expert liquidity is very high, non-incumbent experts will have enough wealth to buy the firm at the full fundamental value without needing to borrow against the firm’s future cash flows. In this case, higher pledgeability has no effect on how much experts will bid to pay for the firm. In other words, *high future liquidity crowds out the need for pledgeability* in enhancing debt repayments. Therefore, we have two influences on pledgeability – the level of outstanding debt taken on to buy the firm and the prospective liquidity of non-incumbent experts.

In normal times, the need to provide the incumbent expert incentives for raising pledgeability keeps leverage and upfront borrowing moderate. As prospective liquidity increases, the incumbent is able to borrow more today to finance the asset, while still retaining the incentive to set pledgeability high. Eventually, though, when prospective liquidity is very high with significant probability – that is, non-incumbent experts will probably have enough wealth to bid full value for the firm in the future without needing to borrow a large amount against its cash flows – debt repayment is enforced largely by the bids made by wealthy non-incumbent experts, and high pledgeability is not needed for them to make their bid. Upfront borrowing will be high, even if it crowds out pledgeability.

Now consider financial intermediaries. Their main role is to improve governance, which allows them to lend more upfront. Specifically, suppose there are ex ante indistinguishable experts who want a loan to make the initial bid, with some of the experts able to raise pledgeability as before, and some completely incapable of doing so. Assume the act of raising pledgeability is observable and verifiable with some noise. So a covenant can be written on it. We examine what happens when the intermediary observes a covenant violation, which signals the incumbent has not raised pledgeability, albeit with noise. The intermediary can monitor after the covenant violation to distinguish experts who simply have no ability to raise pledgeability and correctly tripped the covenant from those who can raise pledgeability but tripped the covenant in error. After monitoring, the financial intermediary can use the covenant violation to raise interest rates and penalize those who cannot raise pledgeability (since there is no adverse effect on their incentives from higher prospective debt payments) while keeping the rate and leverage low for those who tripped it in error. This incentivizes such experts to increase pledgeability. Effectively,
through such monitoring, the intermediary maintains the incumbent expert’s incentive to raise pledgeability while also lending them more than a passive investor can. Of course, the intermediary has to deploy some of its own net worth as skin-in-the-game capital to commit to providing these monitoring services, but it gets an adequate compensation from the fees it can charge the expert for the value it provides.

Depending on the levels of prospective corporate liquidity, the initial equilibrium choice of borrowing will vary. This allows us to describe the prevalence of different forms of lending and intermediation in practice. When prospective expert liquidity is relatively low, the incumbent expert will want to commit to higher pledgeability, which limits how much debt she can take on. It turns out that the low prospective liquidity also limits the extent to which an intermediary can raise rates after monitoring. In this situation, moderate direct lending by investors is the norm. As prospective liquidity rises further, bank monitored finance becomes dominant as it results in the highest amount of up-front lending. In this case, ex-post monitoring conditional on covenant violation is valuable. The intermediary corrects the potential mistake stemming from the noisy signal, while penalizing experts who have no ability to raise pledgeability with higher rates.

As the level of prospective liquidity gets higher still, liquidity substitutes for high pledgeability in forcing debt repayment. This diminishes the value of monitoring and intermediation, and eventually, the actual cost incurred in bank monitoring can outweigh its benefits. At that point, the bank will simply increase the interest rate if the covenant is violated, without monitoring. We call this performance pricing. A few able experts will be penalized when the signal is incorrect, but high liquidity will reduce the loss associated with such errors.

Finally, as the level of prospective liquidity gets very high, the ceiling on debt needed to maintain incentives for pledgeability becomes onerous. Instead, experts will take large loans up front without any intent to raise pledgeability. Intermediaries or even investors will be happy making “covenant-lite” loans without covenants or monitoring. They will rely solely on prospective liquidity for repayment. The intermediaries will effectively become complete “pass through” securitizers, who do not monitor, have no skin in the game, and simply pass through the amounts collected from firm to investor and vice versa. Equivalently, experts will be financed directly by investors through bonds.

Apart from describing when intermediary services are most useful, the model also allows us to discuss the financial intermediary’s capital structure. When it is useful for the intermediary to identify the reliable expert, the intermediary has to make his monitoring credible by having some claims at risk. We describe how the intermediary’s requirement for own at-risk capital changes with the prospective level of future liquidity and thus the demand for the services the intermediary provides. This gives us a theory of the changing demand for intermediary capital (equivalently, a theory of the changing demand for intermediary leverage), where both firm leverage and intermediary leverage are endogenously determined.

In contrast, most of the analysis of intermediary asset pricing and intermediary leverage has studied the effects of variation in the supply of intermediary capital (see Holmström and Tirole (1997), Brunnermeier and Sannikov (2014), Rampini and Vishwanthan (2018) and He and Krishnamurthy (2013) for example). In such models, fluctuations in repayments shock intermediary net worth, which constitutes the supply (or a fixed fraction of the supply) of intermediary capital. Because some types of monitored lending can only occur if the intermediaries have sufficient own net worth, these shocks have pervasive effects of their own (in addition to their direct effects on intermediary net worth). In contrast, our focus on the demand for intermediary capital suggests the fluctuations in its level may be due to fluctuations in the need for intermediation activities.

In particular, when prospective liquidity is very high with significant probability, with little need for intermediation services there is also little need for existing intermediaries to limit leverage and retain much capital as skin in the game. Risky loans to highly leveraged corporates, made by highly leveraged intermediaries, may not be evidence of moral hazard or over-optimism, but may simply be a consequence of high prospective liquidity crowding out intermediation. Such crowding out may, of course, have adverse consequences. As prospective liquidity fades and the demand for intermediation services expands again, the need for intermediary capital also increases. To the extent that intermediary capital is run down in periods when liquidity is expected to be plentiful, it may not be available in sufficient quantities when liquidity conditions turn and demand for capital ramps up. Prospective liquidity breeds a dependence on continued
liquidity for debt enforcement as it crowds out other modes of enforcement, especially governance. This will make debt returns more skewed.

We have associated higher liquidity with high expert/corporate net worth (stemming from an economy-wide boom in the real sector). Higher expert bids could also result from easier monetary policy, easier credit conditions, rapid financial development, lax supervision, or even irrational exuberance in financial markets. Any of these will ease the financing of expert corporations, induce higher future bids, and higher leverage today, which in turn induces lower pledgeability and intermediation. This generalization will be useful in taking the model to the data.

For instance, our model can explain why phenomenon like banks operating with high leverage, securitizations taking place with originators holding little skin in the game, and an explosion in covenant-lite loans, all happened around the same time before the Global Financial Crisis when measures of the ease of financial conditions were sky-rocketing. With plentiful liquidity, there was little demand for intermediation services, so intermediaries tried to transform themselves into pass-through entities. While supervisors could have forced intermediaries to hold more capital, it might simply have moved such lending outside the regulated financial system (as it did before the Coronavirus Pandemic).

We are not the first to describe conditions where intermediary “skin in the game” retention might vary with conditions and possibly be zero, but we show why this may happen during times of high asset valuations. Chemla and Hennessy (2014) presents a signaling model of securitization where retention is zero when asset prices are sufficiently informative about true value, implying that the amount of private information known by securitizers is small. By a similar logic, we get low or zero retention when prospective industry liquidity is high implying little value in providing incentives to securitizers to monitor borrowers. Unless high industry liquidity (high asset valuations) are very highly correlated with informative asset prices, however, the models have very different predictions.

In the rest of the paper, we will formalize our arguments. In section I, we describe the basic framework and the timing of decisions in a two-period model. In section II, we study the firm’s problem in choosing pledgeability, and in section III we examine the intermediary’s decisions of monitoring and lending. In section IV we present a few extensions, relate our paper to the literature, and then conclude in section V.
I. The Framework and Model Setup

Consider an economy with two periods spanning three dates: \( t = 0, 1, 2 \). At date 0, there is an asset, specifically an existing firm, which is up for sale. At date 1, the firm generates interim cash flows \( C_1 \) if the economy is in the good prosperous state \( G \), which occurs with probability \( q \). With probability \( 1 - q \), the economy enters the bad distressed state \( B \), and the firm fails to generate any cash flow. At date 2, the firm will produce final cash flows \( C_2 \) with certainty. Figure 1 illustrates the evolution of the state of nature.

![Figure 1: States of Nature](image)

A. Agents

The economy is populated with three broad groups of agents: expert managers henceforth termed experts, financial intermediaries, and investors. All agents are risk neutral, do not discount the future, and the prevailing gross interest rate is 1. Experts have the ability to produce cash flows with the firm. However, they need to bid against each other for the firm at date 0, as well as possibly at date 1 (see shortly). They can supplement their wealth by borrowing from financial intermediaries.\(^3\) The winning expert becomes the incumbent manager of the firm.

During period 1, the incumbent will learn with probability \( 1 - \theta \) that she will lose her ability in period 2, after date 1 cash flows are produced. In that case, since she has no ability to produce cash flows in period 2, she will want to sell the firm at date 1. Importantly, this offers her a reason

\(^3\) The special role of financial intermediaries can be justified as a way for investors to avoid the duplication of effort, see Diamond (1984).
to increase the date-1 resale value of the firm. $\theta$ is therefore a measure of the firm’s stability – the extent to which the technology or skills needed in the firm are unchanging. Alternatively, the entire model could be reinterpreted as one in which the firm will need to raise essential additional interim financing with probability $1 - \theta$. In this case, $\theta$ is the extent to which the firm can finance itself from internal sources. The event of losing ability is publicly observable but not verifiable and cannot be written into contracts. If it happens, we assume there are plenty of non-incumbent experts around at date 1 whose skills match the firm’s needs and who can bid for the firm.

For much of this paper, we refer to the financial intermediaries as banks for simplicity (we will discuss other interpretations of them later). Banks make loans that may have covenants, and they can monitor borrowers at a cost. Banks have a higher opportunity cost of financing than ordinary investors, in that they have a scalable investment opportunity that returns $R > 1$ at date 1. This assumption implies banks want to raise as much financing as possible from investors against the loans they have made (and thus use as little of their own capital funds as possible). Equivalently, one can assume that banks are less patient or have a limited amount of own capital and want to utilize it as intensively as possible. This gives them a shadow cost of any additional capital invested today that exceeds the market interest rate. $R$ is thus a rough proxy for persistent intermediary capital constraints. For most of the paper, we will focus on $R \to 1$.

Investors have deep pockets. They are willing buy any security as long as they break even in expectation.

**B. Payment Enforcement and Cash-flow Pledgeability**

In general, a bank has two ways of enforcing repayment by the incumbent manager. First, the bank automatically gets paid the “pledgeable” portion of the cash flows produced over the period, up to the amount of the face value of its claim. Second, just before the end of the period, the bank gets the right to seize and auction the firm to the highest expert bidder if it has not been paid in full. This allows it to extract repayment either by the threat of, or by actually, seizing and auctioning the firm. In this auction, both other experts and the incumbent manager are allowed to bid. Implicitly, we assume the incumbent can always bid using other proxies, so contracts that ban her from participating in the auction are infeasible.

We define cash flow pledgeability as the fraction of realized cash flow that can be verified by a court and therefore goes directly to satisfy the lender’s claim. Without loss of generality (because quantity units can be rescaled), we assume the interim cash flows produced during
period 1 are not pledgeable, i.e., $\gamma_1 = 0$. However, during period 1, the incumbent can set pledgeability $\gamma_2 \in [\underline{\gamma}, \overline{\gamma}]$ for the cash flows $C_2$ produced during period 2, where $0 < \gamma < \overline{\gamma} < 1$. The range of feasible values for pledgeability $[\underline{\gamma}, \overline{\gamma}]$ is determined by the economy’s institutions supporting corporate governance, both operating within the firm (such as better auditors, more transparent subsidiary structures, contracts, and accounting, etc.) and through outside institutions (such as regulators and regulations, investigative agencies, laws and the judiciary). Pledgeability can be raised by adopting more informative accounting practices, hiring better accountants, setting up escrow accounts for cash flows, simplifying corporate organizational structures and enhancing their transparency, or putting in place better governance structures such as a more expert and independent board (see Rajan (2012)). Essentially, we can think of raising pledgeability as closing off tunnels, which divert cash flows generated in the firm.

Clearly, the process of improving pledgeability – for example, selecting a reputable and effective accountant -- takes time. The incumbent manager has to invite applications, do due diligence to screen out unsuitable ones, interview the final candidates, and select one at the end. The required actions to improve pledgeability can change over time. Therefore, we assume the incumbent can only affect pledgeability one period ahead. For the same reason, we assume that increasing pledgeability $\gamma_2$ from $\underline{\gamma}$ to $\overline{\gamma}$ requires the incumbent’s continuing effort until the state is realized. If she abandons the process before that time -- for instance, reverting to her existing pliable accounting firm -- her efforts will be in vain, and pledgeability will remain low, i.e., $\gamma_2 = \underline{\gamma}$. In contrast, if she stays the course and finally selects the reputable accountant, she will boost the firm’s accounting transparency next period. Because the accountant cannot be fired without arousing suspicion, this ensures pledgeability stays high next period.\footnote{If next period’s incumbent can reduce pledgeability somewhat from $\overline{\gamma}$ to $\overline{\gamma} - \eta$, the model goes through with $\overline{\gamma}$ substituted by $\overline{\gamma} - \eta$.}

We assume the experts start with no wealth or net worth at date 0, i.e., $\omega_0 = 0$. Let $\omega_{1,s_1}^I$ be the incumbent’s wealth in state $s_1$ at date 1, where $\omega_{1,G}^I = C_1$ and $\omega_{1,B}^I = 0$. Let $\omega_{1,s_1}^E$ be the state-$s_1$ wealth of other experts who do not own any firm. In the spirit of Shleifer and Vishny (1992), we term this state-contingent net worth of industry experts the anticipated future or prospective “liquidity” at date 1. The wealth of these non-incumbent experts (who work
elsewhere when not running a firm) is augmented when the economy is in state G, so 
\( \omega_1^{E,G} > \omega_1^{E,B} = 0 \). Prospective liquidity is exogenous to the model and driven by the economic environment. It will play a key role in determining firm leverage, monitoring, and security issuance.

Since date-1 cash flows are assumed not pledgeable, the key to debt enforcement at date 1 (and thus the amount the incumbent can borrow at date 0) will be the amount other experts will bid for the firm at date 1. To the extent these experts need to borrow against future pledgeable cash flows to add to their liquidity (that is, their personal wealth) in making their bids at date 1, their date-1 borrowing will, in turn, depend on \( \gamma_2 \), the pledgeability of period-2 cash flows.

Finally, experts can be of two types and they differ in the cost incurred in raising pledgeability \( \gamma_2 \). A reliable expert incurs a small cost \( \epsilon \geq 0 \) at the end of date 0 to set \( \gamma_2 \) to any level above \( \gamma \). Throughout the paper, the result will be presented in the limiting case \( \epsilon \to 0 \), as we focus primarily on the benefits of pledgeably (which can be negative), rather than the physical costs to reliable experts. By contrast, we assume the cost of raising pledgeability for an unreliable expert is sufficiently high that she will never do so. The two types of experts can be thought of as having different character traits or scruples. Even the highest quality governance structures can be undermined if the character of the incumbent expert is found unreliable or wanting. We capture this by assuming the cost of binding the unreliable expert is disproportionately high. An equivalent assumption would be that she can set up an ostensibly high pledgeability governance structure, but it would be ineffective in actually directing cash flows to the lender.

### C. Financial Contracts and Intermediation

The fraction of reliable experts in the population is known to be \( \mu \in (0,1) \), but no one (including the expert herself) initially knows the expert’s type.\footnote{Our results stay unchanged if the expert knows her type, because one cannot get unreliable borrowers to self-select. The assumption is for simplicity.} At the beginning of date 0, each expert applies to one bank for funding. We assume the financial contract between the expert and financiers (bank or investors) is a one-period loan contract, which may contain a covenant. At date 0, the financier lends \( l_0 \) in return for which the expert promises to repay \( D_1 \) at date 1. The
inability to write state-contingent contracts can be justified by assuming the aggregate state 
\( s_1 \in \{G, B\} \) is observable but not verifiable.

Once an expert wins the date-0 auction and becomes the incumbent, she can start the process 
of enhancing pledgeability. If she is unreliable, she will find it impossible (which is when she 
discovers she is unreliable). Even if reliable, she may not want to start. However, if she starts 
increasing pledgeability, a noisy and verifiable signal \( \phi \) about \( \gamma_2 \) is realized, before the state of 
the world is revealed at the end of the period. We assume a binary signal structure \( \phi \in \{\phi^H, \phi^L\} \) 
without loss of generality. If \( \gamma_2 = \gamma \), either because the incumbent is unreliable or because she 
has not increased pledgeability, the signal is \( \phi = \phi^L \) with probability \( 1 \). If \( \gamma_2 = \bar{\gamma} \), the signal is 
\( \phi = \phi^H \) with probability \( 1 - e \) and \( \phi = \phi^L \) with probability \( e \), where \( e \) is the probability the 
signal is in error. We assume \( e < 0.5 \) so that the signal is informative. Let 
\[ \pi = \frac{\mu e}{\mu e + (1 - \mu)} \]
be 
the posterior probability of an expert being reliable conditional on \( \phi = \phi^L \).

The loan covenant, if written into the loan, is conditioned on the signal. The covenant is 
violated if and only if \( \phi = \phi^L \), upon which the bank has the right to increase \( D_1 \), the face value 
(and effectively the interest rate) of the loan. The loan contract may itself specify the interest rate 
if the covenant is tripped, which we term a performance pricing loan contract.

In addition, we assume the financial intermediary could pay an additional cost \( \psi \), monitor 
the incumbent, and perfectly learn her type and her action. In this case, she will decide whether to 
increase the interest rate or not based on the information she obtains from monitoring. Note that 
the information from monitoring is privately observed by the bank and therefore is not verifiable. 
We term this a monitored loan.

**D. Timing and Initial Conditions**

The timing of events is described in Figure 2. After funding the project at date 0, the 
incumbent expert begins to set \( \gamma_2 \), the pledgeability for period 2’s cash flow. The noisy and 
verifiable signal \( \phi \in \{\phi^H, \phi^L\} \) is realized, and the covenant (if there is one) is violated if \( \phi = \phi^L \). 
In this case, the bank may monitor the incumbent to learn her type and may increase the face

---

\(^6\) A probability less than one is easily handled, with additional complexity but little additional insight.
value of the loan subsequently. The incumbent finishes raising pledgeability (or not). Then the aggregate state $S_1$ is realized, and the cash flows (if any) are produced. Subsequently, the incumbent’s ability in period 2 becomes known to all. At date 1, the incumbent either pays the remaining debt due or enters the auction. The period ends with potentially a new incumbent in control.

**Figure 2: Timeline and Decisions**

II. Liquidity, Leverage and Pledgeability Choices

We first examine how the face value of debt affects the amount can be raised under different date-1 liquidity conditions. The analysis in period 2 is straightforward. Any non-incumbent expert who buys the firm at date 1, or the incumbent who retains ability and holds on to the firm, can only commit to repay $D_2 = \gamma_2 C_2$ in period 2, where $\gamma_2$ is the pledgeability set at date 0. As a result, they can borrow up to $D_2 = \gamma_2 C_2$ when bidding for control at date 1. During period 2, there is no distinction between a reliable and an unreliable incumbent, since no further pledgeability choice will be made.

We now proceed to the analysis during period 1, with the focus on how the promised payment $D_1$ affects the incumbent’s pledgeability decision. Some parametric restrictions are needed to focus on the most interesting case.

**Assumption 1:**

a. $\omega_{i, G}^{L} \geq \omega_{i, G}^{E, G}, \omega_{i, R}^{L} = \omega_{i, R}^{E, R} = 0$

b. $\omega_{i, G}^{E, G} + \gamma C_2 > \overline{\gamma} C_2$
Assumption 1a requires that in both states, the incumbent has weakly more wealth than non-incumbent experts at date 1. Assumption 1b ensures the difference in non-incumbent expert wealth (that is, prospective liquidity) between the two future states is large enough that regardless of choice of pledgeability, repayment is strictly more in future state G than in future state B.

The enforcement of the debt contracted at date 0 is driven entirely by the face value contracted $D_1$ and the non-incumbent expert’s bid for the firm in a possible date-1 auction. This is determined by $\omega_{E_1}^{E,s_1}$, her wealth in state $s_1$, as well as what she can borrow against future cash flows, which is determined by $\gamma_2$. Since the value of the future cash flows is $C_2$, a rational expert’s date-1 bid for the firm will not exceed that. It will be

$$B_1^{E,s_1}(\gamma_2) = \min\left\{\omega_{E_1}^{E,s_1} + \gamma_2 C_2, C_2\right\}.$$  

Similarly, the maximum the incumbent can bid is

$$B_1^{I,s_1}(\gamma_2) = \min\left\{\omega_{I_1}^{I,s_1} + \gamma_2 C_2, C_2\right\}.$$  

Comparing $B_1^{I,s_1}(\gamma_2)$ and $B_1^{E,s_1}(\gamma_2)$, we see that the incumbent will outbid non-incumbent experts whenever she has (weakly) more wealth ($\omega_{I_1}^{I,s_1} \geq \omega_{E_1}^{E,s_1}$), since both parties can borrow up to $\gamma_2 C_2$ if needed. Under Assumption 1a, the incumbent can retain control in both states by outbidding experts in any possible date-1 auction. Since the continuation value of the firm, $C_2$, is identical for the incumbent and experts, the incumbent is always willing to retain the firm if she retains ability. To do so, she either pays the amount of debt outstanding or outbids other experts. That is, she pays

$$\min\left\{D_1, B_1^{E,s_1}(\gamma_2)\right\} = \min\left\{D_1, \omega_{E_1}^{E,s_1} + \gamma_2 C_2, C_2\right\}.$$

Many of our results stem from this expression, so it is worth noting a few points. First, so long as outside expert bids are below $C_2$, the greater the wealth of outside experts, that is, their prospective liquidity, $\omega_{E_1}^{E,s_1}$, the greater will be the bid of experts, and the greater will be the debt face value that can be enforced. Similarly, the greater the pledgeability $\gamma_2$ chosen, the greater again the date-1 bid, and hence the enforceability of debt payments. However, no rational bidder will pay more than the residual value of the firm, $C_2$. So when liquidity is sufficiently high (that is, $\omega_{I_1}^{I,s_1} \geq (1 - \gamma_2)C_2$), higher pledgeability is no longer needed to enhance debt capacity – bidders have enough wealth of their own to make a bid for full value, without borrowing any
more than the minimum pledgeable cash flows of the asset, $\gamma C_2$. In other words, high liquidity can crowd out the need for pledgeability. We will use all this in what follows.

Let $V^{I,k_1}(D_1, \gamma_2)$ be a reliable incumbent’s payoff when she chooses $\gamma_2$, given the debt payment $D_1$. In state $s_i$

$$V^{I,k_1}(D_1, \gamma_2) = \theta \left( C_2 - \min \left\{ D_1, B^{E,s_i}_1(\gamma_2) \right\} \right) + (1 - \theta) \left( B^{E,s_i}_1(\gamma_2) - \min \left\{ D_1, B^{E,s_i}_1(\gamma_2) \right\} \right) - \varepsilon \Delta_{s_1}.$$  

The terms on the right-hand side are straightforward. With probability $\theta$, the incumbent retains her ability and needs to pay $\min \left\{ D_1, B^{E,s_i}_1(\gamma_2) \right\}$ to retain control and receive cash flows $C_2$ in period 2. With probability $1 - \theta$, the incumbent loses her ability, in which case she has to sell the asset at price $B^{E,s_i}_1(\gamma_2)$, repay creditors $\min \left\{ D_1, B^{E,s_i}_1(\gamma_2) \right\}$, and keep the remaining proceeds. A cost $\varepsilon$ is incurred whenever she sets pledgeability $\gamma_2$ above $\gamma$.

The incumbent faces a tradeoff in raising pledgeability. A higher $\gamma_2$ (weakly) increases the amount the incumbent has to pay the financier when she retains ability and control, therefore (weakly) decreasing the first term, while it (weakly) increases the amount the incumbent gets in the auction if she loses ability, thus (weakly) increasing the second term. In choosing to increase $\gamma_2$, the incumbent therefore trades off being forced to make higher possible repayments – when she buys the firm back from the lender conditional on retaining ability – against the higher possible resale value when she sells the firm after losing ability. More generally, the incumbent trades off the cost of the boost to the value of existing claims on the firm against the benefit from the boost to the value of future new claims. The higher the stability $\theta$, the more the disadvantages loom large relative to the benefits, and higher is the moral hazard associated with raising pledgeability.

The level of contracted debt also clearly shifts how the incumbent sees this tradeoff. The incumbent’s benefit from choosing high versus low pledgeability if state $s_i$ is known to be realized for sure is $\Delta_{k_1}^i(D_1) = V^{I,k_1}(D_1, \gamma_2) - V^{I,k_1}(D_1, \gamma)$, which (weakly) decreases in $D_1$.

The reason is straightforward. If the incumbent retains her ability, she has to pay more on the outstanding debt when she raises pledgeability, and the higher the outstanding debt, the more this is. Similarly, if she loses her ability, she gets the residual value after the selling the firm, and
higher the outstanding debt, the less this is when she raises pledgeability. So higher outstanding
debt reduces the incumbent’s incentive to raise pledgeability. Proposition 2.1 summarizes the
incumbent’s incentive from state $s_i$ for any given $D_1$.

**Proposition 2.1:** Under Assumption 1,

1. A reliable incumbent’s net benefit from choosing high pledgeability in state $s_i \in \{G, B\}$
   \[
   \Delta_i^{s_i}(D_1) = \begin{cases} 
   -\theta [B_i^{E,s_i}(\overline{\gamma}) - B_i^{E,s_i}(\underline{\gamma})] - \varepsilon & \text{if } D_1 > B_i^{E,s_i}(\overline{\gamma}) \\
   \theta B_i^{E,s_i}(\underline{\gamma}) + (1-\theta) B_i^{E,s_i}(\overline{\gamma}) - \varepsilon - D_1 & \text{if } B_i^{E,s_i}(\underline{\gamma}) < D_1 \leq B_i^{E,s_i}(\overline{\gamma}) \\
   (1-\theta) [B_i^{E,s_i}(\overline{\gamma}) - B_i^{E,s_i}(\underline{\gamma})] - \varepsilon & \text{if } D_1 \leq B_i^{E,s_i}(\underline{\gamma}).
   \end{cases}
   \]

2. There exists a unique threshold $D_1^{IC}$ such that the incumbent sets high pledgeability if
   and only if $D_1 < D_1^{IC}$.

3. An unreliable incumbent manager will always choose low pledgeability: $\gamma_2 = \underline{\gamma}$.

The derivation is straightforward (see Lemma 2 in Diamond, Hu, and Rajan (2020), where we
also cover cases in which Assumption 1 is violated). Let us graph $\Delta_i^{s_i}$ as a function of $D_1$ as
described in Proposition 2.1.

![Figure 3: The net payoff to high pledgeability](image)

For $D_1 \leq B_i^{E,s_i}(\underline{\gamma})$, debt repayment is not increased by higher pledgeability because the
face value of outstanding debt is low. Instead, higher pledgeability only increases expert bids,
which is beneficial to the incumbent when the incumbent loses ability and sells the asset. The
benefits of high pledgeability are $(1-\theta) [B_i^{E,s_i}(\overline{\gamma}) - B_i^{E,s_i}(\underline{\gamma})] - \varepsilon$, which is the incumbent’s
expected enhancement in firm sale value by setting pledgeability high versus setting it low. When $D_1$ rises above $B_i^{E,s_1}(\gamma)$, the incumbent has to pay more in expectation to debt holders when she raises pledgeability. So as the face value of debt increases further, $\Delta_1^{s_1}(D_1)$ falls to zero and then goes negative. When $D_1 > B_i^{E,s_1}(\gamma)$, the incumbent has to pay the entire increment in sale price from increasing pledgeability to debt holders when she loses ability – she gets nothing from increasing pledgeability under those circumstances. At the same time, she has to pay debt $B_i^{E,s_1}(\gamma)$ instead of $B_i^{E,s_1}(\gamma)$ if she retains ability. Hence there is no benefit but only cost to the incumbent of increasing pledgeability, and the expected cost is capped at

$$\theta \left[ B_i^{E,s_1}(\gamma) - B_i^{E,s_1}(\gamma) \right] - \varepsilon.$$

Recall that if liquidity in the G state, $\omega_i^{E,G}$, gets sufficiently high such that $\omega_i^{E,G} \geq (1-\gamma)C_2$, non-incumbent experts can pay the full price of the asset $C_2$ even with low pledgeability. In that case, both $B_i^{E,G}(\gamma)$ and $B_i^{E,G}(\gamma)$ equal $C_2$, and $\Delta_1^{G}(D_1) = -\varepsilon$ for any $D_1$. Put differently, when liquidity crosses the threshold of $(1-\gamma)C_2$ in state G, no incentive to raise pledgeability can come from that state. For lower levels of $\omega_i^{E,G}$, i.e., if $\omega_i^{E,G} < (1-\gamma)C_2$, Proposition 2.1 implies there is a maximum debt level for each state where the incumbent has the incentive to set pledgeability high, if that state were to occur with certainty. That debt level, $D_1^{G,PayIC} = \theta B_i^{E,s_1}(\gamma) + (1-\theta)B_i^{E,s_1}(\gamma) - \varepsilon$, is obtained by setting $\Delta_1^{s_1}(D_1^{G,PayIC}) = 0$. Note that the higher the probability the incumbent retains ability, $\theta$, the higher the moral hazard associated with pledgeability, and the lower is $D_1^{G,PayIC}$. Since $\omega_i^{E,G} > \omega_i^{E,B}$, it is easily checked that

$$D_1^{G,PayIC} > D_1^{B,PayIC}.$$

These state-contingent incentive constraints allow us to determine the condition for a reliable incumbent to increase pledgeability, given that she chooses before the period-1 state is known. The risk-neutral incumbent will choose high pledgeability for any given $D_1$ if and only if

$$q\Delta_1^{G}(D_1) + (1-q)\Delta_1^{G}(D_1) \geq 0,$$

where $Q$ is the probability of state G. Let $D_i^{IC}$ be the value of $D_1$
which makes this weak inequality equal zero. At $D^{IC}_1$, when pledgeability is raised a little, the expected (across the two states) increase in debt payments when the incumbent retains ability equals the expected increase in proceeds from selling the firm when she loses ability. Since $\Delta x_s$ is weakly decreasing in $D$, it must be that $D^{IC}_1$, the threshold of debt below which high pledgeability is incentivized before the incumbent knows which state occurs, lies between $D^{B,PayIC}_1$ and $D^{G,PayIC}_1$. One important case is if $\omega^{E,G}_1 \geq (1 - \gamma) C_2$, when higher pledgeability does not enhance expert bids in the G state because the expert has enough wealth to bid full value with pledgeability at $\gamma$. All the incentive to raise pledgeability then comes from state B so that $D^{IC}_1 = D^{B,PayIC}_1$. Importantly, very high prospective liquidity reduces the incentive compatible level of debt.

This implies the maximum expected payments that a borrower can commit to repay may not be achieved by $D^{IC}_1$, the highest promised payment that provides incentives for high pledgeability. Even with low pledgeability choice, the incumbent is able to credibly promise expected repayment of $L = qB^{E,G}_1(\gamma) + (1 - q)B^{E,B}_1(\gamma)$ at date 1. By contrast, to incentivize high pledgeability, the promised payment cannot exceed $D^{IC}_1$, which will imply an expected repayment of $\bar{T} = qD^{IC}_1 + (1 - q) \min\{D^{IC}_1, B^{E,B}_1(\gamma)\}$. If $B^{E,G}_1(\gamma)$ is much larger than $D^{IC}_1$ (either because liquidity in the G state is high or the moral hazard associated with pledgeability is high so that $D^{IC}_1$ is low) and if the probability of the good state $q$ is sufficiently high, the incumbent could commit to more repayment (and thus raise more) by setting $D_1 = B^{E,G}_1(\gamma)$.

The broader point is that the prospect of a highly liquid future state not only makes feasible greater promised payments, but these high promised payments also eliminate incentives to enhance pledgeability. To restore those incentives, debt may have to be set so low that funds raised are greatly reduced – something the incumbent will not want to do if she needs to compete

---

7 Let $D^{IC}_1$ be the highest value if there are multiple solutions to the equation $q\Delta^q_1(D_1) + (1 - q)\Delta^g_1(D_1) = 0$, which only happens in a zero-measure parametric space.
to buy the firm at date 0. Note that this can happen even if the probability of the low state is significant, and even if the direct cost $\epsilon$ of enhancing pledgeability is infinitesimal or even zero.

**Corollary 2.1:** Under Assumption 1, the date-1 face value that enables the incumbent manager to repay the most is either $D_1 = B_1^{E,G}(\gamma)$ or $D_1 = D_1^{IC}$. If

$$q(1-\theta)[B_1^{E,G}(\bar{\gamma}) - B_1^{E,G}(\gamma)] > (1-q)\theta[B_1^{E,B}(\bar{\gamma}) - B_1^{E,B}(\gamma)],$$

$$D_1^{IC} = D_1^{G,PayIC} = \frac{(1-q)\theta}{q}B_1^{E,B}(\bar{\gamma}) - B_1^{E,B}(\gamma) > B_1^{E,G}(\gamma).$$

Otherwise,

$$D_1^{IC} = D_1^{B,PayIC} = \frac{q(1-\theta)}{1-q}B_1^{E,G}(\bar{\gamma}) - B_1^{E,G}(\gamma) < B_1^{E,B}(\gamma).$$

The difference between $\bar{I}$ and $\underline{I}$ shows the value of high pledgeability, which is crucial to our analysis for the rest of this paper.

**Corollary 2.2:** $\bar{I} - \underline{I}$ decreases with $\omega_1^{E,G}$.

### III. Liquidity, Securitization, and Intermediation Services

In the previous section, we described the effect of a bank loan’s outstanding face value as well as the experts’ prospective liquidity on the incumbent’s incentive to increase pledgeability. If a bank would like to encourage a reliable expert to choose high pledgeability, it needs to keep $D_1$, the face value of the loan, below $D_1^{IC}$. If $D_1$ is set higher, even the reliable expert will not choose high pledgeability, and the most any expert will repay is $B_1^{E,G}(\gamma)$ and $B_1^{E,B}(\gamma)$ in state G and B, respectively. In this section, we turn to the analysis of date 0 and study the optimal lending package that enables a competitive bank to lend the most upfront. We will compare different packages, including the use of covenants and bank monitoring. The key tradeoffs will stem from the choice between the endogenously determined support to repayment (higher pledgeability) and the exogenous support (higher prospective liquidity). It turns out that the interaction between these is essential for covenants and monitoring to have bite. This is because covenants do not prevent any actions, they only provide incentives to take actions.

#### A. Securitization and the Bank’s Liability Structure

We assume without loss of generality that the bank sets up a separate capital structure for each loan it makes (equivalently, the bank makes only one loan). Results stay unchanged when a bank
can issue securities backed by an entire pool of loans, since there is no residual risk that could benefit from pooling before tranching.

Given that banks have an alternative investment with return $R > 1$, they would like to fund the entire loan from outside investors. However, banks may need to retain a claim on the loan to commit to monitor the incumbent upon covenant violation. We assume any claim they retain is residual and junior – so they issue a debt contract to investors with face value $F$. Equivalently, the payoff to investors is $\text{Min}[F, x]$, where $x$ is the cash flow collected from expert’s loan repayments. Note that we assume the security $F$ is sold for $P(F)$ before the signal $\phi$ is realized and any potential covenant violation occurs.\(^8\)

\[
\begin{array}{ccc}
\text{Incumbent} & \xrightarrow{\ell_0} & \text{Intermediary} \\
\text{Expert} & x & \xleftarrow{P(F)} \text{Investors}
\end{array}
\]

\textit{Figure 4 Securitization Structure}

\section{B. Lending Packages}

In what follows, we first describe the kind of loans that will be made up front by the bank to the expert. We consider four types of lending packages, which differ in the role of the bank and the resulting pledgeability choices. Subsequently, we will see which package will raise the most and will dominate, given prospective liquidity conditions.

\subsection{B.1 Internal Governance}

The bank can make the initial loan at face value $D_1 = D_1^{IC}$ without a covenant and rely on the incumbent expert’s own incentive, if she is reliable, to raise pledgeability.

If $B_1^{E,G}(\gamma) \leq D_1^{IC}$, an expert is able to raise $l_0^H = \mu \ell + (1 - \mu) \ell$. With probability $\mu$, the expert turns out reliable and can repay $\ell$; otherwise, she is unreliable and can only repay $\ell$.

\(^8\) When the bank securitizes its loans, this can be thought of as shelf registration where the bank files statements with the Securities and Exchange Commission (the “SEC”) to register securities it will issue. In practice, the entire securitization package is typically announced before the underlying loans are originated. For example, more than 90 percent of the agency MBS trading is on a to-be-announced (TBA) basis in which the buyer and seller decide on general trade parameters, such as coupon, settlement date, par amount, and price, but the buyer typically does not know which pools will actually be delivered until two days before settlement (Vickery and Wright, 2013).
If prospective liquidity gets sufficiently high such that $D_i^{IC}$ falls below $B_i^{E,G}(\gamma)$, the expert is able to raise an initial amount of $I_0^H = \bar{T} - (1 - \mu)(1 - q)[D_i^{IC} - B_i^{E,B}(\gamma)]$: the expert is able to repay $\bar{T}$ unless she is unreliable and the realized state is bad. In such a case, she is able to repay only up to $B_i^{E,B}(\gamma)$ as opposed to $D_i^{IC}$.

We call this case “internal governance” since the bank plays no role in enhancing governance (indeed, as we will see, it cannot). Instead, the level of debt issued is low enough that the reliable expert has an incentive herself to raise pledgeability or governance on her own. Since the bank’s services are not needed, it needs no skin in the game (see later). It can lend at $D_i = D_i^{IC}$, issue a security at $F = D_i^{IC}$ and effectively sell the entire loan to investors, retaining no part for itself.

**B.2 Covenant-lite Loans**

Consider next a loan with the face value $D_i = B_i^{E,G}(\gamma)$, again without a covenant.

If $B_i^{E,G}(\gamma) > D_i^{IC}$, even reliable experts have no incentive to set high pledgeability. This package enables the expert to raise $I_0^H = L = qB_i^{E,G}(\gamma) + (1 - q)B_i^{E,B}(\gamma)$ upfront. Of course, if $B_i^{E,G}(\gamma) \leq D_i^{IC}$, it makes sense for the face value of the loan to be set at $D_i = D_i^{IC}$ if the intent is to maximize up-front issuance, hence we are back to the previous internal governance package.

In this package as in the previous one, the loan has no covenant. However, the debt level here is so high that there is little internal governance or pledgeability either. Debt repayment is supported by liquidity. With high levels of debt, low internal governance, and no covenants, it is apt to term this the covenant-lite loan or package.

Note that in the case of covenant-lite loans, the bank again plays a passive role and is essentially a pass-through vehicle: therefore it issues a security $F = B_i^{E,G}(\gamma)$ and sells the entire loan to investors. We could think equivalently of the expert issuing a covenant-lite bond directly to investors, bypassing the bank. Indeed, market participants typically see covenant-lite levered
loans as substitutes for low-covenant bonds.\textsuperscript{9} Typically, these loans will entail higher levels of debt than loans that foster internal governance.

\textbf{B.3 Bank Monitoring}

The next two packages involve the use of covenants. Under bank monitoring, the bank issues an initial loan at face value $D_i = D_i^{IC}$, embedded with a covenant that will be violated if the signal turns out low, i.e., $\phi = \phi^L$. Recall that the signal is noisy and thus can be low for two reasons, either because the expert has not raised pledgeability or because the signal is in error. Upon covenant violation, it may be worthwhile for the bank to monitor and figure out whether the incumbent has raised pledgeability or not. Clearly, monitoring is only useful if the bank is selective – it does not raise the face value (interest rate) after finding the incumbent is reliable, thus providing her with incentives to raise pledgeability despite an erroneous signal, while raising the face value if she is unreliable. If $B_{i}^{E,G}(\gamma) \leq D_i^{IC}$, the bank cannot credibly collect any more by raising the face value of the loan above $D_i^{IC}$, and therefore cannot be selective. In such a case, monitoring is of no value, and we are back to the internal governance package.

If $B_{i}^{E,G}(\gamma) > D_i^{IC}$, monitoring will be particularly valuable if the probability of error, $e$, is high, or equivalently $\pi = \frac{\mu e}{\mu e + (1 - \mu)}$, the posterior probability of an expert being reliable conditional on $\phi = \phi^L$, is high. On seeing the covenant violation and after monitoring, the bank will keep $D_i = D_i^{IC}$ unchanged if the incumbent is reliable but increase $D_i$ to $B_{i}^{E,G}(\gamma)$ if she is not.\textsuperscript{10}

Interestingly, $B_{i}^{E,G}(\gamma) > D_i^{IC}$ if liquidity is high. So the two modes of enforcement interact here – the bank keeps rates lower initially at $D_i^{IC}$ to enhance incumbent incentives to raise

\textsuperscript{9} For example, “The absence of a financial maintenance covenant for the benefit of the term loan lenders is the core feature of a cov-lite loan. Cov-lite loans also often have other borrower-favorable terms that make the restrictions on the borrower more like high-yield bonds than traditional loan transactions with full covenant packages.” See page 2 in https://www.paulweiss.com/media/3978887/goodison_wagner_practicallaw_aug2019_update.pdf

\textsuperscript{10} Note that the reliable incumbent can reduce pledgeability from $\gamma$ to $\gamma$ if she wants to at any time until the state is revealed. Thus, the bank will not always increase $D_i$ upon having private information.
pledgeability and thus enhance repayment, but after seeing the covenant violated, monitors and raises rates on the unreliable, using high liquidity to enforce that higher repayment.

Since monitoring is costly and unobservable, the bank needs skin-in the game to monitor. Otherwise, it might not monitor and might instead always increase $D_i$ upon covenant violation. Indeed, the bank finds it optimal to monitor if and only if

$\left( D_i^{IC} - F \right) - \frac{\psi}{\pi} \geq q \left[ B_i^{E,G} (\gamma) - F \right] + (1 - q) \max \left\{ B_i^{E,B} (\gamma) - F, 0 \right\}.$  \(^{11}\)

On the left-hand side, $\left( D_i^{IC} - F \right)$ is the bank’s net proceeds from monitoring and finding out the expert is reliable. $\frac{\psi}{\pi}$ is the expected cost of monitoring. The right-hand side is the payoff from simply increasing the face value from $D_i^{IC}$ to $B_i^{E,G} (\gamma)$ without monitoring for all those who violate the covenant (the unreliable will be treated similarly under both monitoring and no monitoring and hence drop out from both sides). The set of $F$ that satisfies the constraint is non-empty if and only if $D_i^{IC} \geq 1 + \frac{\psi}{\pi}$, which is the case if we evaluate the constraint at $F = 0$. In general, the constraint reduces to

$$F \leq F^{\text{max}} = \frac{D_i^{IC} - \frac{\psi}{\pi} - qB_i^{E,G} (\gamma)}{1 - q}$$

It is easily verified that $F^{\text{max}}$ falls between $B_i^{E,B} (\gamma)$ and $D_i^{IC}$ so that the bank retains some skin-in the game in the loan repayment.

The expert is able to raise $l_0^{\text{EL}} = \mu T + (1 - \mu) \underline{l} - [\mu e + (1 - \mu)\psi]$ with monitoring. With probability $\mu$, the reliable expert is able to repay $\overline{T}$; otherwise, the unreliable expert repays $\underline{l}$. The bank incurs the monitoring cost $\psi$ with probability $\mu e + (1 - \mu)$.

\(^{11}\) There is also an IC constraint which requires the bank to monitor as opposed to leave the face value unchanged at $D_i = D_i^{IC}$. This constraint can be expressed as

$q \left[ B_i^{E,G} (\gamma) - F \right] - \frac{\psi}{1 - \pi} \geq q \left[ D_i^{IC} - F \right]$, which is always slack under Assumption 2b below.
B.4 Performance Pricing

Finally, the bank can issue a loan at face value $D_1^{IC}$ and choose to always increase $D_1$ to $B_1^{E,G}(\gamma)$ without monitoring upon covenant violation ($\phi = \phi^L$). Such a package can be seen as a loan with a performance-pricing covenant. Of course, if a reliable expert triggers an erroneous covenant violation, the face value will automatically go up. Consequently, she will choose low pledgeability, and behave like an unreliable expert. Therefore, subsequent to covenant violation, all experts will repay only $l$. With probability $\mu(1-e)$, there is no covenant violation, and the reliable expert is able to repay $l$. So the expert is able to borrow the expected value, or $l^p_0 = \mu(1-e)\bar{T} + [\mu e + (1-\mu)]l$ upfront. Compared to bank monitoring, a performance-pricing loan is more valuable if the probability $e$ of an error signal is low, the ex-ante probability of being unreliable is high, or the cost of bank monitoring is high.

In the case of performance pricing, the bank is still active. It tracks the signal and increases the face value of the loan after a low signal. Since the noisy signal is costless, the bank does not need to maintain skin in the game. Instead, it will issue a security at $F = B_1^{E,G}(\gamma)$ and sell the entire loan to investors.

C. Upfront Lending and the Effect of Liquidity

In this subsection, we compare the four lending package and determine which one enables the expert to raise the most amount upfront. Let us make the following parametric assumptions, which enable us to study the most comprehensive cases.

Assumption 2:

a. $q > \theta$

b. $\psi < \min\{\pi\left[B_1^{E,B}(\bar{\gamma}) - l\right], (1-\pi)q\left[B_1^{E,G}(\gamma) - B_1^{E,R}(\bar{\gamma})\right]\}$

---

If the signal were costly, the bank would have to maintain skin in the game so as to monitor covenant violations. We have abstracted from this, but the extension is clear. All such loans with covenants would require bank skin in the game, but less for loans that do not require the additional type of monitoring.

Since $F = B_1^{E,G}(\gamma) > D_1^{IC}$, the face value of the security exceeds that of the loan if the covenant is not violated. In other words, the bank over promises to investors. This result is an artifact from the one-loan bank. In case that the bank issues securities against a pool of loans, it will issue a security with face value $\mu(1-e)D_1^{IC} + [\mu e + (1-\mu)]B_1^{E,G}(\gamma)$. This is what the investor will get in expectation from the bank in the one-loan model but will receive with certainty in the case of a pool of loans.
Assumption 2a requires the probability of the good state \( q \) be higher than \( \theta \), the degree of stability. Without this assumption, it will always be the case that
\[
D_1^{IC} \leq B_1^{E,B}(\bar{\gamma}) < B_1^{E,G}(\gamma),
\]
and the equilibrium will not involve internal governance.

Assumption 2b requires the monitoring cost to be sufficiently low; otherwise the equilibrium never involves monitoring. The first condition \( \psi < \pi \left[ B_1^{E,B}(\bar{\gamma}) - L \right] \) guarantees that a loan with monitoring is not always dominated by a performance-pricing loan. Intuitively, it requires the monitoring cost \( \psi \) to be strictly lower than the benefit of correcting a signal error when
\[
D_1^{IC} = B_1^{E,B}(\bar{\gamma}).
\]
The second part of this condition guarantees that a loan with monitoring enables the expert to borrow more than a loan with internal governance when \( D_1^{IC} = B_1^{E,B}(\bar{\gamma}) \).

This is assumed for expository reasons.

Let us compare \( \{l_0^H, l_0^L, l_0^M, l_0^P\} \), the amount that the expert can raise under the four packages. Proposition 3.1 is our main result.

**Proposition 3.1:** There exists a unique triple \( \{\omega_1^*, \omega_1^{**}, \omega_1^{***}\} \) such that

1. If \( \omega_1^{E,G} < \omega_1^* \), \( \max \{l_0^H, l_0^L, l_0^M, l_0^P\} = l_0^H \) so that a loan package with internal governance enables the expert to raise the most upfront.
2. If \( \omega_1^{E,G} \in [\omega_1^*, \omega_1^{**}] \), \( \max \{l_0^H, l_0^L, l_0^M, l_0^P\} = l_0^M \) so that a loan package with a covenant and monitoring enables the expert to raise the most upfront.
3. If \( \omega_1^{E,G} \in [\omega_1^{**}, \omega_1^{***}] \), \( \max \{l_0^H, l_0^L, l_0^M, l_0^P\} = l_0^P \) so that a loan package with a performance-pricing covenant enables the expert to raise the most upfront.
4. If \( \omega_1^{E,G} \geq \omega_1^{***} \), \( \max \{l_0^H, l_0^L, l_0^M, l_0^P\} = l_0^L \) so that a covenant-lite loan enables the expert to raise the most upfront.

For the remainder of this subsection, we offer a proof in a few steps. Along with the proof, we will also illustrate the economic forces behind the results. Recall
\[
L = qB_1^{E,G}(\gamma) + (1-q)B_1^{E,B}(\gamma)
\]
is the amount lent under the covenant lite loan and

---

14 Intuitively, the probability of high liquidity is higher than the degree of expert moral hazard, which ensures \( D_1^{IC} \) is relatively high.
\[ \bar{T} = qD_1^{IC} + (1-q) \min \left\{ D_1^{IC}, B_1^{E,G}(\gamma) \right\} \]

is the amount lent under internal governance. Corollary 2.2 has shown that the difference between \( \bar{T} \) and \( \underline{l} \) decreases with \( \omega_{1}^{E,G} \), the prospective future liquidity.

**Lemma 3.1:** If \( D_1^{IC} \geq B_1^{E,G}(\gamma) \), \( \max \left\{ I_0^H, I_0^L, I_0^M, I_0^P \right\} = I_0^H \) so that internal governance always enables to expert to raise the most upfront.

As discussed earlier, when \( D_1^{IC} \geq B_1^{E,G}(\gamma) \), the lending bank is unable to raise rates following a low signal on pledgeability. Therefore, the bank can lend the most by lending at face value \( D_1 = D_1^{IC} \) and never intervening afterwards. This face value offers incentives for a reliable expert to raise pledgeability, yet it still collects the highest amount from an unreliable expert.

**Lemma 3.2:** If \( \underline{l} \geq \bar{T} \), \( \max \left\{ I_0^H, I_0^L, I_0^M, I_0^P \right\} = I_0^L \) so that a covenant-lite loan enables the expert to raise the most upfront.

Intuitively, when prospective industry liquidity gets very high, a high face value \( D_1 = B_1^{E,G}(\gamma) \) generates more repayment in expectation, even when it incentivizes low pledgeability from the reliable expert, than a lower face value \( D_1 = D_1^{IC} \) that incentivizes high pledgeability. The high-debt covenant-lite package thus dominates other packages that implement high pledgeability with a lower face value, including internal governance, monitoring, and performance pricing.

**Lemma 3.3:** When industry liquidity increases from below such that \( \underline{l} \uparrow \bar{T} \),
\[
\max \left\{ I_0^H, I_0^L, I_0^M, I_0^P \right\} \rightarrow I_0^P
\]
so that a loan with performance pricing covenant raises the most upfront.

Intuitively, the value of high pledgeability gets reduced as \( \omega_{1}^{E,G} \) increases. It eventually shrinks to zero when \( \underline{l} \uparrow \bar{T} \). Therefore, the value from monitoring (gross of monitoring cost) and internal governance also converge to zero as \( \underline{l} \uparrow \bar{T} \). Meanwhile, as long as \( \underline{l} \) still falls strictly below \( \bar{T} \), a performance-pricing loan dominates a covenant-lite loan, because the former package enables the lender to lend more in the case that the covenant is not violated. It also dominates
internal governance so long as the signal is informative. Finally, so long as the cost of monitoring is finite and positive, performance pricing will generate higher lending than monitoring.

**Lemma 3.4:** \( l_0^M - l_0^H \) monotonically increases with \( \omega_1^{E,G} \).

The lemma indicates the value from monitoring compared to internal governance is higher as industry liquidity increases. Intuitively, compared to internal governance, monitoring enables the bank to collect more from an expert if she has set low pledgeability. This allows the banks to collect additional payment in state G, which amounts to \( B_1^{E,G} (\gamma) - D_1^{IC} \). As long as \( D_1^{IC} \leq B_1^{E,B} (\bar{\nu}) \), \( B_1^{E,G} (\gamma) - D_1^{IC} \) increases with \( \omega_1^{E,G} \).

While Lemma 3.4 shows internal governance is dominated by monitoring as \( \omega_1^{E,G} \) increases, Assumption 2b guarantees that at the liquidity level that satisfies \( D_1^{IC} = B_1^{E,B} (\bar{\nu}) \), \( l_0^M > l_0^H \). Thus, internal governance is never optimal if \( D_1^{IC} \) is any lower.

**Lemma 3.5:** \( l_0^M - l_0^P \) decreases with \( \omega_1^{E,G} \).

Intuitively, the value of monitoring compared with performance pricing is to correct the covenant signal error. Such correction becomes less valuable when \( \omega_1^{E,G} \) increases and \( \omega \uparrow \). Therefore, there exists a threshold for \( \omega_1^{E,G} \) such that monitoring raises less than performance pricing once \( \omega_1^{E,G} \) rises above the threshold.

From these lemmas, we can easily see how the thresholds are obtained. When \( \omega_1^{E,G} \uparrow \omega_1^* \), \( D_1^{IC} \downarrow B_1^{E,B} (\bar{\nu}) \). So \( \omega_1^* \) is the value of \( \omega_1^{E,G} \) where \( D_1^{IC} = B_1^{E,B} (\bar{\nu}) \). Above this level of liquidity, costly monitoring is dominant until \( \omega_1^{E,G} = \omega_1^{**} \) which is obtained by solving \( (\bar{T} - \underline{T}) = \frac{\nu}{\pi} \). At yet higher levels of liquidity, performance pricing is dominant, until \( \omega_1^{E,G} = \omega_1^{***} \), which is obtained by setting \( \bar{T} = \underline{T} \). Above this, covenant-lite loans dominate. We write \( \{ \omega_1^*, \omega_1^{**}, \omega_1^{***} \} \) in terms of primitives.

**Corollary 3.1:** in the equilibrium characterized by Proposition 3.1,
\[
\omega_1^* = (1 - \gamma)C_2 - \frac{(1 - q)\theta}{q(1 - \theta)}(\bar{\gamma} - \gamma)C_2,
\]
\[
\omega_1^{**} = (1 - \gamma)C_2 - \frac{(1 - q)\theta}{q(1 - \theta)}\left\{q + (1 - q - \theta)\gamma - (1 - \theta)\bar{\gamma}\right\}C_2 + \frac{\psi}{\pi},\text{ and}
\]
\[
\omega_1^{***} = (1 - \gamma)C_2 - \frac{(1 - q)\theta}{q(1 - \theta)}\left[q + (1 - q - \theta)\gamma - (1 - \theta)\bar{\gamma}\right]C_2.
\]

Figure 5 offers a numerical comparison on how upfront lending in different packages vary with the level of prospective liquidity. Before \(\omega_1^{E,G}\) reaches 0.2314, \(B_1^{E,G}(\gamma) \leq D_1^{IC}\) and therefore the internal governance (blue solid line) enables the expert to borrow the most. Once \(\omega_1^{E,G}\) rises above 0.2314, \(D_1^{IC}\) also falls below \(B_1^{E,G}(\gamma)\), in which case monitoring (yellow dotted line) enables the highest amount of borrowing. Clearly, \(l_0^M\) falls with \(\omega_1^{E,G}\) because \(D_1^{IC}\) falls, while both \(l_0^P\) and \(l_t^C\) increase because \(I\) increases. In this example, \(\omega_1^* = 0.4231, \omega_1^{**} = 0.4881,\) and
\[ \omega^* = 0.5025. \] Other parametric values are: \( q = 0.52, \theta = 0.5, \bar{\gamma} = 0.6, \gamma = 0.3, C_2 = 1, e = 0.45, \mu = 0.8, \psi = 0.01. \]

### D. Discussion

#### D1. Lending and Liquidity

It is well known that current net worth or liquidity matters in reducing borrowers’ agency problems. In our model, we focus on the future net worth of bidders as the exogenous source of liquidity which, following the Shleifer-Vishny (1992) intuition, should also enhance borrowing today. Any exogenous boost to prospective bids – easier monetary policy, a more aggressive lending environment, lax prudential supervision, or even financial development – will play a similar role, so the model is more general than the specific Shleifer-Vishny net worth mechanism. As we show in Diamond, Hu, and Rajan (2020), even though prospective liquidity can lead to greater borrowing today, it also can crowd out borrower incentives to enhance pledgeability or governance. As a result, the firm’s borrowing becomes overly dependent on the persistence of strong liquidity conditions. This is a downside to high prospective asset valuations that is not explored in canonical models of asset-based borrowing (e.g., Kiyotaki and Moore (1997)).

When liquidity is very low, the expert raises the most by setting \( D_1 = D_1^{IC} \), where the reliable expert will raise pledgeability on her own. Since the incentive-compatible level of debt is the highest face value collectible, and it is not possible to collect any more from firms that do not raise pledgeability, the kind of monitoring/performance pricing we have described so far is pointless. There is no role for monitoring. Traditional bank roles in corporate finance are no longer viable.

Monitoring starts to get more valuable as prospective liquidity, \( \omega^E = \omega^G \), goes further up. It allows the bank to set different levels of leverage for borrowers that have chosen different levels of pledgeability, in order to collect more repayment. This also increases the size of the loan that an expert can take out upfront, though the bank will have to be compensated for its cost of monitoring. As liquidity increases further, this cost, coupled with the lower additional benefit to rectifying errors, makes loans with monitoring less valuable than loans with performance pricing. Even though the latter package is subject to error, the consequence of this error gets mitigated as the importance of high pledgeability is reduced by high liquidity.

Eventually when liquidity gets sufficiently high, the amount that can be lent while inducing high pledgeability falls below that from a loan that induces low pledgeability – given that the
former requires debt face values to be capped below a certain level. There is no need for internal governance, or external mechanisms that help implement higher pledgeability such as bank monitoring and performance pricing. Covenant-lite loans, with little reliance on internal governance but instead entirely reliant on liquidity for repayment, predominate at such times. Firm leverage will also be high at those times since liquidity, not pledgeability, is the source of recovery. Interestingly, loan risk may also be higher, but this is because debt is higher. Lending will also be more dependent on continued liquidity.

D2. Intermediation and Intermediary Leverage

At lower levels of liquidity, the lender sets debt at a level where experts voluntarily set pledgeability high. In practice, much of the intermediary’s effort will be up front in doing due diligence on whom to lend to based on observable characteristics. The intermediary can disengage fully ex post, and issues little capital.

At intermediate levels of liquidity, banks play an important role in incentivizing pledgeability while also lending more because they can alter the terms of the loan ex-post based on their monitoring. To make credible their ex post monitoring, they will have to have skin in the game, and hence they will have capital at risk and moderate leverage, as will the firms they lend to.

At higher levels of liquidity, costly monitoring does not contribute much since feasible levels of leverage with or without higher pledgeability are similar. Performance pricing dominates, until liquidity gets so high that pledgeability is irrelevant. Since the intermediary contributes little ex post, it can be a complete pass-through, with high leverage and no equity.

Periods with such high liquidity would thus be associated with either entry by highly levered intermediaries (pass through securitization vehicles) who do not screen or monitor, or a switch by banks to higher leverage and a suspension of monitoring. The securitization vehicles can buy loans originated by banks or other intermediaries and the loans will be covenant lite. Thus covenant lite loans will have a large market share at times when intermediary leverage is high and non-bank lending increases.

Practically, then, our model suggests changes in intermediary capital are a response to the changing demand for monitoring services, and stem from changing liquidity in the corporate sector. Thus, for example, bank leverage was very high before the global financial crisis when liquidity, as measured by the IMF Financial Conditions Index, was high. While regulations forced
bank book capital ratios higher after the crisis, the value of bank capital in Europe, as reflected in market to book ratios for bank stocks, was also low during the recovery. This perhaps reflected low liquidity in the corporate sector, and low need for intermediation services. Other work tends to focus on the state-contingent variation in the supply of intermediary capital, as shocks hit intermediary profits and net worth, which can disrupt the process of intermediation. We hope to take these different predictions to the data in future work.

IV. Extensions, Empirical Relevance, and Related Literature

We now discuss extensions, robustness, some empirical implications, and the related literature.

A. Banks Contributing to Pledgeability

In the baseline model, banks are relatively passive. All they do is to choose whether to monitor and intervene if there is a covenant violation. This may limit the range of situations in which they are seen to be effective. In practice, banks may do more in directly affecting the governance of their borrower (this may be a role played by banks in less financially developed markets or for small young firms in developed ones, and by venture capitalists or private equity in developed markets for young or larger firms). For example, banks may ask for high quality audited financial statements from small borrowers, forcing them to find reputable auditors (Lisowsky, Minnis, and Sutherland (2017)). Venture capitalists often hold seats on the borrowing firm’s board and voting rights far in excess of their cash flow rights, and work to make the enterprise more transparent, governable, and acceptable to the share market (see Hellmann and Puri (2002), Rajan (2012)). In the context of our model, one can think that such an intermediary as not simply checking for covenant violations but also adding to pledgeability through its monitoring.

We extend our model to study this role of financial intermediaries. For simplicity, we will assume all experts are reliable, $\mu = 1$, and there is no uncertainty with respect to the future state at date 1, $q = 1$. Before the incumbent sets pledgeability, the bank could raise the firm’s cash flow pledgeability from $\gamma_2$ to $\gamma_2 + \beta_2$ after paying an additional cost $\delta < \beta C_2$ where $\beta_2 \in \{0, \bar{\beta}\}$. Given $\beta_2 \in \{0, \bar{\beta}\}$, the date-1 bid of experts is

$$B_1^{H,G}(\gamma_2, \beta_2) = \min \left\{ \alpha_1^{E,G} + (\gamma_2 + \beta_2)C_2, C_2 \right\}.$$

15 Key studies of the effects of varying supply of intermediary capital are Holmstrom-Tirole (1997), He-Krishnamurthy (2013) and Rampini-Vishwanthan (2018).
Following the analysis in section II, it is straightforward to show that the incumbent expert will choose high pledgeability \( \gamma_2 = \overline{\gamma} \) if and only if \( D_1 \leq D^{G, PayIC}_1 (\beta_2) \), where

\[
D^{G, PayIC}_1 (\beta_2) = \theta B^{H, G}_1 (\overline{\gamma}, \beta_2) + (1 - \theta) B^{H, G}_1 (\gamma, \beta_2).
\]

Clearly, \( D^{G, PayIC}_1 (\overline{\beta}) \geq D^{G, PayIC}_1 (0) \), with the inequality being strict so long as \( B^{H, G}_1 (\gamma, 0) < C_2 \). So high bank-determined pledgeability \( \beta_2 = \overline{\beta} \) will increase the incentive-compatible level of debt that can be taken by the incumbent (unless liquidity is high enough that pledgeability is never needed).

Next, let us turn to the bank’s incentive in choosing pledgeability. As before, let \( F \) be the exogenous amount of claims on the bank that are due at date 1, which is sufficiently small. The bank chooses \( \beta_2 = \overline{\beta} \) if and only if

\[
\max \left\{ \min \left[ D_1, B^{E, G}_1 (\gamma_2, \overline{\beta}) \right] - F, 0 \right\} - \max \left\{ \min \left[ D_1, B^{E, G}_1 (\gamma_2, 0) \right] - F, 0 \right\} \geq \delta.
\]

It is straightforward to show that the banker’s incentive to increase pledgeability (the left hand side of the constraint above) is maximized at \( D_1 = D^{G, PayIC}_1 (\overline{\beta}) \). This illustrates an interesting double-blessing effect. Intuitively, \( D_1 = D^{G, PayIC}_1 (\overline{\beta}) \) is the highest claim that incentivizes \( \gamma_2 = \overline{\gamma} \) under \( \beta_2 = \overline{\beta} \). However, such a high level of debt will induce the incumbent to choose \( \gamma_2 = \gamma \) under \( \beta_2 = 0 \). This means the incumbent expert will choose high pledgeability if and only if the bank also increases pledgeability. Therefore, in addition to giving the bank the maximum benefit of raising \( \beta_2 \), \( D_1 = D^{G, PayIC}_1 (\overline{\beta}) \) will also induce high firm pledgeability. In this sense, firm internal governance and bank governance can complement each other. Any higher debt level will induce only the bank to raise pledgeability (a “single-blessing”, which lowers the bank’s incentive to do so), while a lower debt level will diminish the bank’s incentive to incur \( \delta \) directly.

How does industry liquidity affect the bank’s choice of \( \beta_2 \)? When industry liquidity \( \omega^{E, G}_1 \) is low or moderate, the bank has no incentive to increase \( \beta_2 \) because \( D_1 \) is too low – the low level of debt either does not require additional pledgeability to enforce or does not justify the bank’s cost of enhancing pledgeability.
At the other extreme, as prospective liquidity $\omega^E_G$ gets sufficiently high, the bank has no incentive again if the asset will be fully priced even without the bank raising $\beta$. Put differently, when $\omega^E_G$ is very high, expert bids are close to full pricing even without the bank enhancing pledgeability. It may not be worthwhile then for the bank to pay the cost $\delta$ to increase it, which in turn may reduce the incumbent’s incentive to increase pledgeability. Once again, high prospective liquidity will crowd out both internal governance but also bank-induced governance. The broad lesson again is that during periods of relatively high industry liquidity, firms don’t rely on bank intermediary services.

The above analysis could also be reinterpreted as the difference in intermediation in a developed country versus a developing country. In countries with better-developed legal systems and more transparent accounting rules, $\bar{\gamma}$ is high and financial intermediaries can do very little to increase pledgeability above $\bar{\gamma}$. Arguably, their role is to monitor, taking governance possibilities as given, as in the earlier sections of the paper. In less developed countries (or for small and young firms), governance can be improved by intermediaries over and above what the institutional environment allows for, and the extension in this section may be more applicable.

**B. Screening Financial Intermediaries**

Our earlier analysis focused on the financial intermediary’s function of monitoring and intervening after covenant violation. In practice, screening is also an important element of the financial intermediary’s service. In the online appendix, we show that our essential results stay unchanged if the financial intermediary can pay a cost to screen experts upfront. Here is a sketch of the model and its important results.

We modify the model presented in section I by dropping the noisy signal $\phi$ and bank’s choice of monitoring. Instead, we assume that when the experts apply for loans, the bank can pay a physical cost $\psi$ to screen borrowers and figure out whether they are reliable or not. Before screening, the bank announces a set of contracts $(l^r_0, D^r_1, F^r)$ and $(l^u_0, D^u_1, F^u)$, from which it chooses after screening. $l_0$, $D_1$, and $F$ refer to the size of the loan, the associated debt payments, and the subsequent security issued to investors respectively. The superscripts stand for the contracts offered to reliable and unreliable experts, respectively.

Let $\omega_0$ be the amount of net worth that experts have at date 0 – termed current liquidity, so that if experts have access to a fixed investment technology which requires initial funding $I$,
they need to borrow \( I - \alpha_b \). Note that in the earlier setup, the borrower sought to maximize the upfront loan amount so as to maximize her upfront bid for the firm. In this setup, she seeks to minimize the cost of borrowing the threshold amount, \( I - \alpha_b \), to start the firm.

If \( l \geq I - \omega_l \) and \( \alpha^{E,G}_l \) gets sufficiently high, banks will lend without screening. In this case, all banks offer the same contract \( l^u_0 = l, \ D^u = B^{E,G}_l (\gamma), \ F^u = B^{E,G}_l (\gamma) \), and the loan is sold entirely to investors – this is the least cost borrowing because no bank capital is required. For the rest of this subsection, we focus on the case of \( I - \alpha_b > l \) so that an unreliable expert is unable to borrow enough to make the investment. As a result, banks will not find it profitable to lend to unreliable experts after screening.

In the parametric regions such that \( D^{IC}_1 < B^{E,G}_l (\gamma) \), two types of equilibria can emerge in the contracts offered to experts upfront, depending on whether these contracts are separating or pooling, i.e., whether \( (l^u_0, D^u_r, F^r) = (l^u_0, D^u_r, F^u) \) or not.

Let us first describe the separation and loan rejection after screening equilibrium. In this case, an unreliable expert’s loan will be rejected after screening. If the bank has found the expert reliable, however, \( D^r \) needs to be set at \( D^{IC}_1 \) to ensure the expert has the incentive to choose high pledgeability. Moreover, we need to check the bank’s incentives. First, the bank must be incentivized to choose the appropriate contract, having screened. In particular, if the expert turns out unreliable, the bank must reject the application instead of offering \( (l^u_0, D^r, F^r) \). This leads to the following self-selection constraint:

\[
0 \geq \left[ -l^u_0 + P\left(F^r\right) \right] R + q\left(D^{IC}_1 - F^r\right) + (1-q)\max\left\{ B^{E,B}_1 (\gamma) - F^r, 0 \right\},
\]

where \( P\left(F^r\right) \) is the price of the security sold to investors, which equals \( F^r \) for \( F^r \leq D^{IC}_1 \) (assuming the incentive constraint are all satisfied). The second incentive constraint requires a bank to be better off screening and lending rather than rejecting all the loans so

\[
-\psi R + \mu \left[ -l^u_0 + P\left(F^r\right) \right] R + \left( D^{IC}_1 - F^r \right) \geq 0.
\]
These two constraints characterize a set of loan contracts that are incentive compatible, as illustrated in the shaded area in the figure below. The preferred contract is the one that allows the highest \( \left( I_0^r, F^r \right) \), which lends the most upfront to a reliable expert and minimizes the call on costly bank capital.

![Figure 6 The Set of IC-Compatible Contracts by Screening Intermediaries](image)

The second type of equilibrium involves pooling, in which banks only offer one type of contract so \( I_0^r = I_0^u = I_0 \), \( D_1^r = D_1^u = D_1 \), and \( F^r = F^u = F \). In other words, the bank may lend the same amount at face value \( D_1 = D_1^{IC} \) and sell securities against part of its balance sheet even though the experts turn out unreliable. In a pooling equilibrium, the bank still faces a few IC constraints. First, it must be indifferent between lending to an unreliable expert or not:

\[
\left[ -I_0 + P(F) \right] R + \left\{ q \left( D_1^{IC} - F \right) + (1-q) \max \left\{ B_1^{E,B} (\gamma) - F, 0 \right\} \right\} = 0.
\]

Even though the contracts offered to both types are identical, the bank still screens because the probability of offering the contract will be different. Specifically, banks will lend for sure to the reliable with positive ex-post profits. Let \( \sigma^u \) be the probability of lending to an unreliable expert and \( \lambda = \frac{\mu}{\mu + \sigma^u (1-\mu)} \in [\mu, 1] \) be the equilibrium expected share of reliable experts who receive

---

16 There are also incentive and participation constraints faced by the experts as we described in the online appendix.
loans. \( P(F) = \left[ \lambda + (1 - \lambda)q \right] F + (1 - \lambda)(1 - q) \min \left\{ B^{E,B}(r), F \right\} \) is the price of security issued. The second IC constraint requires the bank to screen borrowers and then lend based on the screening, rather than not screen and lend:

\[
\left[ -\psi - l_0 + P(F) \right] R + \left[ \mu + (1 - \mu)q \right] \left( D^{IC}_1 - F \right) + (1 - \mu)(1 - q) \max \left\{ B^{E,B}_1(r) - F, 0 \right\} \geq \\
\left[ -l_0 + P(F) \right] R + q \left( D^{IC}_1 - F \right) + (1 - q) \max \left\{ B^{E,B}_1(r) - F, 0 \right\}.
\]

Intuitively, in the pooling equilibrium the unreliable expert effectively gets subsidized by the reliable expert through the high price paid for securities by the uninformed investors. This then serves as a form of ex-ante insurance, for it ensures that some of the unreliables get funding for the project. Pooling can be socially efficient, for it allows experts who might turn out to be unreliable to get funding with some probability for a privately valuable project that would not otherwise be funded because the returns to the lender, absent pooling, are less than breakeven.

The bank here provides the valuable service of certification (for the reliable), but also does not reveal the information it acquires during screening. It thus is able to credibly provide credit to the unreliable also. The bank shares the rents it gets from certifying the reliable with the unreliable (through the pooled price it issues securities at). By offering terms which commit it to pool, we will see the pooling bank could attract experts away from separating but rationing banks because the pooling bank offers the uninformed expert a higher probability of getting funding for the project. From the expert’s perspective, the pooling bank offers partial insurance against her turning out to be unreliable.

This result -- that the bank’s information about the creditworthiness of individual borrowers is not fully used in lending, thus providing insurance to borrowers – is related to the ideas in Dang, Gorton, Holmstrom and Ordoñez (2017). In their paper, banks keep the details about their borrowers secret, which keeps the bank total net worth secret (pooled) over time, allowing the bank to sometimes issue new and overpriced claims. This keeps their old claims safe. The bank is assumed benevolent, and has no conflict of interest which need to be resolved (unlike in our model). In both models, the pooling contract allows banks to fund themselves on different terms than if all of their information was known by outside investors.

The ability to cross-subsidize imperfectly correlated projects within an institution is used in the theories of banks as delegated monitors (Diamond, 1984), where private information is ex-post. It is also seen in models where a diversified firm uses internal capital markets to allocate
funding to ex-ante relatively good projects within the firm (Stein, 1997). The cross subsidy allows internal capital markets to provide more efficient project funding within an institution. Somewhat similarly, pooling in our model works because the claim issued to outside investors is priced to reflect the equilibrium probabilities that loans will be made to a mix of experts, such that the intermediary finds it just profitable to make a loan to unreliable (and strictly profitable to make a loan to reliable) experts. This is a decentralized way of providing greater ex-ante access to funding. The screening model has other interesting features, but the broad implication that prospective liquidity crowds out intermediation is also present in that model.

C. Alternative Interpretations of Liquidity and Agency Problems

We have defined prospective liquidity as the prospective net worth of expert bidders. The reason it is such a powerful support to debt issuance today is that bidders, using their own net worth, can diminish the agency problems associated with external finance, whether they stem from moral hazard or asymmetric information. The diminished need to resolve such problems then reduces the need for internal or intermediary governance, raises the level of external claims that can be issued, which further disincentivizes incumbent management from raising governance.

What is critical is that future expert bids are high for exogenous reasons. We have assumed it comes from high expert net worth (an economy-wide boom in the real sector) but those higher bids could also come from easier monetary policy (as in Bernanke and Gertler (1995)), easier credit conditions associated with easier monetary policy or credit market competition, financial development, or even irrational exuberance in financial markets or lax supervision. Any of these will induce easier financing conditions, higher future bids, and higher current leverage, which in turn induces lower internal and intermediary governance. This broader interpretation is useful in taking the model to the data.

Note that a range of agency problems can be addressed by future bidders bidding with their own wealth. Underlying this is the assumption that there is no problem with the underlying asset (the firm), and its full production possibilities can be obtained in large measure by seizing and selling the asset to the right expert. If, however, the asset itself (and the cash flows that can be produced with it), rather than the human capital that controls it, is of variable and hidden quality, future liquidity will be of little help in correcting malfeasance. Similarly, if the incumbent can make away with the asset in period 1 (which essentially assumes a very low level of period 1 pledgeability), date-1 liquidity will be of little use. The broader point is that prospective liquidity can correct a variety of problems of asymmetric information and moral hazard, not all.
Nevertheless, the focus of our paper is on the effect of variations in prospective liquidity on variation in the need for intermediation services. The implication that in some circumstances intermediation can be dispensed with entirely should be modulated with the appropriate caveats.

D. Empirical Relevance

Studies do find fluctuations in intermediary governance over the cycle. Lisowsky, Minnis, and Sutherland (2017) examine the construction sector over the financing cycle. Specifically, they study the housing boom-bust cycle over the 2002 to 2011 period and find that banks required, on average, fewer high-quality, audited financial statements from construction firms (relative to other firms) before lending to them during the housing boom than prior to 2008. This trend was reversed, however, during the subsequent downturn between 2008 and 2011. Interestingly, banks that collected fewer high-quality financial reports also experienced larger loan losses during the downturn.

Fluctuation in the demand for screening and certification by financial intermediaries has implications for the capital structure of traditional banks, of banks that sell syndicated loans, and of non-bank intermediaries such as securitization structures. The reduced demand for intermediary services (such as monitoring or screening) reduces the need for intermediary capital. This would show up as respectively as increased bank leverage, push banks to retain a smaller fraction of their loans and lead to a reduced fraction of junior claim retained as “skin in the game” by sponsors of loan securitization vehicles. This would imply a positive correlation in changes in these data. For instance, it is not surprising that just before the Global Financial Crisis, retention levels on securitization structures went down even as bank leverage and covenant lite loans increased. Our model would suggest the common factor driving all was expectations of high prospective liquidity.

Bord and Santos (2015) use data from 2004 to 2008 and find loans sold to collateralized loan obligations (CLOs), which are pooled vehicles for securitized loans, underperform unsecuritized loans originated by the same bank. On the other hand, Benmelech et al. (2012) also study CLOs and find little evidence of adverse selection before 2005 – securitized loans performed no differently from loans held on bank balance sheets. However, the evidence is more mixed in the 2005-2007 sample. Much like Begley and Purnanandam (2016) on mortgage-backed securities, they suggest that appropriate structuring (with greater skin in the game) helped give originators the right incentives; CLOs primarily held syndicated loans, where lead bank loan originators retained substantial skin in the game by holding on to a fraction of the originated loans on their balance sheets.
While indeed greater skin in the game may have been associated with better intermediary screening, another important environmental factor was also at work: changes in the underlying liquidity for the assets being securitized. This may explain some of the differences in the empirical evidence described above. Arguably, liquidity was moderate but increasing as the economy recovered from the Dot Com bust. Securitizers did substantial due diligence, and securitization structures reflected their desire to signal their commitment, as suggested by Begley and Purnanandam (2016). As the recovery picked up and policy interest rates stayed lower than normal (see Taylor (2010)), liquidity increased, and the need for intermediary services diminished, until very little screening or monitoring was done just before the crisis, as suggested by Benmelech et al. (2012) and Keys et al. (2010). Seen with the benefit of hindsight from the depth of the crisis, this may have seemed to be an aberration, and some indeed was. Yet it was also consistent with the kind of behavior induced by expectations of high liquidity, reducing the demand for intermediary services, even while increasing the amount of borrowing and its riskiness. It is also possible that the expectations were too extreme, with the probabilities of the low liquidity state underestimated as in Gennaioli, Shleifer and Vishny (2015), yet that does not take away from the fundamental thrust of our arguments.

From a policy perspective, demanding that intermediaries hold more skin in the game during the period of high liquidity may have been futile. It would simply have accentuated the move toward lending structures that minimized intermediary involvement.

V. Conclusion

While this paper describes how financial intermediation varies with prospective liquidity in the underlying real borrowing sector, there is a more general point here. Liquidity tends to diminish the consequences of many kinds of moral hazard over repayment. Internal governance matters little if the firm can be seized and sold for full repayment in a chapter 11 bankruptcy. Similarly, liquidity can also diminish the consequences of adverse selection over borrower types. Once again, it matters less if the manager is reliable or unreliable if the firm she manages can be seized and sold for full value. Therefore, prospective liquidity encourages leverage at both the borrower and intermediary level, even while requiring less governance. Equivalently, because the intermediary performs fewer useful functions, high prospective liquidity encourages disintermediation.

Evidence that intermediaries abandon their natural functions of monitoring (or screening) when markets are very liquid does not mean their functions are without value at other times. Similarly, it may not be appropriate to look back after liquidity collapses to claim securitization is
problematic. Both borrowing and securitization may have been optimized for the high liquidity states ex ante, and that may have been the best thing for the borrower and securitizer to do. Effectively, both may have neglected the low liquidity state, but that is a consequence of the liquidity leverage nexus.

We have not explored the adverse consequences of the reliance on liquidity. Essentially, it induces a kind of liquidity addiction, since other supports to borrowing are crowded out. We have also not explored how intermediaries might behave as the demand for their services wax and wane. Can they shrink easily after having expanded, or do they compete for mandates even when they have little comparative advantage relative to other forms of finance? Does this contribute to financial fragility? Finally, at a macroeconomic level, liquidity is endogenous and deserves to be explored further. These are important areas for future research.

References


Proof of Proposition 2.1

The first statement naturally follows by taking the difference between $V_{1}^{I,S_{1}}(D_{1}, F)$ and $V_{1}^{I,S_{2}}(D_{1}, F)$. The second statement follows by solving the equation

$$q\Delta_{1}^{g}(D_{1}^{x}) + (1-q)\Delta_{1}^{s}(D_{1}^{x}) = 0.$$  

The solution is unique unless
\[ q(1 - \theta)\left[ B_1^{E,G}(\overline{\gamma}) - B_1^{E,G}(\underline{\gamma}) \right] = (1 - q)\theta\left[ B_1^{E,B}(\overline{\gamma}) - B_1^{E,B}(\underline{\gamma}) \right], \]

in which case we pick the highest solution.

**Proof of Corollary 2.1 and 2.2**

Corollary 2.1 naturally follows by solving \( q\Delta^G_i\left(D_i^{IC}\right) + (1 - q)\Delta^B_i\left(D_i^{IC}\right) = 0 \). Note that

\[ \Delta^G_i\left(D_i^{IC,PostIC}\right) = \Delta^B_i\left(D_i^{IC,PostIC}\right) = 0. \]

\[ \overline{T} - \underline{L} = \left[ qD_i^{IC} + (1 - q)\min\{D_i^{IC},B_i^{E,B}(\overline{\gamma})\} \right] - \left[ qB_i^{E,G}(\overline{\gamma}) + (1 - q)B_i^{E,B}(\overline{\gamma}) \right]. \]

If we collect terms involving bids in state G, the relevant part is \( q(1 - \theta)\left(B_i^{E,G}(\overline{\gamma}) - B_i^{E,G}(\underline{\gamma})\right) \) if

\[ D_i^{IC} > B_i^{E,G}(\overline{\gamma}) \]

and

\[ \frac{q(1 - \theta)}{1 - q}\left[ B_1^{E,G}(\overline{\gamma}) - B_1^{E,G}(\underline{\gamma}) \right] - qB_1^{E,G}(\overline{\gamma}) \]

if \( D_i^{IC} \leq B_i^{E,B}(\overline{\gamma}) \). Clearly, both fall with \( \omega_i^{E,G} \).

Finally, \( D_i^{IC} \) falls from some level above \( B_i^{E,G}(\overline{\gamma}) \) to some level below \( B_i^{E,B}(\overline{\gamma}) \) as \( \omega_i^{E,G} \) increases, validating global monotonicity.

**Proof of Proposition 3.1 and Corollary 3.1**

Naturally follows from Lemma 3.1-3.5, where \( \omega_i^{*} \) is the value of \( \omega_i^{E,G} \) where \( D_i^{IC} \) falls to

\[ B_i^{E,B}(\overline{\gamma}). \]

\( \omega_i^{**} \) is obtained by solving \( \overline{T} - \underline{L} = \frac{\psi}{\pi} \). \( \omega_i^{***} \) is obtained by setting \( \overline{T} = \underline{L} \).

**Proof of Lemma 3.1-3.5**

If \( D_i^{IC} \geq B_i^{E,G}(\overline{\gamma}) \), the initial borrowing amount from different packages is as follows:

- \( l_0^M = \mu\overline{T} + (1 - \mu)\underline{L} - \left[ \mu e + (1 - \mu) \right] \psi \).
- \( l_0^L = \underline{L} = qB_i^{E,G}(\overline{\gamma}) + (1 - q)B_i^{E,B}(\overline{\gamma}). \)
- \( l_0^p = \mu(1 - e)\overline{T} + \left[ \mu e + (1 - \mu) \right] \underline{L} \).
- \( l_0^H = \mu\overline{T} + (1 - \mu)\underline{L} \).

Lemma 3.1 clearly follows. If \( D_i^{IC} \leq B_i^{E,B}(\overline{\gamma}) < B_i^{E,G}(\overline{\gamma}) \), the initial borrowing amount from different packages becomes:
Lemma 3.2-3.5 are thus straightforward from these expressions.

- \( l_0^M = \mu \bar{T} + (1 - \mu) \underline{L} - [\mu e + (1 - \mu)] \psi \).
- \( l_0^L = \underline{L} = q B_i^{E,G} (\gamma) + (1 - q) B_i^{E,B} (\gamma) \).
- \( l_0^P = \mu (1 - e) \bar{T} + [\mu e + (1 - \mu)] L \).
- \( l_0^H = \bar{T} - (1 - \mu)(1 - q) \left[ D_i^{E,C} - B_i^{E,B} (\gamma) \right] \).