Correlated Default and Financial Intermediation

Gregory Phelan*

This version: September 13, 2016

Abstract

Financial intermediation naturally arises when knowing how loan payoffs are correlated is valuable for managing investments but lenders cannot easily observe that relationship. I show this result using a costly enforcement model in which lenders need ex-post incentives to enforce payments from defaulted loans and borrowers' payoffs are correlated. When projects have correlated outcomes, learning the state of one project (via enforcement) provides information about the states of other projects. A large, correlated portfolio provides ex-post incentives for enforcement; thus, intermediation dominates direct lending, and intermediaries are financed with risk-free deposits, earn positive profits, and hold systemic default risk.

Keywords: Financial intermediation, systemic risk, default

*Williams College, Department of Economics, Schapiro Hall, 24 Hopkins Hall Drive, Williamstown MA 01267. Email: gp4@williams.edu. Website: `https://sites.google.com/site/gregoryphelan/`. I have benefited from feedback from Quamrul Ashraf, Bruno Biais, William Brainard, Gerard Caprio, Maximilian Eber, John Geanakoplos, Johannes Hörner, Michael Kelly, David Love, Peter Montiel, Guillermo Ordoñez, Peter Pedroni, Stanislav Rabinovich, Ashok Rai, Alexis Akira Toda, and anonymous referees. The views and errors are my own.

As is well known, banks and other financial intermediaries frequently hold assets whose risks are correlated and hard to value.In other words, loans are jointly correlated but it is either unclear how they correlate with easily observable information, or that information has low explanatory power. Banks do not simply pool risk: banks' balance sheets are risky and banks produce information to efficiently monitor investments, yet banks can nonetheless produce safe deposits from their assets. This paper proposes a mechanism that would explain intermediation with this risk exposure.[1]

The key contribution of this paper is to show that financial intermediation with correlated risks naturally arises when knowing how loan payoffs are correlated is valuable for managing investments but lenders cannot easily observe this relationship (as might be the case for investments in non-traded securities). When a complete description of a portfolio's correlation structure is unknown, an investor with a large portfolio can learn the relevant information by monitoring or servicing the assets in the portfolio—and the investor has incentives to do so because this information is valuable. Small investors necessarily cannot learn much from their portfolios (the sample size is small). As a result, it is natural for an intermediary to borrow from small investors and to hold a large portfolio, which provides incentives to learn the condition of their loans and to improve the portfolio performance. In addition to motivating why intermediaries hold loans with correlated risks, my model rationalizes empirical regularities regarding returns to scale and the use of hard and soft information in lending (discussed below).

I illustrate my results in a simplified costly enforcement model as in Krasa and Villamil (2000). Lenders must pay a cost to enforce payments from borrowers, and lenders cannot commit to enforce payments (or to monitor, service, or audit defaulted loans) and so must have ex-post incentives to do so. I depart from their setup in two principal ways. First, I suppose that borrowers' payoffs are correlated: there are two "aggregate states" that index the distribution of payoffs. Knowing the aggregate state (how loan payoffs are related) typically allows lenders to pursue improved enforcement strategies. Second, I suppose large enforcement costs: if a borrower defaults

[1] Economists have developed a number of other explanations for why intermediaries are often exposed to more correlated, or aggregate, risk than standard models of intermediation would predict: fixed costs of developing expertise in an industry or asset class provide incentives to specialize; limited liability gives incentives to take risk, passing downside risk to depositors (Hellwig, 1998); not monitoring as a form of ex-ante risk-shifting encourages investment in aggregate risk (Kahn and Winton, 2004); banks have strategic incentives to fail at the same time when they expect to be bailed out (Farhi and Tirole, 2012).

only in the states in which the borrower cannot repay the loan, the enforcement cost exceeds the expected gain. As a result, the borrower defaults stochastically when able to repay. Knowing the joint condition of loans is valuable when enforcement costs are low, but even more valuable when costs are high. When costs are high, without knowing the joint condition of loans a lender does not have ex-post incentives to enforce payments. But a lender will have incentives to enforce payments ex-post when the likelihood of a "good aggregate state" is sufficiently high (Proposition 3).

In this setup, correlation increases portfolio values because lenders must take costly actions to manage their investments (Proposition 4), and financial intermediation endogenously arises when borrowers have correlated payoffs. I show that intermediation can arise when intermediaries diversify idiosyncratic risks in order to isolate correlated (sectoral/aggregate) risks (Proposition 6). By diversifying away all risks *except* correlated risks, the intermediary can commit to monitor more frequently than a single investor can, and as a result borrowers repay more frequently. Intermediation serves to decrease monitoring costs in the economy because correlation minimizes expected monitoring costs when investors cannot commit to monitor. As a result, intermediaries earn positive expected profits and hold correlated risk.[2]

The model simplifies many important features to illustrate the point, and it is worth understanding some of the richer features of the real world we have in mind. First, information about current conditions is useful for managing investments. In general, any loan that requires interim or ex-post monitoring will benefit from having information regarding the appropriate action. In particular, knowing the likely condition of defaulted borrowers provides useful information for monitoring or auditing those loans. Lenders have many options available when dealing with defaulted borrowers—e.g., restructuring the loan, delaying foreclosure, Chapter 11 vs 7—and the best choice may depend nontrivially on current conditions. Servicing a loan to restructure payments—e.g., de-

[2]This result is closely related to the literature on diversification and intermediation. Diamond (1984) and Williamson (1986) show that when multiple lenders are needed to fund a single project and there is costly state verification, financial intermediation decreases monitoring costs. The intermediary diversifies idiosyncratic risks, which allows an intermediary to offer risk-free deposits to its investors, thus eliminating the need to "monitor the monitor." In Diamond (1984) and Williamson (1986), intermediaries earn sure portfolio returns and zero profits. Compared to Hellwig (1998), in my model the demand for aggregate risk is not driven by limited liability; it is instead a feature that enables intermediaries to offer risk-free deposits. In contrast to Kahn and Winton (2004), the decision to monitor in my paper is ex-post rather than ex-ante. My paper also relates to Boyd and Prescott (1985), in which intermediaries naturally arise as coalitions to address an information problem.

ferring interest payments or decreasing the principal value of a loan—may improve the value to the lender, but servicing is a costly process that could prove fruitless. The liquidation value of a borrower's assets depends on the market for those assets; the market value of a firm's assets depends on the condition of the industry; the market value of a house depends on the local housing market.[3] In a corporate default, it may be obvious that proceedings must wipe out equity holders, but should a restructuring also take the costly decision to replace management? If the default was caused by aggregate conditions rather than bad management, then no.

Second, there are valuable components of the correlation structure of some assets that are sufficiently difficult to learn. Even when an investor can perfectly observe aggregate conditions, the way investment payoffs depend on that state may still be uncertain—in other words, investors may not know the sensitivities of their investments to what is easily measured. (In contrast, the standard assumption is that "knowing the aggregate state" means knowing the aggregate state *and* all the implications of that state.) For instance, an investor may know the economy is in a boom or a recession without knowing what that means for the loans in her portfolio. Does she have loans that will weather the storm, or will they turn south? There is no doubt *that* an aggregate shock occurred 2007–2009, but economists have spent years debating the implications.[4] The issue is further complicated when investors must also discern the appropriate action for managing an investment, particularly if the appropriate action of the past is not obviously appropriate for current conditions.

Additionally, signals of current conditions may not provide all the relevant information about

[3]Bernstein et al. (2015) highlights the importance of local markets and asset specificity in resolving financial distress: whether the indirect costs of Chapter 7 bankruptcy exceed those of Chapter 11 are concentrated in thin local asset markets with few potential buyers; in contrast, no differences occur in thick asset markets. Cantor and Varma (2004) show that recovery rates depend on contemporaneous industry effects, such as capacity utilization, and that macro factors are more important for Chapter 11. Importantly, Gupton et al. (2000) finds no ex-ante difference in recovery rates by industry, providing evidence that contemporary "aggregate" information is useful. Woo (2009) shows that the whether residential developers ought to file Chapter 7 or Chapter 11 depends on aggregate conditions. Acharya et al. (2007) finds that creditors of defaulted firms recover significantly lower amounts in present-value terms when the industry of defaulted firms is in distress.

[4]For an interesting example, Lubben (2012) argues that observing an aggregate shock, even one as clear as Lehman, nonetheless is insufficient to know how to proceed for liquidation or reorganization for related complex firms. Lubben examines a legal and financial structure of Bank of America and argues that no matter how complex Lehman was, the remaining "too big to fail" financial institutions were significantly more complex, and thus the aggregate implications of Lehman's failure would remain quite unclear.

aggregate conditions. The economy is a complex and multi-dimensional object and it is rarely the case that a single index or even a combination of indices—such as the unemployment rate, growth rate of GDP, house prices, or inflation—would span all the necessary information for understanding the behavior of hard-to-value investments.[5] The pertinent information might be how the condition of an industry or asset class loads on other aggregate information. Signals of aggregate conditions are at best noisy and may be biased, as credit ratings in the mid-2000s appear to have been (given the high sensitivity of CDO's to the correlation of subprime loans, even small noise observing the true correlation would lead to vastly different valuations). In other words, a noisy signal of aggregate conditions does not reveal the correlation matrix for the multi-dimensional state.

Empirical Evidence

Banks' assets are hard-to-value but correlated. DeYoung et al. (2001) find that government examinations produce new, value-relevant, information which is eventually revealed in bank subordinated debt prices. Berger and Davies (1998) find that information from unfavorable examinations is eventually revealed in banks' stock prices. Haggard and Howe (2007) find that banks have less firm-specific information in their equity returns than matching industrial firms. Morgan (2002) and Iannotta (2006) find that bond rating agencies are more likely to disagree on the ratings of banks compared to other firms, suggesting that banks are harder to understand.

Turning to the empirical literature on enforcement costs and recovery—as noted by Acharya et al. (2007), direct (administrative and legal) costs of formal bankruptcy are rather small, and so economists have shifted attention to indirect costs arising from the loss of intangibles and growth opportunities, bargaining inefficiencies, and fire-sale liquidations during industry-wide distress.[6] First, indirect costs depend on the enforcement/bankruptcy option. Bris et al. (2006) find that Chapter 7 is neither cheaper nor faster than Chapter 11 (after accounting for selection effects), but that

[5] Bruche and Gonzalez-Aguado (2010) provide statistical evidence that "credit downturns" do not perfectly align with NBER recessions, often starting before and ending after, or not occurring during a recession at all. (Neither does NBER know in real-time when a recession has started.) Li and White (2009) show that bankruptcy and mortgage default are substitutes in some cases and complements in others, and local neighborhood conditions affect borrowers' decisions.

[6] Djankov et al. (2006) find that globally enforcement costs are very time consuming, costly, and inefficient, with average losses of 48%. Enforcement costs are very heterogeneous, with richer countries more efficient, and in poor countries reorganizations typically fail so the best procedure is foreclosure.

Chapter 11 preserves assets better.Gupton et al. (2000) find that bankruptcy length is 60% longer for Chapter 11 than for prepackaged Chapter 11. Second, recovery rates and enforcement costs differ depending on the number of creditors and whether borrowers had substantial bank loans. Gilson et al. (1990) find that firms that owe more of their debt to banks, or that owe fewer lenders, are most likely to restructure their debt privately. Gupton et al. (2000) find that the senior unsecured recovery rates is 63.4% for single-loan defaulters, but 36.8% for multiple-loan defaulters. Chatterjee et al. (1996) find that firms with lots of bank debt more likely to use Chapter 11.

Empirically, there is mixed-evidence that banks experience returns to scale in credit activities (economies of scale are more important for market-based activities): early empirical studies find that scale is limited to relatively small banks with less than \$10-50 billion in asset (Laeven et al., 2014); however, more recently, Hughes and Mester (2013) find returns to scale in large banks by explicitly considering how banks choose risk, potentially taking more risk as size increases, and Anderson and Joeveer (2012) find economies of scale by considering that rents are captured by bankers, and hence not picked up by analyses that assume competitive factor markets.

In my model, returns to scale diminish quickly because learning about correlation risks occurs quickly, and large intermediaries can experience returns to scale precisely because they can benefit from correlated risk exposure while small investors cannot. While in my baseline model financial intermediaries experience returns to scale and so the optimal size is infinite, the returns to scale are most prominent for small size and diminish quickly.[7] Additionally, in my model large banks experience returns only because their size allows them to take on investment risk (correlated risk exposure for the purpose of learning) that is inaccessible to small banks, which is consistent with the hypothesis in Hughes and Mester (2013) that banks achieve scale by taking on new types of risk as their size grows.

Empirically, large banks make loans based on quantifiable, "hard" information, while small

[7] Intermediation also arises because a large intermediary can diversify idiosyncratic risks; however, Krasa and Villamil (1992a) show that this result kicks in very quickly and that as few as 30 projects are enough to sufficiently diversify. Furthermore, when enforcement costs are increasing and optimal intermediary size is finite, I show that correlation increases portfolio value and the maximum value-per-loan is non-monotonic in correlation. Future work might incorporate aggregate risk that can take on multiple values, as in Krasa and Villamil (1992b), who show that finite-sized intermediaries are optimal because aggregate risk leads to non-zero monitoring costs. I conjecture that adding my learning mechanism to their story would deliver finite-sized, but larger, intermediaries.

banks tend to make loans based on qualitative, idiosyncratic, "soft" information (Strahan, 2008). In my model large banks capitalize on having better knowledge about how loans are correlated (hard information) not by having better idiosyncratic knowledge (soft information), and access to hard information about the correlation of loans enables intermediaries to earn rents.[8] In light of these issues my model presents a mechanism for how large banks, lending based on hard information, acquire information. In my model large banks acquire information from managing assets, not from observing the behavior of their liabilities (deposits), providing a rationale for how large banks with large borrowers can nonetheless acquire an informational advantage.[9]

Related Literature

My paper relates to the literature on delegated monitoring, which includes Diamond (1984) and Williamson (1986). In these canonical models, intermediation arises because scale and complete diversification allow banks to reduce monitoring costs; these models do not predict bank portfolios of the type we see. Krasa and Villamil (1992a) consider delegated monitoring with finite sized intermediaries and show that (i) two-sided debt contracts with delegated monitoring dominate direct investment, (ii) two-sided debt is optimal, and (iii) only about 30 investments are needed for diversification and delegation to work well. Krasa and Villamil (1992b) consider a costly state verification model with minimum project size and non-diversifiable aggregate risk, and show that finite-sized intermediaries are optimal because aggregate risk leads to non-zero monitoring costs. These results are driven by declining costs rather than increasing monitoring. The new and important result of my paper is that correlation allows the intermediary to earn rents while providing the

[8] Small banks tend to have small borrowers who are geographically near, providing better access to soft information, (Berger et al., 2005). Furthermore, large banks enjoy scale economies that permit them to succeed with price competition, although only when they provide a standardized product based primarily on verifiable, "hard" (quantifiable) information (DeYoung, 2008; Cole et al., 2004); nonetheless, large banks earn profits.

[9] Gorton and Winton (2003) argue that a potentially important aspect of information production by banks concerns whether the information is produced upon first contact with the borrower or is instead learned through repeated interaction with the borrower over time. One way banks acquire information about lenders is by monitoring check account activity for their borrowers, but this applies only to small banks. Nakamura (1993) argues that small banks lending to small businesses are especially well suited to use checking account information; in contrast, the payments activities of large firms are both too complex and too dispersed to be of much value to a potential bank lender (findings by Cole et al. (2004) supported this claim). Botsch and Vanasco (2015) provide evidence that banks acquire valuable private information about borrowers via lending relationships, and private bank learning about firm quality particularly benefits higher-quality borrowers, who receive lower interest rates on subsequent loans.

return promised to investors; intermediaries may have an incentive to hold some risk, and this can be efficient.

My paper relates to the literature on costly state verification and costly enforcement. Townsend (1979) and Gale and Hellwig (1985) show that when lenders can commit to monitor deterministically, the optimal contract is a standard debt contract; however, with randomized monitoring, the optimal contracts are no longer standard debt. Mookherjee and Png (1989) consider a model with randomized auditing and a moral-hazard problem so that contracts must provide incentives for auditing and incentives for borrowers to take desirable actions. In a modified setting with costly enforcement instead of costly verification, Krasa and Villamil (2000) show that simple debt contracts are optimal when there is limited commitment to initial decisions and enforcement is costly and imperfect (with full commitment, stochastic contracts are optimal).

In my model, investing in correlated loans functions as a commitment device. My result relates to Khalil and Parigi (1998), who show that loan size acts as a commitment device when auditing costs are fixed, because a recovering from larger loan provides ex-post incentives to pay the fixed cost. I show that even weak correlation can create commitment for large portfolios (with perfect correlation many loans behave just like one very large loan). My result is not a direct extension of Khalil and Parigi (1998)—in my model, while auditing costs increase with the number of loans (as more loans need to be audited), expected auditing costs per loan decrease with a large, correlated portfolio, enabling commitment. Additionally, while Khalil and Parigi (1998)'s results have implications for optimal loan size, my model makes a statement about optimal loan composition. In particular, Khalil and Parigi (1998) implies that lenders make loans that are larger than what they would have made in a world with commitment (but smaller than in a frictionless world); my theory implies that due to the lack of commitment, lenders will choose to invest in a portfolio with large degrees of correlated risk.[10]

My paper also relates to the literature on the production of risk-free deposits from risky as-

[10] Melumad and Mookherjee (1989) show that delegation can serve as a commitment device, and in my paper lenders delegate monitoring responsibilities to a large intermediary. My result is also related to Ben-Porath and Dekel (1992) in which costly actions can signal future actions. The idea that lenders may not always find it optimal to enforce repayment is related to Zhao (2008) and Chen (2012) who provide environments in which a principal may not want to reward every action an agent takes even when those actions are observable.

sets. Gorton and Pennacchi (1990) show that risk-free, information-insensitive debt is useful for mitigating informational asymmetries, but acknowledge that in principal firms, not just banks, can produce risk-free debt by financing themselves with debt and equity. Still, intermediaries seem to have a "monopoly" on producing money-like securities. Dang et al. (2014) argue that banks can produce information-insensitive deposits because they are opaque: they invest in hard-to-value assets and keep information secret. In my model, because intermediaries have access to information that is useful for managing loans, borrowers are less likely to default when borrowing from intermediaries. This equilibrium result implies that banks can offer better returns even in states of the world when many investments perform poorly.[11]

1 The Basic Model

I consider a simplified "reduced-form" version of the economy in Krasa and Villamil (2000) in which lenders cannot commit to an enforcement strategy. The economy consists of two risk-neutral agents and three periods, $t = 1,2,3$. Agents derive utility from consumption in the last period and there is no discounting. The lender, or investor, has consumption/investment goods to lend in period 1. The borrower, or entrepreneur, has no endowment in goods but has access to a production technology that transforms 1 unit in period 1 into a stochastic value s from $S = \{s_1,...,s_n,...,s_N\}$ in period 2. Agents share a common prior belief $f(s)$ of probabilities. I interpret s_1 to be the "liquidation value," or collateral value, of the investment, and thus $s_n - s_1$ can be interpreted as non-collateralizable, or non-seizable, cash flows in state s_n.

Because the entrepreneur has a technology but no input goods, and the investor has input goods but no technology, the entrepreneur will borrow 1 unit from the investor in exchange for payments in later periods specified by a contract. Nature determines the the project payoff s, which is known

[11] My paper also relates to the "fragility is good" literature: Diamond and Rajan (2000) argue that financial fragility is a disciplining feature; Farhi and Tirole (2012) show that systemic bailouts make it profitable for banks to adopt risky balance sheets; Acharya (2009) provides a model in which banks undertake correlated returns as a result of a systemic risk-shifting externality. In my paper, financial intermediation arises precisely because intermediaries hold portfolios with correlated, systemic default risks, which is ex-ante desirable. Van Nieuwerburgh and Veldkamp (2010) and Gârleanu et al. (2013) emphasize that investors may want to hold correlated portfolios because information acquisition is costly and so optimal portfolios are not fully diversified.

to the borrower in period 2. In period 2 the borrower then makes a voluntary payment $r \leq s$ to the lender, which cannot exceed the total project payoff s. In period 3 the lender, having observed the payment r but not the realized state s, decides whether to enforce a final payment—the lender cannot commit in period 1 to an enforcement strategy in period 3. Enforcement is provided by an outside agent such as a court and costs the lender γ.[12]

Contracts take the form of standard debt contracts with exogenous face value F, which can be interpreted as the interest rate and satisfies $s_1 < F < s_N$. The exogenous interest rate is to ease the exposition and one could endogenize the interest rate without changing results (suppose the borrower has a reservation utility and the lender offers the contract). If the borrower pays $r < F$, then the lender can enforce payment and claim the full project value s_n. The enforcement action is denoted by $e \in \{0,1\}$, where $e = 0$ is no enforcement and $e = 1$ is enforcement. Final consumption values for the borrower are given by

$$x_B(s) = \begin{cases} 0 & \text{if } r(s) < F \text{ and } e = 1 \\ s - r(s) & \text{otherwise} \end{cases} \tag{1}$$

and for the lender

$$x_L(s) = \begin{cases} s - \gamma & \text{if } r(s) < F \text{ and } e = 1 \\ r(s) - \gamma & \text{if } r(s) \geq F \text{ and } e = 1 \\ r(s) & \text{if } e = 0 \end{cases} \tag{2}$$

A contract defines a game summarized by a set of players, strategies, a production technology, beliefs, and payoffs. The set of (mixed) strategies are Σ_B, Σ_L. Strategy $\sigma_B \in \Sigma_B$ is the conditional distribution of payments $r(s) \in [s_1, s_N]$ given the state s. Formally, $\sigma_B(r;s) = \Pr(r|s)$. Strategy $\sigma_L \in \Sigma_L$ is the conditional probability of enforcement action $e \in \{0,1\}$ given the payment of the borrower, i.e. $\sigma_L(r) = \Pr(e = 1|r)$. After seeing the payment r, the lender has posterior beliefs $\tilde{f}(s|r)$ about the realization s. The strategies are used to choose $r(s)$ and $e(r)$ optimally as part of

[12] The setup simplifies a number of features in Krasa and Villamil (2000). In their setup, payoffs take a finite number of positive values, both agents face dead-weight loss costs of enforcement, borrowers can protect a small fraction from being seized in the case of enforcement, and borrowers have a reservation utility. I leave these issues out to simplify the analysis.

a Perfect Bayesian Nash Equilibrium (PBE).

Definition 1 (PBE)**.** *A collection of strategies σ_B, σ_L and beliefs $f, \tilde{f}$ are a Perfect Bayesian Nash Equilibrium if and only if*

1. *$\sigma_B \in \Sigma_B$ maximizes $\mathbb{E}_{\sigma_B,\sigma_L}[x_B(s)]$ for every s.*
2. *$\sigma_L \in \Sigma_L$ maximizes $\mathbb{E}_{\tilde{f},\sigma_L}[x_L(s)]$ for every r.*
3. *$\tilde{f}(s|r)$, the posterior belief that the state is s given payment r, is derived using Bayes' rule whenever possible, where $f(s)$ is the prior belief.*

In period 1 agents choose strategies and the lender gives 1 unit to the borrower. In period 2 the project realizes and the borrower makes a payment to the lender. In period 3, the lender, according to σ_L chooses whether to enforce payment.

Broadly, one can think of servicing a loan as all of the financial and legal considerations that would go into reducing debt payments and restructuring a loan rather than, when possible, just expediently taking a borrower into court and seizing assets. In other words, a lender can costlessly seize s_1, but getting a higher payment may require taking a costly action. Delaying foreclosure can be costly—the homeowner may stop taking care of his house—and, of course, delays repayment; however, doing so may, in a good state of the world, provide the homeowner the ability to repay some of his remaining debt. The option to not service a loan could mean selling collateral quickly in a liquidation and getting a cheap value, rather than trying harder to get more either out of the collateral or out of the borrower. Hence, the enforcement decision could capture deciding between filing Chapter 7 and Chapter 11 (or 13). The empirical literature suggests that these options can have very different indirect costs, that the optimality of each of these may depend on current conditions, and that different lenders pursue these options with different frequencies.

Parametrization. To simplify exposition, without loss of generality, we can consider the payoffs as taking a stochastic value s from $S = \{0, 1, G\}$, with probabilities as given in Table 1, where G is understood to be large. Let the required payment R ("the interest rate") satisfy $1 < R < G$. As a result, the borrower must default in states $s_1 = 0$ and $s_2 = 1$, which I refer to as "default states." It is convenient to define the following probabilities:

$$\pi = f(s_1) + f(s_2), \quad \kappa = \frac{f(s_2)}{f(s_1)+f(s_2)}. \tag{3}$$

That is, π is the probability that the project realizes in one of the two default states, and κ is the conditional probability that the default state is the good default state s_2.[13]

Table 1: Probabilities of Project Payoffs

State	0	1	G
$f(s)$	$(1-\kappa)\pi$	$\kappa\pi$	$1-\pi$

where $\kappa \in [0,1], \pi \in (0,1)$

I assume throughout that (i) $\gamma < 1$ so that enforcement is possibly optimal for $s = 1$, (ii) $\gamma < \pi\kappa + (1-\pi)G$ so that enforcement is always optimal if the borrower never repays, and (iii) $\gamma < G - R$ so that enforcement is optimal if $\tilde{f}(G) = 1$ and $r(G) < R$.

Equilibrium With Commitment

As is well-known, if the lender can commit to an enforcement strategy then the borrower will truthfully reveal when he can repay R and agents use pure strategies (see Townsend (1979)). The equilibrium strategies with commitment are: $r(0) = 0$, $r(1) = 1$, $r(G) = R$, and $e(r) = 1$ for $r < R$ and $e(r) = 0$ for $r \geq R$. Expected payoffs to the borrower and lender are

$$V_B^C = (1-\pi)(G-R), \tag{4}$$

$$V_L^C = (1-\pi)R + \pi(\kappa - \gamma), \tag{5}$$

and total utility is given by

$$V_S^C = G\left(1 - \pi + \pi\frac{\kappa}{G}\right) - \pi\gamma, \tag{6}$$

[13] To see why the normalization is without loss of generality, consider a 3-state model, $S = \{s_1, s_2, s_3\}$: s_1 is the scrap value of the project; normalize $s_2 - s_1 = 1$; define $G = s_3 - s_1$. Set $F = s_1 + R$ so that R is the payment required in excess of the liquidation value. Thus the model is about payoffs in excess of the minimum value. In Appendix B I consider a model with a continuum of states.

which differs from the expected project payoff by $\pi\gamma$, which is the deadweight loss associated with expected enforcement costs.

Equilibrium Without Commitment

The heart of the paper is what happens when the lender cannot commit to enforce payments. In this case, enforcement only occurs when there are ex-post incentives. I consider two cases: when enforcement costs are low ($\gamma < \kappa$) and high ($\gamma > \kappa$). All proofs are in Appendix A.

Proposition 1. *Let $\gamma < \kappa$. The unique equilibrium in deterministic (pure) strategies is: $r(0) = r(1) = 0$ and $r(G) = R$ (the borrower defaults in full in states 0 and 1 and pays R in G); $e(r) = 1$ if and only if $r < R$ (the lender enforces payment in the case of default); off-equilibrium beliefs are chosen appropriately.*[14]

The intuition is that because enforcement costs are lower than the expected project value in states in which the borrower cannot repay R, the lender has ex-post incentives to monitor and thus equilibrium is the same as if the lender has ex-ante commitment to monitor. Furthermore, when $\gamma < \kappa$ expected payoffs are given by equations (4), (5), and (6).

When $\gamma > \kappa$, if full default only occurs for states $s \in \{0, 1\}$, the lender has no ex-post incentive to enforce payment. In this case equilibrium cannot be deterministic (the assumptions of Krasa and Villamil (2000) are not satisfied) and equilibrium is in mixed strategies.

Proposition 2. *Let $\gamma > \kappa$. Equilibrium consists of mixed strategies: $r(0) = r(1) = 0$ (the borrower defaults in full in states 0 and 1), $\sigma_B(0; G) = \frac{\pi(\gamma-\kappa)}{(1-\pi)(G-\gamma)}$ and $\sigma_B(R; G) = 1 - \sigma_B(0; G)$ (the borrower stochastically pays either 0 or R in state G); $\sigma_L(0) = \frac{R}{G}$, $e(r) = 1$ for $r \in (0, R)$, and $e(r) = 0$ for $r > R$ (in the case of default the lender enforces payment with probability $\sigma_L = \frac{R}{G}$); off-equilibrium beliefs are chosen appropriately. The equilibrium is unique.*

The intuition is that strategic default in G provides incentives for enforcement, which must be stochastic to sustain equilibrium. Given the equilibrium payment and enforcement strategies, the

[14] Technically equilibria in mixed strategies may exist; however, Krasa and Villamil (2000) show that equilibria in mixed strategies are not time-consistent when agents can renegotiate the contract at $t = 2$. Since my environment is a reduced-form version of their contracting environment (I do not explicitly consider contracting choices), I ignore mixed strategy equilibria.

associated values to the borrower and the lender for an interest rate R when $\gamma > \kappa$ are

$$V_B = \left(1 - \pi + \pi \frac{\kappa}{G}\right)(G - R), \tag{7}$$

$$V_L = \left(1 - \pi \frac{G-\kappa}{G-\gamma}\right) R. \tag{8}$$

The total total utility $V_S(R)$ is given by

$$V_S = G\left(1 - \pi + \pi \frac{\kappa}{G}\right) - R\left(\frac{G-\kappa}{G-\gamma}\frac{\pi\gamma}{G}\right). \tag{9}$$

Compared to when the lender can commit to enforcement or to when $\gamma < \kappa$, the borrower gets higher utility and the lender gets lower utility.

2 Correlated Default

The setup is as before except that (i) there are N_b borrowers and one lender endowed with N_b units, and (ii) there are two aggregate states of the world, $\omega \in \{\alpha, \beta\}$, with $\Pr(\alpha) = q_0$. The distribution of project payoffs depends on the aggregate state of the world: projects take on values $s \in \{0, 1, G\}$ with probabilities $f(s;\omega)$ given in Table 2. The aggregate state ω is *not* directly observed by any agent and therefore contracts *cannot* condition on the aggregate state ω.

Table 2: Probabilities of Project Payoffs

State	0	1	G
α	$(1-a)\pi_\alpha$	$a\pi_\alpha$	$1-\pi_\alpha$
β	$(1-b)\pi_\beta$	$b\pi_\beta$	$1-\pi_\beta$

where $a, b \in [0,1], \pi_\alpha, \pi_\beta \in (0,1)$

Conditional on project realizations being in $\{0, 1\}$, when $\omega = \alpha$ the project yields 1 with probability a, and when $\omega = \beta$ the project yields 1 with probability b. Let $a \geq b$ so that α is the "good state" and β is the "bad state." How different are a and b—how different are the aggregate states—determines the correlation of projects.[15]

[15]Notice that we are interested in project correlation conditional on default states. This does not rule out more

The timing is slightly modified. As before, in period 1 borrowers invest 1 in their projects, which pay off in period 2. Now, within period 3 the lender has a sequential choice of how many projects to enforce, and his enforcement actions are only known to defaulted borrowers. In this way, the lender can learn about the aggregate state ω by observing the states of defaulted loans. As before, borrowers repay before the lender enforces, which is critical. This timing assumption rules out any (potentially interesting and important) dynamic strategic behavior for borrowers.

For exposition the model necessarily abstracts from the important issues related to observing and using information about the joint condition of loans: there is a clear and obvious signal of the aggregate state and it is perfectly obvious how loans depend on the state. In that vein, though we tend to use costly enforcement/verification models for idiosyncratic rather than for aggregate risks, using the (very tractable) costly verification framework can easily be applied in this context once we keep in mind the richer intricacies involved with observing or using aggregate information. What is important is that there is *some* information that the lender would like to learn, such as current conditions or how loans depend on those conditions.

2.1 The Lender's Strategy with Many Borrowers

Define $\kappa = q_0 a + (1 - q_0)b$ to be the ex-ante expected value conditional on the state being zero or one. When $\gamma < b$ the lender has ex-post incentives to enforce payments regardless of the aggregate state. When $b < \gamma < \kappa$ the lender has incentives to enforce payments when the aggregate state is unknown, but the lender will stop enforcing payments if the lender learns that the state is $\omega = \beta$. The most interesting parameter values to consider are $a > \gamma > \kappa$: if default only occurs for $s \in \{0, 1\}$, the lender does not have incentives to enforce payments without knowing the aggregate state, but does have incentives when $\omega = \alpha$ is known for sure (enforcing defaulted loans has negative expected value when the state is unknown, but enforcement has positive expected value when it is likely that $\omega = \alpha$).

general correlation structures. Of course, to the extent that distributions differ in other states, projects are correlated everywhere, but what matters for a debt contract is precisely the distribution of payoffs near and below the repayment level. Loan payoffs are correlated in either case, but our interest here is correlation in default. Interestingly, Bastos (2010) finds bimodal recovery rates for bank loans, with modes around zero and one, and other studies also find bimodal distributions.

We saw in the previous section that equilibrium with $\gamma > \kappa$ required stochastic default for $s = G$. If all projects were uncorrelated, then this would still be the case. I will show that when the lender can contract with enough borrowers, there is no default in equilibrium for $s = G$. I will start by supposing that in equilibrium $r(G) = R$ and $r(1) = 0$ and then state the necessary parameters for this to be true.

Lenders can learn the aggregate state from the fraction of defaulted loans and by enforcing defaulted loans, but neither borrowers nor lenders can directly observe the aggregate state. Suppose, given the fraction of defaulted loans, that the lender believes the posterior likelihood that $\omega = \alpha$ is q. Consider the value to holding a portfolio of N defaulted loans with belief $q = \Pr(\omega = \alpha)$ (possibly different from the prior q_0).

For each loan, the lender can choose to service—or enforce—a loan or not. If it is optimal to not service one loan it is optimal to not service any loans. If the lender enforces payment, she learns the state for that loan, receives a payment net the enforcement cost, and infers something about the aggregate state. Denote by ς the state of the loan serviced (the signal received after servicing; $\varsigma \in \{0,1\}$). Denote by $q(\varsigma)$ the posterior belief when a lender receives ς and has prior q. Denote $\Pr(\varsigma = 1)$ by $p_1(q) = qa + (1-q)b$. Thus we can write the value of the portfolio recursively as

$$V(q,N) = \max\{0, p_1(q)\,(1 + V(q(1),N-1)) + (1-p_1(q))\,V(q(0),N-1) - \gamma\}, \tag{10}$$

where the value to not servicing is 0, the payment received from defaulted loans; $p_1(q) = qa + (1-q)b = \Pr(\varsigma = 1)$ is the probability the serviced loan is in state-1; $1 - p_1(q) = q(1-a) + (1-q)(1-b) = \Pr(\varsigma = 0)$ is the probability the serviced loan is in state-0; $q(1) = \frac{qa}{qa+(1-q)b} > q$ is the posterior probability of $\omega = \alpha$ after servicing a loan in state-1, i.e $\Pr(\omega = \alpha | \varsigma = 1)$; $q(0) = \frac{q(1-a)}{q(1-a)+(1-q)(1-b)} < q$ is the posterior probability of $\omega = \alpha$ after servicing a loan in state-0, i.e $\Pr(\omega = \alpha | \varsigma = 0)$.

It is easy to verify that $V(q,N)$ is increasing in both its arguments. And because $p_1(q)$, $q(1)$, and $q(0)$ are all increasing in q, the servicing policy is easy to characterize. Define t_N as the minimum belief q such that the lender will enforce if $q \geq t_N$, where t_N depends on the number of remaining loans—t_N is the enforcement threshold. The more the loans, the lower the threshold.

This is our first result.

Lemma 1. *The enforcement threshold satisfies $t_{N+1} < t_N$.*

Importantly, with a sufficiently large portfolio, initial enforcement is always worthwhile no matter how low the prior likelihood of the good state. The intuition is that enforcing many loans in the good state is very valuable, and so even a small possibility that the true state is good is enough to encourage enforcement.

Proposition 3 (Enforcement Asymptotically Optimal)**.** *Suppose $\gamma < a$. Then $\lim_{N\to\infty} t_N = 0$. In other words, for any q_0 enforcement is optimal if N is sufficiently large.*

The intuition is that with a large enough portfolio of loans, lenders are willing to pay enforcement costs to learn the aggregate state and then to optimally enforce once the state is known almost surely. How many loans are needed depends on how quickly enforcement reveals the aggregate state—more correlation reveals the state more quickly and so fewer loans are needed. The more correlated are outcomes—the more distinct are the two aggregate states—the more valuable is learning. Learning happens more quickly the greater the difference between a and b, i.e., the more different are the two states. Importantly, learning has asymmetric payoffs. When learning reveals the bad state, this stops servicing, which saves costs. When learning reveals the good state, this leads to further (valuable) enforcement.[16]

Larger enforcement costs require more loans, since even in the good state enforcement is less valuable. A lower prior requires more loans because the likelihood of enforcement paying off depends on what can be gained in the good state, which depends on the number of loans. If the good state is less likely, it better be the case that the good state "pays off more" with more loans.

[16] Another implication is that if N is finite, enforcement can "cascade" to the wrong outcome. Even when the true state is the good state, a sequence of bad outcomes can push $q < t_N$ because beliefs decrease after servicing a bad loan. Thus a lender may stop enforcement even when it is in fact profitable to do so. When b is very small and a is large, this is unlikely. However, because learning happens so fast in that case, a lender may give up on enforcement very quickly. Furthermore, sufficient correlation cannot ensure enforcement for any loan size. The best one could hope for is $a = 1, b = 0$, so that servicing perfectly reveals ω. With perfectly correlated defaulted loans, servicing reveals $\omega = \alpha$ with probability q_0 and the lender gets $1-\gamma$ for each loan. With probability $1-q_0$ servicing reveals that $\omega = \beta$, and there is no use servicing. Thus, the value to enforcing the first loan is $V(q_0, N) = q_0(1-\gamma)N - (1-q_0)\gamma$. And so, for correlation to incentivize initial enforcement, the portfolio must be at least as large as N_{min} where $N_{min} = \frac{1-q_0}{q_0}\frac{\gamma}{1-\gamma}$. An corollary of Proposition 4 is that for $N > N_{min}$ there is a minimum amount of correlation—a smallest a and largest b—that incentivizes enforcement.

From the Law of Large Numbers, as $N \to \infty$ the lender will enforce loans until she learns the aggregate state almost surely (within ε of the true state). Asymptotically, when $\omega = \alpha$ every defaulted loan will be enforced with probability 1; and when $\omega = \beta$, with probability 0. Thus, when the investor lends to an asymptotically large number of borrowers, the lender will service every loan with ex-ante probability approaching q_0.

The value of a portfolio of loans increases the more distinct are the aggregate states. The intuition is that more correlation improves the speed of learning about the aggregate state because information is more valuable. Since information cannot hurt the lender, more information leads to a higher value. I investigate this possibility by holding fixed the ex-ante value of payoffs and increasing a/b. Let $\rho = a/b$ denote the level of correlation. We parameterize the value function by ρ, writing $V(q,N;\rho)$.

Proposition 4 (Correlation Increases Portfolio Value). *Suppose $\gamma < a$. Fix κ and consider two probability distributions given by ρ_1 and ρ_2 with $\rho_1 < \rho_2$, and let $V(q,N;\rho_1)$ and $V(q,N;\rho_2)$ be the value functions given those probability distributions. Then $V(q,N,\rho_2) \geq V(q,N;\rho_1)$.*

The intuition is that learning happens more quickly the greater is the difference between a and b, and learning has an asymmetric value. When learning leads to higher beliefs, servicing continues for longer. But when servicing leads to lower beliefs, enforcement stops since enforcement is not worthwhile. Thus, faster learning saves unwanted enforcement costs more quickly when the state is the bad state—even though the bad state is worse. And when the state is the good state—well, the good state is better. The proposition is true whether or not $\kappa < \gamma$ and whether or not there is stochastic default for projects in state G. Thus, even when the lender already has incentives to enforce payments, correlation is better.

The following corollary summarizes the lender's strategy.

Corollary 1. *Suppose that borrowers default only for $s \in \{0,1\}$. Then (i) when $\gamma < b$ the lender has ex-post incentives to enforce every loan and will always enforce; (ii) when $\gamma \in (b,\kappa)$ the lender will initially monitor for any loan size N, and if $N_b \to \infty$ the lender will monitor until learning the aggregate state almost surely and will stop monitoring when $\omega = \beta$; (iii) when $\gamma \in (\kappa,a)$, the* *lender will initially monitor only if $N_b \to \infty$, and then the lender will monitor until learning the aggregate state almost surely and will stop monitoring when $\omega = \beta$.*

2.2 Numerical Examples

To see how quickly the returns to a large portfolio diminish, I numerically consider the per-loan value as a function of the number of defaulted loans for several levels of correlation. In particular, I show that convergence to the asymptotic value occurs very quickly. I consider two numerical versions: (i) with constant enforcement costs, as we've seen, and (ii) with increasing enforcement costs, as a reduced-form way of capturing decreasing returns to scale.

2.2.1 Constant Enforcement Costs

As a baseline, I consider $q_0 = .8$ so that "recessions" occur 20% of the time. Cantor and Varma (2004) find average annual recovery rates from 1983-2003 ranging from 28.7% to 53.9% with first and third quartiles of 38.2 and 48.7%. As a result, I set $a = .5$ so that in good times 50% of defaulted firms have additional cash, and I set $b = .35$ so that in bad times only 35% have additional cash. This implies that $\kappa = .47$.

I set $\gamma = .485$; however, the convergence results are not qualitatively sensitive to γ. While this number may seem high (48.5% of claimable cash-flow), as a fraction of assets this number may actually be very small. For example, if $s_1 = 6$ so that 85% of the firm's value is collateralizable assets in state s_2, then γ would be only 8.1% of the value of assets, which is consistent with evidence (Bris et al. (2006) estimate bankruptcy costs to be between 0 and 20% of assets).

To establish the sensitivity to aggregate correlation I vary a and b, letting $b = .25, .15$ and fixing κ and q_0 to increase correlation. When $b = .25$ then $a = .525$, corresponding roughly to the average range of recovery rates in Cantor and Varma (2004). Figure 1 plots the value-per-loan $\frac{V(q,N)}{N}$ as a function of the number of defaulted loans N. Figure 1 also plots the maximum attainable ex-ante value, which is the asymptotic per-loan value.

As is clear, correlation significantly increases the per-loan value. More importantly, convergence occurs very quickly, and the gains from a larger portfolio taper off after relatively small

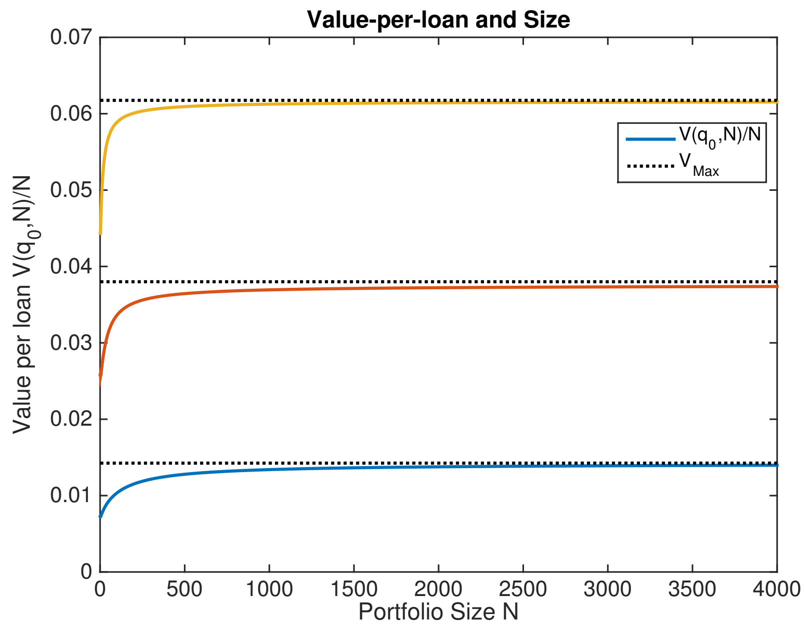


Figure 1: Value-per-loan, Loan size, and correlation. Blue line: low correlation ($b = .35$); red line: medium correlation ($b = .25$); yellow line: high correlation ($b = .15$)

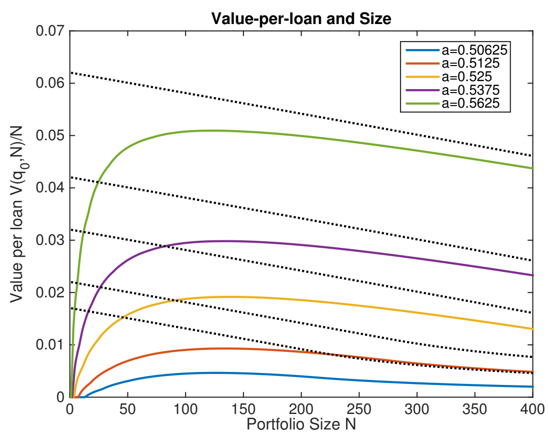
portfolios. In particular, to get within 10% of the maximum attainable value requires 516 defaulted loans in the base case (which if anything is a conservative estimate on the rate of learning), and only 125 and 36 defaulted loans in the cases with medium and high levels of correlation.

This result is important because there is mixed evidence that banks experience returns to scale, with evidence suggesting that returns to scale taper out for small portfolios. My model is consistent with this evidence, showing that the returns to strategic monitoring arising from correlation are present for small portfolios but diminish (though not completely) for large N. This is because typically lenders can quickly learn the aggregate state; as a result, the gains from learning diminish when N is large, since the state is already known for N not so large. The lender can pursue an optimal enforcement strategy without needing an asymptotically large portfolio.

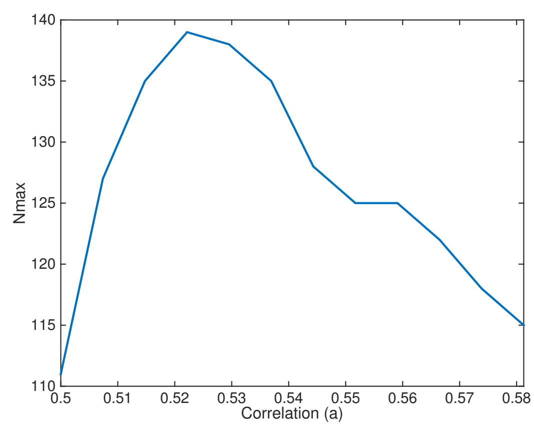
2.2.2 Increasing Enforcement Costs

The results of this section suggest that the optimal lender size is infinite, but evidence suggests that returns to scale in banking are limited to small size or to market-based activities (Laeven et al., 2014). As a reduced-form way of capturing diminishing returns to scale in credit-based activities, I consider when the enforcement cost is an increasing function of the number of enforced loans (e.g., Cerasi and Daltung 2000). As a result, the marginal cost of managing a portfolio is increasing.

Theoretically, Proposition 3 no longer applies since costs may increase too quickly to make enforcement optimal. However, Proposition 4 continues to hold, so correlation increases portfolio value even when costs are increasing. This is because (i) more correlation means a higher a so monitoring in the good state is worthwhile for a larger number of loans, and (ii) more correlation enables faster learning so that the lender can pursue optimal enforcement quickly. Finally, with increasing enforcement costs the optimal portfolio size is finite. (I consider the number of defaulted loans, but since defaulted loans are just a fraction of the total loans this focus is without loss of generality.)



(a) Value-per-loan, Loan size, and correlation. Solid lines are average loan value and dashed lines are maximum ex-ante value.



(b) Portfolio size that maximizes average loan value given correlation.

Figure 2: Increasing Enforcement Costs.

For a numerical example, I suppose enforcement costs are given by $\gamma(n) = \gamma e^{\frac{2(n-1)}{10000}}$; however, results are robust to other specifications (whether costs are linear, concave, or convex). Figure 2a plots value-per-loan as a function of portfolio size for different degrees of correlation. As expected, value-per-loan decreases as portfolios grow because enforcement costs are increasing, but higher correlation increases portfolio value. In addition, the maximum average loan value occurs for a finite portfolio, and the optimal portfolio size varies with correlation.

Figure 2b plots the maximize average loan value as a function of the correlation, with a higher a corresponding to higher correlation. For low levels of correlation, more correlation increases the viability of lending (by providing higher returns to learning), but the relationship is non-monotonic. This is perhaps surprising, but the intuition is that with more correlation the lender can quickly learn the aggregate state before returns diminish. This result is robust to specifications for the cost function, and could imply that lenders hold smaller portfolios when facing greater degrees of aggregate risk; this result is complementary to Krasa and Villamil (1992b) who show that more aggregate risk increases the monitoring costs for depositors in an intermediary.

2.3 Equilibrium with Many Borrowers

For a very large number of borrowers, the probability of being monitored after default is q_0 since the lender will enforce all loans for sure when the aggregate state is α and will stop enforcement when the aggregate state is β. Enforcement depends on the aggregate state even though the state is not directly observed. From equation (14), borrowers will not default in G if the enforcement probability is at least $\frac{R}{G}$. I suppose that the α-state is sufficiently likely, requiring $q_0 > \frac{R}{G}$. Because borrowers repay in $t = 2$, before lenders can learn the aggregate state, borrowers will *always* repay regardless of the realization of the aggregate state. With this condition, given our early results equilibrium is as follows.

Proposition 5. *Let $N_b \to \infty$. (i) Suppose $\gamma \in (b,a)$ and $q_0 > \frac{R}{G}$. Then borrowers repay $r(0) = r(1) = 0$, $r(G) = R$. The lender enforces payment so long as $q \geq t_N$, enforcing all payments when $\omega = \alpha$ and leaving an infinite number of unenforced loans when $\omega = \beta$. (ii) Suppose $\gamma < b$. Then borrowers repay $r(0) = r(1) = 0$, $r(G) = R$ and the lender enforces all defaulted loans.*

The result requires that N_b be sufficiently large that the lender will enforce enough loans before stopping (when $\omega = \beta$); in this case, the probability of enforcement is sufficiently high to induce repayment from borrowers.

Welfare Consider the case when $\pi_\alpha = \pi_\beta = \pi$. Because a fraction π of all borrowers default, enforcement occurs with ex-ante probability πq_0. Denote expected payoffs by W. The total utility per loan is

$$W_S = G\left(1 - \pi + \pi\frac{\kappa}{G}\right) - \pi q_0 \gamma, \tag{11}$$

which is larger than (9) if $\frac{R}{G} < q_0 < \frac{R}{G}\frac{G-\kappa}{G-\gamma}$ (the upper bound is more likely to be satisfied the higher is the cost γ). Given R, borrowers get value W_B and lenders get value W_L per loan given by

$$W_B = (1-\pi)(G-R) + \pi(1-q_0)b, \tag{12}$$

$$W_L = (1-\pi)R + q_0\pi(a-\gamma). \tag{13}$$

Relative to (7) and (9), for the same R the lender is better off and borrowers are worse off; however, total utility is higher. This is not surprising. First, the lender is better off and borrowers worse because borrowers do not default strategically since the lender can commit to audit with a higher probability. Second, auditing creates deadweight losses and in this case auditing occurs less frequently, increasing the total resources in the economy (because there is a high probability that a defaulted loan will be audited, defaults occur less frequently and the total number of audits decrease).

In this example, once agents are in an equilibrium in which there is no default in G, lowering R does not decrease the number of states in which default occurs and thus it does not decrease the probability of default and the expected enforcement costs. In an environment with more states (see Appendix B), lowering R decreases the number of states in which default occurs, which decreases the expected enforcement costs and further improves the possibility for trade.

If $\pi_\alpha \neq \pi_\beta$, equilibrium is unchanged, but there are a few things to note. First, the fraction defaulting reveals information about the aggregate state. When the fraction defaulting perfectly reveals the state (which need not occur with a finite portfolio), with probability q_0 every defaulted loan will be monitored. Correlation is still valuable when default rates change—the result is even stronger since the lender learns the state simply by observing default rates. Second, the welfare results hold so long as π_α and π_β are not so different.

Equilibrium When Borrowers Observe the Aggregate State One might worry that the assumption that borrowers repay without knowing the aggregate state is overly restrictive. If borrowers could learn the state gradually, by simply observing what happens to other defaulting borrowers, the strategy not to monitor in the bad state would simply result in a massive strategic default wave. If borrowers observe whether they are monitored or not, this may create an incentive to default. In a dynamic setting, and definitely in a continuous time setting, the proposed strategies could not be equilibrium. Perhaps there is no realism in the assumption that lenders can economize on the cost of monitoring by simply not monitoring anybody in some states of the world. In fact, the result that a large lender can economize on costs continues to go through precisely because equilibrium strategies will incorporate how borrowers know the aggregate state (though strategies will of course necessarily change). I provide the details in the appendix.

3 Correlation and Financial Intermediation

Increasing the correlation in the portfolio increases the value of the portfolio, but for the same portfolio size, correlation increases the variance of the portfolio. The lender cannot diversify away all risk because correlation implies that there is aggregate risk, not just idiosyncratic risk. One might wonder if this spells trouble for intermediation. Diamond (1984) and Williamson (1986) show that when multiple lenders are needed to fund a single project, financial intermediation can decrease verification costs because with a large enough portfolio, intermediaries can diversify away all idiosyncratic risk to get a portfolio with a sure, risk-free return. With a risk-free return, investors who lend to the intermediary will never have to monitor. In both of these papers, intermediaries and lenders can commit to monitor. Without commitment, correlation improves the value to monitoring, even when monitoring is ex-ante worthwhile.

In fact, financial intermediation should emerge in equilibrium. A large and correlated portfolio serves as a commitment device to service loans with a particular probability. Individual lenders may not be able to commit to such a strategy, but by pooling their resources together they can. Thus, an intermediary can function, in a way, as a coalition of lenders in order to get borrowers to repay in full in state G.

Lenders are willing to grant rents to the intermediary because the intermediary can more effectively commit to audit. The intermediary offers a risk-free return to its investors even though investors have no a priori preference for risk-free securities. This is because non-contingent promises never require enforcement; investors do not need to pay a cost to ensure that they receive the correct payment. Much like in Diamond (1984) and Williamson (1986), intermediation arises as a way of economizing on enforcement costs. In those cases, the intermediary creates a benefit because she can diversify all risks. In this case, the intermediary creates a benefit because borrowers must behave differently when dealing with a large lender. Diversification of idiosyncratic risks remains crucial—it is the only way an intermediary can offer a risk-fee payment—but it is the isolation of aggregate risks that facilitates intermediation. Correlation allows the intermediary to earn rents while still providing an attractive return to investors.

3.1 Many Borrowers and Lenders

There are N_b borrowers and N_b lenders each endowed with 1 unit. I assume that $\gamma \in (b,a)$. If each lender contracts with an individual borrower, they play an equilibrium with stochastic default in state G and the lender receives V_L. Suppose instead that one lender emerges as an intermediary so that all lenders invest through the intermediary. With a large enough portfolio, the intermediary will learn the aggregate state almost surely. When $\omega = \alpha$, the intermediary will enforce every defaulted loan and collect $a - \gamma$ per defaulted loan. When $\omega = \beta$, the intermediary will enforce a finite number of loans before stopping (having learned the state with sufficient certainty) and will not enforce the remaining (because we assume $\gamma > b$). Thus, the monitoring costs per loan asymptotically go to zero, and so the lender will on average only receive a payment from non-defaulting loans.

Suppose for now that $\pi_\alpha = \pi_\beta = \pi$ and $\gamma > \kappa$. With a large enough number of agents ($N_b \to \infty$), the per-loan value of the portfolio in each state is given by Table 3. The intermediary can offer a sure payment of $R(1-\pi)$ to investors.[17] Lenders will prefer to invest in the intermediary because

[17]Technically the payment is ε smaller than this because of the initial loans that were monitored before stopping. Since we take the interest rate R as exogenous, competition among intermediaries is not meaningful. However, if intermediaries could compete through R, rates would be driven down to the point that intermediaries could just offer a

lending on their own yields expected value of $R\left(1-\pi\frac{G-\kappa}{G-\gamma}\right) < R(1-\pi)$. This is because with direct lending, borrowers will sometimes default in state G, whereas they will not default with intermediated lending. The emergence of an intermediary is quite natural in this case.

Table 3: Value-per-loan

State	Value
α	$R(1-\pi)+\pi(a-\gamma)$
β	$R(1-\pi)$

Proposition 6. *Suppose $\gamma \in (\kappa, a)$ and let $N_b \to \infty$. Suppose an intermediary can be set up at no cost. Then in equilibrium, lenders deposit with the intermediary in exchange for a payment of $R(1-\pi)$ in period 3. The intermediary lends to borrowers, who behave as in Proposition 5.*

Proof. The intermediary borrows from individual lenders and agrees to repay $R(1-\pi)$. Individual lenders contracting directly with borrowers would get an expected value of $R\left(1-\pi\frac{G-\kappa}{G-\gamma}\right)$, which is less than $R(1-\pi)$, because an individual lender cannot commit to monitor and thus their borrowers would strategically default. Hence, individual lenders prefer to invest with the intermediary.

From Proposition 5, a lender with a large portfolio can commit to audit in the good state, and thus borrowers default in $0,1$ but never in state G. Thus, the intermediary can deliver a value of $R(1-\pi)$ to investors (because the loan portfolio is asymptotically large). When $\omega=\beta$ the intermediary gets nothing, and the intermediary gets all payments above $R(1-\pi)$ when $\omega=\alpha$. □

As in Proposition 5, while the result holds asymptotically what is required are sufficiently large intermediaries. More correlation is clearly better in this context as well. Since the intermediary will enforce every payment when $\omega=\alpha$, a higher a is strictly better. And since loans are not enforced when $\omega=\beta$, a lower b does not hurt. More correlation conditional on default is better. It might not be surprising to see intermediaries who specialize in particular markets.

If $\pi_\alpha \neq \pi_\beta$, then the value-per-loan of a large portfolio is now given by Table 4. In this case, an intermediary can guarantee a payment of $R(1-\pi_\beta)$, and the emergence of intermediaries depends

risk-free rate. Thus, profits (arising from payoffs in α) would persist.

Table 4: Value-per-loan

State	Value
α	$R(1-\pi_\alpha)+\pi_\alpha(a-\gamma)$
β	$R(1-\pi_\beta)$

on whether $V_L(R) < R(1-\pi_\beta)$, which depends on parameter values. If individual lenders prefer direct lending to a sure payment of $R(1-\pi_\beta)$, then intermediation is possible only if investors also get a larger payment in $\omega = \alpha$. When investors cannot force the intermediary to make payments, investing through an intermediary would require "monitoring the monitor" when $\omega = \beta$ to verify that payments are low because of the aggregate state and not because the intermediary is shirking.

Is more correlation still better? There is a trade-off if one cannot change a without changing π_α. So long as investors prefer the risk-free payment, more correlation is better. In fact, the optimal level of correlation is such that intermediaries can just guarantee a preferred risk-free payment. However, if for some reason intermediated lending is possible even without risk-free promises, more correlation is better because investors can learn the aggregate state more quickly and therefore spend less resources monitoring the monitor in the case of default. Thus, it is both easier for *the* monitor *to* monitor, and for investors *to* monitor *the* monitor.

If investors do not have a problem monitoring the monitor, then an intermediary emerges even when $\gamma \in (b, \kappa)$, when enforcement costs are relatively small. This is because the intermediary can strategically enforce upon learning the aggregate state (in particular, the intermediary can stop enforcements when $\omega = \beta$) and so can offer a higher expected payoff, though not a risk-free payoff, to its investors.[18]

Finally, there are surely other considerations for why an intermediary might want to avoid aggregate risk. The result of this model is that given the circumstances considered in this paper, correlated investments increase portfolio value *on the margin* compared to any other considerations outside of the model. Of course there are reasons why intermediaries would want to diversify

[18] Khalil et al. (2007) consider an environment when multiple lenders face coordination problems monitoring a single agent. They show that if coordination problems are severe such that financiers choose their monitoring efforts independently, free riding in monitoring efforts reduces the incentive to monitor, and free riding may be so strong that there may even be less monitoring compared to when the financiers fully cooperate and merge as one. Thus, my results provide an additional mechanism for why an intermediary can improve monitoring.

aggregate risks, but this paper suggests that, even so, intermediaries may have some incentive to hold some risk and doing so could be advantageous.

3.2 Implications

In this section I consider empirical implications, discuss how alternative arrangements could approximate or replace a financial intermediary, and discuss related implications of the model when considered in richer theoretical settings.

3.2.1 Empirical Implications

The model provides a number of empirical implications that could be valuable for future research. First, the model implies that large intermediaries deal with loans differently from individual lenders:

(i) Ex-post, intermediaries will monitor and enforce non-performing loans differently, either more frequently or using costlier actions (the model predicts deterministic enforcement by intermediaries in cases when direct lenders audit stochastically). There is some empirical evidence that bankruptcy procedures differ when the lender is a bank (Gilson et al., 1990; Chatterjee et al., 1996; Gupton et al., 2000), and future research should determine if these differences constitute an important source of the investment return.

(ii) Intermediaries' decisions will depend on current conditions (state contingent) in ways that the decisions of individual lenders will not. For example, do we see banks and individual lenders differentially choosing Chapter 7 or Chapter 11 when aggregate conditions change?

(iii) Intermediaries' decisions may change as information as revealed (in the model auditing ceases when intermediaries receive sufficient unfavorable information). Chapter 7 and 11 filings change as events develop and information improves, but can some of the change in behavior be attributed to what is learned from previous monitoring results?

Broadly interpreting the model, *interim* monitoring decisions should differ as well. It would be valuable to understand qualitatively how lenders' decisions differ, and how ex-ante lending arrangements reflect these differences. Crucially, since the model implies that these differences reflect intermediaries' informational advantage, these differences should be concentrated in asset-classes with correlated risks that are difficult to learn. Care must be taken because these differences will lead to ex-ante selection for borrowers.

Second, the model predicts that in equilibrium borrowers behave differently when dealing with large lenders, truthfully revealing information (repaying) in good states of the world with greater frequency than when contracting with direct lenders. As well, since borrowers may face different interim monitoring actions from intermediaries, borrowers who are financed by intermediaries may take different interim actions while managing projects, perhaps adopting different reporting systems of accounting rules to self-discipline to avoid costly monitoring. Empirical research might attempt to identify differences for identical borrowers who exogenously end up with large or small lenders, perhaps using regulatory environments or bank failures as instruments.

Third, the model implies that these equilibrium differences in lenders' and borrowers' behaviors constitute an important source of intermediaries' returns. In other words, intermediaries are distinct because of how they manage the assets, not just because of the classes of assets in which they invest. How intermediaries behave differently, or elicit different behavior from their lenders, ought to have some explanatory power for investment returns. Furthermore, banks have a greater ability to supply risk-free debt as a result of the informational advantage. One might find evidence of this by looking at US banking in the 1800s.

Fourth, the model implies that bank size and profitability should be related to the amount of correlated risk in banks' assets, though this relationship may be non-linear. Assets with more correlated risk allow for faster learning, implying that banks need not be as large in order to capture informational advantages. But more correlation also provides a greater informational advantage. Consistent with the evidence that banks achieve scale by changing the composition of risk in their portfolios, empirical research ought to find that banks take on different degrees of correlated risk as portfolios grow.

Fifth, in recognition that bank assets are correlated but hard-to-value, identifying the sources of intermediaries' returns could proceed as follows. The econometrician could estimate factors determining the returns to bank assets (perhaps using principal-component analysis). The factors of interest are those which do not correlate highly with information that is easily observable in real time. These factors need not have the most explanatory power for returns overall but should have explanatory power for the differential returns that intermediaries earn for these assets. However, because bank assets are opaque and the model implies that banks profit from their information about these assets, a purely statistical exercise may find that the factors of interest look a lot like white noise (the econometrician, like the outside investor, lacks the information that banks presumably use to identify and manage these assets, information learned from managing these assets). For this reason, instruments and natural experiments, such as regulation, deregulation, and bank failures, might be critical for identifying the informational source of intermediaries' returns.

Finally, the model provides implications for the value of bank mergers. Since informational rents arising from correlated risks are central to intermediation in the model, mergers between banks with correlated or uncorrelated portfolios have different implications for the complementarity of informational advantages. In the case of same-sector mergers, informational advantages for one portfolio are valuable for the other, thus creating complementarities; for mergers between uncorrelated banks, information about one portfolio is not useful for the other. Future empirical can analyze how deregulations affected the compositions of bank risks following deregulation. Do banks consolidate in order to diversify, or do same-sector mergers occur? Put differently, when mergers to diversify occur, do banks' behavior change following the merger (because there are no complementarities between informational advantages for uncorrelated portfolios), and does the merged bank change the composition of the balance sheet after the fact? (i.e., do they do more than simply combine portfolios?) When same-sector mergers occur and informational advantages compound, what is the effect on loan pricing?

3.2.2 Related Implications

These broader implications emerge when the mechanisms of the model are considered in richer theoretical settings. Thus, these implications provide additional theoretical and empirical insights that may be valuable for future research.

Markets for information Because intermediaries benefit from informational rents, it is worth considering whether markets for information could implement similar results. For a market for information to function, the market must exhibit frictions—perhaps lenders must search for buyers of information. In this case, one interpretation of the intermediary is that it solves the "search and matching" problem that would exist in a decentralized market.[19] An alternative institution would be for lenders to coordinate the sequence in which they enforce and provide information to the following lender(s). Of course, such a coordination problem is immense (consider also the holdup problem that would arise). Hence, intermediation should arise precisely in asset-classes where markets for information are least effective.

There are several reasons to expect that markets for information may not always function as effectively as an intermediary. Consider an economy without an intermediary, with only direct investors; suppose $\gamma > \kappa$ and consider an equilibrium (if it exists) in which borrowers do not strategically default in G. A lender would only be willing to enforce payment only if it could profit from selling that information. Since the expected payoff from enforcing is $\kappa - \gamma$, the lender must be able to sell the information for at least $\gamma - \kappa$.

First, for this arrangement to be feasible, enforcement must be observable (the lender could claim to have enforced payment and report the information that it received nothing). Second, consider a perfectly competitive market for information. Information is priced according to the quality of the information, which is to say the number of observations n. Let the price of n observations be τ_n. Define $\bar{N} = \max\{N | V(q_0, N) \leq 0\}$, which is the minimum portfolio size for enforcement to have positive value. An agent will be willing to buy $n < \bar{N}$ observations in order to then sell $n+1$ observations after enforcing her own project. But because an agent can sell the $n+1$ observations for τ_{n+1} to an arbitrarily large number of agents, paying τ_n, enforcing, and selling $n+1$ observations could yield infinite profits. Thus, in equilibrium information must have zero price (so that profits are finite), implying that no agent would have incentive to produce information to start.

[19] See Hirshleifer (1971) and also Campbell and Kracaw (1980), who show that intermediaries emerge as information producers because the production of information, the protection of confidentiality, the provision of transactions services, as well as other intermediary services, are naturally complimentary activities. In particular, moral hazard problems are resolved by the information producer having a large enough position in the market so as to be reliable.

Multi-sector Risk The model suggests that intermediaries can earn informational rents by investing in loans that are subject to correlated risk. Intermediaries may optimally invest in multiple sectors and capture informational advantages in each sector. (Suppose loans are index by sector i with payoffs distributed according to $f_i(s;\omega_i)$, where ω_i is a shock to the condition of sector i; avery large intermediary can invest sufficiently large portfolios in many sectors to learn the ω_i for each sector.) In reality there may be factors that would limit the return to investing entirely in one sector (including desires to diversify to minimize default risk). Thus, intermediaries should invest in loans with large sector-specific risk, without necessarily needing to invest exclusively in one sector. Additionally, to the extent that sector-specific shocks are correlated, informational rents would spill-over across sectors.

Opacity The model implies an additional reason why banks may choose to be "opaque" and hard to manage, or in the language of Dang et al. (2014), why banks may choose to keep the realization of projects secret. If the intermediary's profits and actions are observable, then the condition of the loan portfolio is observable and there is a "free-rider problem" for individual investors (rather than depositing with the intermediary, invest directly, observe the profits of the banks, and optimally decide to enforce in light of the information received). This free-rider problem could unravel equilibrium—unless banks could maintain an opaque balance sheet. Additionally, free riding could take the form of financial intermediaries choosing to diversify and learning aggregate information by observing banks that do not. Investing in many sectors with large sector-specific shocks could serve to increase opacity.

Secondary Markets The model assumes that defaulted debt cannot be resold to a third party, but debt collectors could specialize in particular types of loans and economize on the cost of monitoring on behalf of the lenders. This institutional arrangement is isomorphic to (splitting) the intermediary described above: individual lenders lend to borrowers and sell defaulted loans to the debt collector. Because the collector can pursue the optimal auditing strategy for a large number of loans, the collector will audit with the same strategy as an intermediary and borrowers will not strategically default. Gains from correlation will go to either the financial intermediary or to the secondary markets, which in part get passed on to lenders by means of lower default rates from borrowers. How these institutions arise would depend on features outside of the model.[20]

Hedging Aggregate Risk In the model, the gains to intermediation are due to the lender knowing/learning the aggregate state, not to having *exposure* to the aggregate state per se. There is nothing in principle that would prevent intermediaries from hedging the aggregate state if possible. However, if hedging instruments existed, the aggregate state would necessarily be observable and so gains from intermediation would disappear. The underlying issue is precisely that the aggregate state may be hard to learn through either indicators or markets. In fact, even when indicators of aggregate states exist, markets on the state often do not.[21]

Federal Deposit Insurance One argument for federal deposit insurance (rather than private markets) is that tax revenue can cover any shortfall in the event of systemic failures. Additionally, a federal agency can learn about the condition of the banking system in order to act effectively and protect the system. Crucially, the FDIC decides not only how much to charge for deposit insurance, but also when to shut down non-performing banks. The FDIC is responsible for identifying, monitoring, and addressing risks to financial institutions; and for limiting the effect of bank failure on the economy and the financial system. When a bank fails, the FDIC becomes the receiver for its assets and is responsible for selling those assets. When the FDIC takes a bank into receivership and learns the true value of its assets, which often differ from the book value,[22] the FDIC learns

[20] I am indebted to an anonymous referee for pointing out this issue. Interestingly, in the U.S. such a market exists. While the market focuses on consumer loans, it also involves some (small) business loans. The volume in this market is fairly high and it is growing pretty rapidly. Second, since early 90s lenders increasingly outsource debt collection or even sell them. Today even large lenders sell uncollected debt or outsource debt collection, closing their in-house department or making them do the first round of collection. Debt collection agencies typically deal with many lenders and in the 90s there has been major consolidation among them. The market is now fairly concentrated. For another example consider trade credit (see Eisfeldt and Rampini (2009)).

[21] Consider the mortgage market. Investors who wanted to hedge aggregate risk would like to buy credit-default swaps on an index of subprime securities; however, contracts on the ABX didn't emerge in liquid form until very late in the mortgage market, and so for many years investors had no way of hedging this risk. Similarly, contracts on aggregate house prices, as in the Case-Shiller index, emerged very late and the market is highly illiquid. Why these markets would emerge late and be illiquid is beyond the scope of the paper, but the fact that these markets were either nonexistent or unusable precisely underlies the exercise of this paper.

[22] The FDICIA requires "prompt corrective action" to shut down a bank when equity is 2% of assets. For the 433 banks closed from 2007-2014, the FDIC estimates actual losses of -26%, implying that assets commanded a significant haircut compared to book value.

about the likely condition of similar banks and can use that information to intervene effectively.

Foreclosure incentives Many economists have emphasized how securitization provides bad incentives for modifying existing loans (see Mian et al., 2011; Geanakoplos and Koniak, 2008, 2009), and it is clear from the model why securitizing loans that used to remain on banks' balance sheets would decrease incentives to monitor or modify loans ex-post. One might suspect that securitization arose precisely because investors believed that these incentives were no longer as important, that loan services would have incentives to monitor. A possible explanation for the collapse of securitization in subprime loans could be the recognition that these incentives are important.

4 Conclusion

Banks invest in hard-to-value assets with correlated risks that require information production to effectively manage. When investors require information about their assets, and this information is costly to observe, correlation increases the value of that information. This paper presents a theory of financial intermediation in which, because borrower's payoffs are correlated, a large intermediary can commit to an enforcement policy that individual lenders cannot. By holding a large, correlated portfolio of loans, an intermediary can commit to monitor ex-post with a potentially higher frequency. As a result, an intermediary can offer more competitive interest rates and borrowers default less frequently. In this environment, an intermediary can offer its investors a risk free payment which dominates the expected payments lenders could get by investing on their own, because the intermediary's borrowers default less frequently. Furthermore, intermediaries earn positive expected profits and take on systemic risk.

Lenders have every incentive to decrease expected loan-servicing costs in order to offer more affordable rates. When lenders cannot observe all current information that is relevant for managing investments, correlation can be a desirable portfolio feature and intermediation may arise naturally when borrowers have correlated outcomes. By making loans with correlated risks, the lender can minimize the costs of servicing defaulted loans. Lenders have incentives to take on correlated risks, while still diversifying away idiosyncratic risks.

Finally, there is an old banking joke that banker's work according to "3-6-3": borrow at 3%, lend at 6%, on the golf course by 3pm; and the "on the golf course" part of the joke is, in fact, not a joke.[23] Bankers spend a lot of time in the community, forming networks, getting new business, and gathering information about business conditions. The information acquired from existing loans is useful for making and managing loans to other clients. This is entirely consistent with my model of a financial intermediary who makes a large number of correlated loans and uses information about one loan to learn about the condition of others.

[23] I am grateful to Michael Kelly for this anecdote, who was first told it by a banker at JP Morgan.

References

ACHARYA, V. V. (2009): "A Theory of Systemic Risk and Design of Prudential Bank Regulation," *Journal of Financial Stability*, 5, 224–255.

ACHARYA, V. V., S. T. BHARATH, AND A. SRINIVASAN (2007): "Does industry-wide distress affect defaulted firms? Evidence from creditor recoveries," *Journal of Financial Economics*, 85, 787–821.

ANDERSON, R. W. AND K. JOEVEER (2012): "Bankers and bank investors: Reconsidering the economies of scale in banking," *CEPR Discussion Paper No. DP9146*.

BASTOS, J. A. (2010): "Forecasting bank loans loss-given-default," *Journal of Banking & Finance*, 34, 2510–2517.

BEN-PORATH, E. AND E. DEKEL (1992): "Signaling future actions and the potential for sacrifice," *Journal of Economic Theory*, 57, 36 – 51.

BERGER, A. N. AND S. M. DAVIES (1998): "The information content of bank examinations," *Journal of Financial Services Research*, 14, 117–144.

BERGER, A. N., N. H. MILLER, M. A. PETERSEN, R. G. RAJAN, AND J. C. STEIN (2005): "Does function follow organizational form? Evidence from the lending practices of large and small banks," *Journal of Financial economics*, 76, 237–269.

BERNSTEIN, S., E. COLONNELLI, AND B. IVERSON (2015): "Asset Reallocation in Bankruptcy," *Available at SSRN*.

BOTSCH, M. AND V. VANASCO (2015): "Relationship Lending: Do Banks Learn?" Tech. rep., Stanford University.

BOYD, J. H. AND E. C. PRESCOTT (1985): "Financial Intermediary-Coalitions," Staff Report 87, Federal Reserve Bank of Minneapolis.

BRIS, A., I. WELCH, AND N. ZHU (2006): "The Costs of Bankruptcy: Chapter 7 Liquidation versus Chapter 11 Reorganization," *The Journal of Finance*, 61, 1253–1303.

BRUCHE, M. AND C. GONZALEZ-AGUADO (2010): "Recovery rates, default probabilities, and the credit cycle," *Journal of Banking & Finance*, 34, 754–764.

CAMPBELL, T. S. AND W. A. KRACAW (1980): "Information production, market signalling, and the theory of financial intermediation," *The Journal of Finance*, 35, 863–882.

CANTOR, R. AND P. VARMA (2004): "Determinants of Recovery Rates on Defaulted Bonds and Loans for North American Corporate Issuers: 1983-2003," *Journal of Fixed Income*.

CERASI, V. AND S. DALTUNG (2000): "The optimal size of a bank: Costs and benefits of diversification," *European Economic Review*, 44, 1701–1726.

CHATTERJEE, S., U. S. DHILLON, AND G. G. RAMÍREZ (1996): "Resolution of Financial Distress: Debt Restructurings via Chapter 11, Prepackaged Bankruptcies, and Workouts," *Financial Management*, 25, 5–18.

CHEN, B. (2012): "All-or-nothing payments," *Journal of Mathematical Economics*, 48, 133 – 142.

COLE, R. A., L. G. GOLDBERG, AND L. J. WHITE (2004): "Cookie cutter vs. character: The micro structure of small business lending by large and small banks," *Journal of financial and quantitative analysis*, 39, 227–251.

DANG, T. V., G. GORTON, B. HOLMSTRÖM, AND G. ORDONEZ (2014): "Banks as secret keepers," Tech. rep., National Bureau of Economic Research.

DEYOUNG, R. (2008): "Safety, Soundness, and the Evolution of the U.S. Banking Industry," *Handbook of financial intermediation and banking*, 347–374.

DEYOUNG, R., M. J. FLANNERY, W. W. LANG, AND S. M. SORESCU (2001): "The information content of bank exam ratings and subordinated debt prices," *Journal of Money, Credit and Banking*, 900–925.

DIAMOND, D. W. (1984): "Financial Intermediation and Delegated Monitoring," *The Review of Economic Studies*, 51, 393–414.

DIAMOND, D. W. AND R. G. RAJAN (2000): "A Theory of Bank Capital," *Journal of Finance*, 55, 2431–2465.

DJANKOV, S., O. HART, C. MCLIESH, AND A. SHLEIFER (2006): "Debt enforcement around the world," Tech. rep., National Bureau of Economic Research.

EISFELDT, A. L. AND A. A. RAMPINI (2009): "Leasing, Ability to Repossess, and Debt Capacity," *Review of Financial Studies*, 22, 1621–1657.

FARHI, E. AND J. TIROLE (2012): "Collective Moral Hazard, Maturity Mismatch, and Systemic Bailouts," *American Economic Review*, 102, 60–93.

GALE, D. AND M. HELLWIG (1985): "Incentive-compatible debt contracts: The one period problem," *Review of Economic Studies*, 52, 647–663.

GÂRLEANU, N., S. PANAGEAS, AND J. YU (2013): "Financial Entanglement: A Theory of Incomplete Integration, Leverage, Crashes, and Contagion," Tech. rep., National Bureau of Economic Research.

GEANAKOPLOS, J. AND S. P. KONIAK (2008): "Mortgage Justice is Blind," *New York Times, Op-Ed*.

——— (2009): "Matters of principal," *New York Times, op-ed*.

GILSON, S. C., K. JOHN, AND L. H. LANG (1990): "Troubled debt restructurings: An empirical study of private reorganization of firms in default," *Journal of financial economics*, 27, 315–353.

GORTON, G. AND G. PENNACCHI (1990): "Financial Intermediaries and Liquidity Creation," *Journal of Finance*, 45, pp. 49–71.

GORTON, G. AND A. WINTON (2003): "Financial intermediation," *Handbook of the Economics of Finance*, 1, 431–552.

GUPTON, G. M., D. GATES, AND L. V. CARTY (2000): "Bank-Loan Loss Given Default," *Moody's Investors Service, Global Credit Research*.

HAGGARD, K. S. AND J. S. HOWE (2007): "Are banks opaque?" *International Review of Accounting, Banking and Finance ()*.

HELLWIG, M. (1998): "Allowing for Risk Choices in Diamond's" Financial Intermediation as Delegated Monitoring"," Tech. rep., Sonderforschungsbereich 504, Universität Mannheim & Sonderforschungsbereich 504, University of Mannheim.

HIRSHLEIFER, J. (1971): "The private and social value of information and the reward to inventive activity," *The American Economic Review*, 561–574.

HUGHES, J. P. AND L. J. MESTER (2013): "Who said large banks don't experience scale economies? Evidence from a risk-return-driven cost function," *Journal of Financial Intermediation*, 22, 559–585.

IANNOTTA, G. (2006): "Testing for opaqueness in the European banking industry: evidence from bond credit ratings," *Journal of Financial Services Research*, 30, 287–309.

KAHN, C. AND A. WINTON (2004): "Moral hazard and optimal subsidiary structure for financial institutions," *The Journal of Finance*, 59, 2531–2575.

KHALIL, F., D. MARTIMORT, AND B. PARIGI (2007): "Monitoring a common agent: Implications for financial contracting," *Journal of Economic Theory*, 135, 35–67.

KHALIL, F. AND B. M. PARIGI (1998): "Loan Size as a Commitment Device," *International Economic Review*, 39, pp. 135–150.

KRASA, S. AND A. P. VILLAMIL (1992a): "Monitoring the monitor: an incentive structure for a financial intermediary," *Journal of Economic Theory*, 57, 197–221.

——— (1992b): "A theory of optimal bank size," *Oxford Economic Papers*, 725–749.

——— (2000): “Optimal Contracts when Enforcement is a Decision Variable,” *Econometrica*, 68, 119–134.

LAEVEN, M. L., L. RATNOVSKI, AND H. TONG (2014): *Bank size and systemic risk*, 14, International Monetary Fund.

LI, W. AND M. J. WHITE (2009): “Mortgage default, foreclosure, and bankruptcy,” Tech. rep., National Bureau of Economic Research.

LUBBEN, S. J. (2012): “Resolution, Orderly and Otherwise: B of A in OLA,” *University of Cincinnati Law Review, Forthcoming*.

MELUMAD, N. D. AND D. MOOKHERJEE (1989): “Delegation as Commitment: The Case of Income Tax Audits,” *The RAND Journal of Economics*, 20, pp. 139–163.

MIAN, A., A. SUFI, AND F. TREBBI (2011): “Foreclosures, House Prices, and the Real Economy,” Working Paper 16685, National Bureau of Economic Research.

MOOKHERJEE, D. AND I. PNG (1989): “Optimal Auditing, Insurance, and Redistribution,” *The Quarterly Journal of Economics*, 104, 399–415.

MORGAN, D. P. (2002): “Rating Banks: Risk and Uncertainty in an Opaque Industry,” *American Economic Review*, 92, 874–888.

NAKAMURA, L. I. (1993): “Commercial bank information: information for the structure of banking,” *Structural change in banking. Homewood: Business One Irwin*.

STRAHAN, P. E. (2008): “Bank structure and lending: what we do and do not know,” *Handbook of financial intermediation and banking*, 107–131.

TOWNSEND, R. M. (1979): “Optimal Contracts and Competitive Markets with Costly State Verification,” *Journal of Economic Theory*, 21, pp. 265–293.

VAN NIEUWERBURGH, S. AND L. VELDKAMP (2010): “Information Acquisition and Under-Diversification,” *The Review of Economic Studies*, 77, 779–805.

WILLIAMSON, S. D. (1986): "Costly monitoring, financial intermediation, and equilibrium credit rationing," *Journal of Monetary Economics*, 18, 159–179.

——— (1987): "Costly Monitoring, Loan Contracts, and Equilibrium Credit Rationing," *The Quarterly Journal of Economics*, 102, pp. 135–146.

WOO, S. P. (2009): "Simultaneous distress of residential developers and their secured lenders: An analysis of bankruptcy & bank regulation," *Fordham J. Corp. & Fin. L.*, 15, 617.

ZHAO, R. R. (2008): "All-or-Nothing Monitoring," *American Economic Review*, 98, 1619–28.

Internet Appendix

A Proofs

Proof of Proposition 1.

Proof. First, $r(0) = r(1) = 0$ and $r(G) = R$ together imply that the posterior probability of payoffs, conditional on default, is $\tilde{f}(0) = 1-\kappa$, and $\tilde{f}(1) = \kappa$. Since $\gamma < \kappa$ enforcement is optimal. Second, $e(r) = 1$ for $r < R$ implies that $r(G) = R$ to avoid enforcement. For the same reason the borrower has no incentive to pay $r(1) > 0$.

No other pure-strategy equilibrium exists. In any equilibrium, $r(0) = 0$ by the resource constraint. For monitoring to occur, $r(s) = 0$ with some probability for some s, which is required or else the borrower will never repay. If either $r(G)$ or $r(1)$ is zero monitoring will occur for sure because $\gamma < \kappa$, but it is obvious that $r(G) = 0$ is not optimal for the borrower because repaying a positive amount to prevent monitoring yields a strictly better payoff. Thus if must be that $r(1) = 0$, which is the equilibrium we constructed. □

Proof of Proposition 2.

Proof. Stochastic default implies that enforcement is also stochastic. Note that $r(0)=0$. If $r(1)>0$ and $r(G)>0$, then $\tilde{f}(s=0|r=0)=1$ and it is not optimal to for the lender to enforce if $r=0$. But if there is no enforcement when $r=0$, the borrower would never voluntarily make a positive payment knowing that there would be no enforcement. Thus, in equilibrium the borrower must repay 0 in another state with positive probability.

Suppose that $r(1)>0$ and fully reveals the state (i.e., $r(1)\neq r(G)$). Then enforcement will occur only if $r(1)\leq 1-\gamma$. Since the borrower gets nothing if the lender enforces payment, the borrower is indifferent among $r\in(0,1-\gamma)$. For $r>1-\gamma$, enforcement will not occur if $r(G)>r$, and thus the borrower prefers $1-\gamma$. But if there is no enforcement in this case, then optimally $r(G)=1-\gamma$, too, in which case enforcement is optimal.

Suppose instead that $r(1)=0$ and $r(G)\neq 0$. Then $\tilde{f}(s=1|r=0)=\kappa$, which by assumption this is less than the enforcement cost, and so there will not be enforcement. Thus in equilibrium with some probability $r(G)=1$.

Consider σ_B with $r(0)=r(1)=0$ and with stochastic default in G. The borrower is willing to randomize when indifferent between defaulting in full and with some probability getting caught, or repaying in full,

$$G(1-\sigma_L)=G-R. \tag{14}$$

Thus, if $r<R$ the lender will stochastically enforce with probability $\sigma_L=\frac{R}{G}$, otherwise the borrower will not randomize. The lender is willing to randomize only if

$$\frac{\pi\kappa}{\pi+(1-\pi)\sigma_B}+G\frac{(1-\pi)\sigma_B}{\pi+(1-\pi)\sigma_B}=\gamma. \tag{15}$$

That is, enforcement has expected payoff zero. Thus, $\sigma_B=\frac{\pi(\gamma-\kappa)}{(1-\pi)(G-\gamma)}$.

Off-equilibrium beliefs ensure that the strategies are rational. This holds so long as $\tilde{f}(G|r)$ for $r\in(0,R)$ is sufficiently large so that the lender will enforce payment in this case. Thus, if $s=1$ the borrower has no incentive to offer $r\geq 1-\gamma$ to prevent enforcement. And if $s=G$ the borrower will not offer $r<R$ because then enforcement will occur for sure and the borrower will get nothing.

Finally, this equilibrium is unique. We have established that $\sigma_B(0:G) > 0$ in equilibrium. Suppose there exists $r_1 > 0$ played with positive probability in equilibrium. I show this cannot be.

First, suppose r_1 is played only if $s = 1$ but not for $s = G$. For stochastic monitoring to occur (which is required), it must be that $r_1 = 1-\gamma$, since monitoring would yield $1-\gamma$ and not would yield r_1. Let the lender enforce payment with $\sigma_L(r_1)$. Rationality for the borrower requires $(1-\sigma_L(r_1))(1-r_1) = (1-\sigma_L(0))1$, which equates the expected payoff to paying r_1 or zero. But because $\gamma < 1$, this means that $(1-\sigma_L(r_1))(G-r_1) > (1-\sigma_L(0))G$, because $(1-\sigma_L(r_1)) = \frac{(1-\sigma_L(0))}{\gamma}$ and so $(1-\sigma_L(r_1))(G-r_1) = (G-(1-\gamma))\frac{(1-\sigma_L(0))}{\gamma}$, and $G-1+\gamma > G\gamma$. As a result, in G the borrower strictly prefers to play r_1 and so $\sigma_L(r_1;G) = 1$, contrary to our hypothesis.

Second, suppose that r_1 is played for $s \in \{1, G\}$. For the borrower to be willing to randomize in each state requires

$$(1-\sigma_L(r_1))(1-r_1) = (1-\sigma_L(0))1, \quad (1-\sigma_L(r_1))(G-r_1) = (1-\sigma_L(0))G, \quad G-R = (1-\sigma_L(0))G.$$

Hence, $(1-\sigma_L(r_1)) = \frac{(1-\sigma_L(0))}{1-r_1}$ and $(1-\sigma_L(r_1)) = \frac{(1-\sigma_L(0))G}{G-r_1}$, which requires $\frac{1}{1-r_1} = \frac{G}{G-r_1}$, which requires $r_1 = 0$.

□

Proof of Lemma 1.

Proof. By the definition of t_N, t_N is the largest q such that $V(q,N+1) \leq 0$. Since $V(q,N) \leq V(q,N+1)$ and V is increasing in q, we have

$$t_N = \sup\{q|V(q,N) \leq 0\} \geq \sup\{q|V(q,N+1) \leq 0\} = t_{N+1}.$$

□

Proof of Proposition 3.

Proof. For any prior q_0, for any small $\varepsilon > 0$, there exists M such that if the lender services M loans and finds them all in the 1-state, then the posterior belief grows to $\tilde{q} > 1-\varepsilon$. This occurs with a small probability $\delta < q_0$. Hence,

$$V(\tilde{q}, N-M) \geq (N-M)\left[(1-\varepsilon)(a-\gamma)+\varepsilon(b-\gamma)\right],$$

which is the value from enforcing every remaining loan (this strategy is suboptimal and the lender could only do better). This value can be made arbitrarily large by increasing N.

The net cost of enforcing M loans is at worst $-M\gamma$, which would be the payoff from enforcing M loans in the 0-state. Thus, if a lender enforced M times without stopping, the initial value would be at least

$$V(q_0, N) \geq \delta(N-M)\left[(1-\varepsilon)(a-\gamma)+\varepsilon(b-\gamma)\right]-M\gamma.$$

The first term is the small probability that she now believes the aggregate state is α and enforcement is valuable and the second term is the what she paid to get there. Even if δ is very small and M is very large, this value can be made positive if N is sufficiently large. □

Proof of Proposition 4.

Proof. I prove the proposition inductively in steps. The following lemmas show that the proposition holds for $N = 1, 2$. I proceed by fixing b and letting q_0 vary with ρ, which is without loss of generality.[24] Since $\kappa = q_0 a + (1-q_0)b = q_0(\rho-1)b + b$, we have $q_0 = \frac{\kappa - b}{b(\rho-1)}$. The proof for $N = 1$ is trivial since there is no option value to continuing. For $N = 2$, the lender may monitor even when the expected flow payment is negative, and so it needs to be shown that the change in the option value outweighs the change in the flow payment.

Lemma 2 ($N = 1$)**.** $V(q, 1; \rho_2) \geq V(q, 1; \rho_1)$.

Proof. Define $V_1(q, 1) = q(a-b)+b-\gamma = p_1(q)-\gamma$, which is the value to monitoring one loan. Then $V(q, 1) = \max\{0, V_1(q, 1)\}$ with $V(q, 1) = V_1(q, 1)$ for $q > t_1$. Writing V_1 in terms of ρ we

[24] I do it this way because if we fix q_0, then increasing ρ increases $p_1(q)$ for $q \geq q_0$, and decreases it otherwise. For the same q_0, more correlation leads to lower expected flow payoffs when beliefs are low. Despite negative flow payoffs, the lender will monitor because the the continuation value to monitoring, compensates for the flow loss. It turns out that increasing correlation has a larger effect on the continuation value (optimally using the information from monitoring) than on the flow payoff.

have, $V_1(q,1) = q(\rho-1)b + b - \gamma$, which is increasing in ρ.[25] □

Lemma 3 ($N = 2$)**.** $V(q,2;\rho_2) \geq V(q,2;\rho_1)$.

Proof. I will show that

$$V(q,2) = \max\{0, V_1(q,2), V_2(q,2)\},$$

where $V_1(q,2) = q(a-b)(a+b+1-\gamma) + (b-\gamma)(1+b)$ and $V_2(q,2) = 2q(a-b) + 2(b-\gamma)$.

There are two possibilities for the value function when monitoring occurs. When $V(q,2) = V_2(q,2)$ it is optimal to monitor the second loan no matter what the signal for the first. When $V(q,2) = V_1(q,2)$ it is optimal to monitor the first loan and to stop if the signal is $\varsigma = 0$. Monitoring twice provides value $2V(q,1)$ and therefore $V_2(q,2) = 2V(q,1)$. This is the value for large q where $q(0) > t_1$. From the previous lemma, we know that $V_2(q,2;\rho_2) > V_2(q,2;\rho_1)$ since $V_2(q,2) = 2V_1(q,1)$.

For small q, $q(0) < t_1$, and so the second loan will not be monitored if $\varsigma = 0$. In this case

$$V_1(q,2) = (q(a-b)+b)(1+V(q(1),1)) - \gamma.$$

That is, monitoring yields $\varsigma = 1$ with probability $q(a-b)+b$ and so the monitor receives 1 and pays γ to monitor the next loan. Otherwise, the lender receives nothing and does not monitor the second loan. In either case she pays γ. Plugging in for $V(q,1)$ we have

$$\begin{aligned} V_1(q(1),1) &= \frac{qa}{q(a-b)+b}(a-b) + b - \gamma \\ &= \frac{qa^2}{q(a-b)+b} + \frac{(1-q)b^2}{q(a-b)+b} - \gamma. \end{aligned}$$

Plugging these equations in and simplifying yields the expression. Writing $V_1(q,2)$ in terms of ρ yields

$$V_1(q,2) = q(\rho-1)b(\rho b + b + 1 - \gamma) + (b-\gamma)(1+b),$$

[25] Notice that if we instead fix q_0 and vary a and b we have $V_1(q,1) = q\frac{\kappa-b}{q_0} + b - \gamma$ and so $\frac{dV_1(q,1)}{db} = 1 - \frac{q}{q_0}$, which is negative for $q > q_0$. Thus, $V(q,1;\rho_2) > V(q,1;\rho_1)$ wherever $V(q,1;\rho_1) = V_1(q,1)$, which would be true whenever $\kappa > \gamma$.

which is increasing in ρ. Thus, $V_1(q,2;\rho_2) > V_1(q,2;\rho_1)$. $\square$

Finally, assume the proposition is true for $N-1$. We can write the value function as consisting in two terms: the expected flow payoff from servicing this loan and a convex combination of continuation values. That is, for $V \geq 0$

$$
\begin{aligned}
V(q,N) &= [p_1(q) - \gamma] + [p_1(q)V(q(1),N-1) + (1-p_1(q))V(q(0),N-1)] \qquad (16)\\
&= p_1(q) - \gamma + \mathbb{E}\left[V(\tilde{q},N-1)\right].
\end{aligned}
$$

I show that $V(q,N)$ is convex in q and that more correlation acts as a mean preserving spread over $\tilde{q}$, and thus correlation increases the value. The next lemma establishes convexity.

Lemma 4 (V(q,N) convex, piecewise linear). *$V(q,N) = \max\{V_n(q,N)\}_{0\leq n\leq N}$, where $V_n(q,N)$ is linear in q and n is the number of $\varsigma = 0$ until enforcement stops. (The optimal n depends on q.) Furthermore, $V(q,N)$ is convex in q.*

Proof. The proof is by induction. First, I've already shown this is true for $N = 1,2$. Suppose true for $N-1$. Define $V_n(q,N-1) = C^n_{N-1}q + D^n_{N-1}$. Let $V(q,N) = V_m(q,N)$. For $m = N$, $V_N(q,N) = N(q(a-b)+b-\gamma) = N(p_1(q)-\gamma)$. For $m < N$ $V_m(q,N)$ can be written as

$$
\begin{aligned}
V_m(q,N) &= (q(a-b)+b)(1+V_m(q(1),N-1)) + (1-(q(a-b)+b))(V_{m-1}(q(0),N-1)) - \gamma,\\
V_m(q,N) &= q\left[(a-b)(1+D^m_{N-1} - D^{m-1}_{N-1}) + C^{m-1}_{N-1} + a(C^m_{N-1} - C^{m-1}_{N-1})\right]\\
&\qquad + b(D^m_{N-1} - D^{m-1}_{N-1}) + b + D^{m-1}_{N-1} - \gamma.
\end{aligned}
$$

Which is linear. Additionally, because of the ordering and linearity of the functions it must be that $C^m_N > C^{m-1}_N$ and $D^m_N < D^{m-1}_N$. Given an enforcement policy, a higher q linearly improves the expected value. Since the lender can only do better by choosing a different policy for a higher q, the value increases at least linearly and thus V is convex. $\square$

The next lemma shows that increasing $\rho = \frac{a}{b}$ increases the spread between $q(1)$ and $q(0)$ in a "mean-preserving" fashion. That is, given prior q, the expected posterior is always q but the variance in the posterior is higher when ρ is larger.

Lemma 5 (Learning Quickens with Correlation). *A larger ρ implies a larger $q(1)$ and smaller $q(0)$.*

Proof. Plugging in for posterior beliefs in terms of ρ and simplifying we have

$$q(1) = \frac{1}{1+\frac{1-q}{q}\frac{1}{\rho}} \quad \text{and} \quad q(0) = \frac{1}{1+\frac{1-q}{q}\left(\frac{1-b}{1-\rho b}\right)}$$

It is immediately clear that $\frac{dq(1)}{d\rho} > 0$ and $\frac{dq(1)}{d\rho} < 0$. Furthermore,

$$\mathbb{E}[\tilde{q}] = (qa + (1-q)b)q(1) + (q(1-a) + (1-q)(1-b))q(0) = q.$$

□

Because $V(\tilde{q}, N-1)$ is convex, and because increasing ρ yields a mean-preserving spread over the posterior beliefs, and by the induction hypothesis that $V(\tilde{q}, N-1; \rho_2) \geq V(\tilde{q}, N-1; \rho_1)$, then:

$$\mathbb{E}_{\rho_2}\left[V(\tilde{q}, N-1; \rho_2)\right] \geq \mathbb{E}_{\rho_1}\left[V(\tilde{q}, N-1; \rho_1)\right]. \tag{17}$$

Thus, more correlation acts as a mean-preserving-spread over the continuation value, which is beneficial because of convexity.

Finally, define $\tau(q) = q\rho + 1 - q - \frac{\gamma}{b}$. We can write

$$V(q,N) = \tau(q)b + \mathbb{E}\left[V(\tilde{q}, N-1)\right]$$

Since $\tau(q)$ is increasing in ρ and the expectation over continuation values increases with higher ρ, the induction hypothesis implies that $V(q,N;\rho_2) \geq V(q,N;\rho_1)$. □

The proposition is true whether or not $\kappa < \gamma$ and whether or not there is stochastic default for projects in state G. Thus, even when the lender already has incentives to enforce payments, correlation is better.

Proof of Proposition 5.

Proof. From Corollary 1, when $\gamma \in (b,a)$, the lender will optimally enforce loans until learning the aggregate within ε of the truth, which implies the probability of a defaulted loan getting enforced is q_0. Since borrowers do not observe the aggregate state, the probability of getting enforced if they default is q_0, which by assumption is high enough to discourage strategic default. Borrowers have no incentive to deviate because doing so would lead to enforcement for sure. When $\gamma < b$ the lender always has ex-post incentives to enforce. $\square$

Equilibrium When Borrowers Observe the Aggregate State

To simplify how dynamic interactions could lead to learning, suppose that borrowers in fact can observe the aggregate state ω at no cost, but lenders do not observe the aggregate state.[26]

With one borrower and one lender, the equilibrium is just as in Proposition 2: the borrower defaults in G with probability $\sigma_B = \frac{\pi(\gamma-\kappa)}{(1-\pi)(G-\gamma)}$ and the lender audits with probability $\sigma_L = \frac{R}{G}$. Remember that the borrower's default rate is chosen so that the lender is indifferent between auditing or not, which depends on the lender's beliefs about the state. Since the lender does not know the aggregate state, the lender's beliefs are unchanged and so the borrower will use the same default rate as before.

However, the equilibrium with a very large number of borrowers and a single lender is different. Notice that if all players behave as just described, then the lender would eventually learn the aggregate state from auditing a fraction of the defaulted loans. But once the lender learns the aggregate state, the lender would audit with probability 1 when $\omega = \alpha$ and with probability 0 when $\omega = \beta$ because the distributions do and don't provide incentives to audit. Of course borrowers know the true state, and so would know the likelihood (asymptotically) of being monitored. When $\omega = \alpha$ they will be monitored almost surely, and when $\omega = \beta$ they would almost surely not be. Thus the strategy couldn't be an equilibrium.

In fact, the equilibrium strategies are as follows. When $\omega = \beta$ borrowers default stochastically

[26] An additional consideration is whether the dynamic game should be considered a repeated game, in which case reputational concerns could allow lenders to sustain commitment, and these reputational concerns might be different for large versus small lenders.

in G with probability $\sigma_B(\beta) = \frac{\pi(\gamma-b)}{(1-\pi)(G-\gamma)}$, and always repay in G when $\omega = \alpha$. The lender audits until learning the state almost surely and then stochastically audits with probability $\sigma_L(\beta) = \frac{R}{G}$ when $\omega = \beta$, and audits for sure when $\omega = \alpha$. Notice that the lender can learn the aggregate state both from the fraction of loans in state 1, but also from any loans in state G. With probability q_0 the aggregate state is good and the lender collects $(1-\pi)R + \pi(a-\gamma)$ per loan, and with probability $1-q_0$ the aggregate state is bad and the lender gets $\left(1-\pi\frac{G-b}{G-\gamma}\right)R$, which is a modified version of equation (8). Thus, the lender's value per loan is

$$q_0\left((1-\pi)R + \pi(a-\gamma)\right) + (1-q_0)\left(1-\pi\frac{G-b}{G-\gamma}\right)R. \tag{18}$$

It is easy to check that higher correlation improves the value to the lender in this case.

B The Model with a Continuum of States

I have shown that intermediation can arise because a large intermediary can commit to monitor with a higher probability, which decreases default rates and benefits lenders. However, the result that intermediation arises when borrowers have correlated payoffs is more robust. In this section I show that intermediation arises when interest rates are endogenous: in this case, a large intermediary can offer a lower interest rate, which benefits both borrowers and lenders. I show this in a model in which borrowers' payoffs can take a continuum of values, and as a result the optimal contract is debt.

B.1 One Borrower and One Lender

The setup is as before except now projects take a stochastic value s in period 2 from $S = [s_{min}, s_{max}]$ with probability density $f(s)$, which is continuous, positive, and differentiable, and with cumulative density $F(s)$. I assume throughout that $\gamma < \mathbb{E}[s; f(s)]$ so that enforcement is optimal if the borrower never repays.

As proved by Krasa and Villamil (2000), the optimal contract is deterministic, i.e., a debt contract with face value R (principal plus interest). Borrowers default in full whenever $s \leq R$, i.e., they pay zero, and lenders enforce payment, claiming s; otherwise borrowers repay R and no enforcement occurs. Since the optimal contract is debt, we can state the lender's problem as choosing an interest rate R to maximize the lender's payoff subject to an incentive compatibility constraint that ensures ex-post incentives to audit

$$\max_{R} \quad V_L(R) = \int_{s_{min}}^{R} sf(s)ds - \gamma F(R) + R(1-F(R)), \tag{19}$$

subject to

$$\int_{s_{min}}^{R} s\frac{f(s)}{F(R)}ds \geq \gamma, \tag{20}$$

where equation (20) states that the expected benefit to monitoring a defaulted loan is at least as large as the monitoring cost.

Assuming that $V_L(R)$ is concave[27] (which requires $f(s) + \gamma f'(s) > 0$ for all $s \in S$), then when the incentive compatibility constraint does not bind the optimal interest rate R^* solves

$$1 - F(R^*) - \gamma f(R^*) = 0. \tag{21}$$

The lender's payoff is not strictly increasing in R, though the borrower's payoff is strictly decreasing in R. The reason is, as pointed out by Williamson (1987), that a higher R increases the default probability, which increases the expected auditing cost.

I assume that the incentive compatibility constraint is not satisfied at R^* so that

$$\mathbb{E}[s|s \leq R^*] = \int_{s_{min}}^{R^*} s\frac{f(s)}{F(R^*)}ds < \gamma \tag{22}$$

As a result, for the contract to provide incentives to monitor, the interest rate must exceed R^*. The equilibrium interest rate is $\hat{R} > R^*$ satisfying $\mathbb{E}\left[s|s \leq \hat{R}\right] = \int_{s_{min}}^{\hat{R}} s\frac{f(s)}{F(\hat{R})}ds = \gamma$. Since $1 - F(\hat{R}) - \gamma f(\hat{R}) < 0$, the lender's profit is strictly less than it would be if she could commit to a deterministic

[27]Concavity ensures that the optimal interest rate has an interior solution. In particular, it ensures that $R = s_{max}$ is not optimal because the deadweight cost of monitoring would be too great.

auditing strategy. Default occurs more often, leading to a large expected auditing cost; specifically, default occurs in states in which the payoff is higher, which provides incentives for the lender to audit.[28]

Proposition 7. *The lender will offer R^* whenever $\int_{s_{min}}^{R^*} s\frac{f(s)}{F(R^*)}ds > \gamma$. Otherwise, the lender will offer $\hat{R} > R^*$ determined by the incentive compatibility constraint.*

B.2 Correlated Default

Aggregate states index the distributions of payoffs: projects take on values $s \in S$ with probabilities $f_\alpha(s)$ and $f_\beta(s)$, with corresponding CDFs $F_\alpha(s)$ and $F_\beta(s)$. By definition, $f(s) = q_0 f_\alpha(s) + (1-q_0) f_\beta(s)$. I make the following assumptions about the distributions:

Monotone Likelihood Ratio: For $s < \hat{s} = \frac{\hat{R}}{q_0}$, the likelihood ratio is (weakly) monotonically increasing:

$$\frac{f_\alpha(s)}{f_\beta(s)} \quad \text{is weakly increasing in } s \text{ for } s < \hat{s}.$$

This property has two implications. First, given a realization $s < \hat{s}$, a higher s is relatively more likely when $\omega = \alpha$, so that the conditional likelihood that $\omega = \alpha$ given s is higher when s is higher. Second, for any $R < \hat{s}$, the conditional expected payoff below R is greater under α, i.e.,

$$\mathbb{E}_\alpha [s|s \leq R] > \mathbb{E}_\beta [s|s \leq R].$$

I assume that

$$\mathbb{E}_\alpha [s|s \leq R^*] > \gamma > \kappa = \mathbb{E} [s|s \leq R^*], \tag{23}$$

which implies that when the interest rate is R^*: (i) enforcing defaulted loans has negative expected value when the state is unknown, but (ii) enforcement has positive expected value when $\omega = \alpha$ is sufficiently likely.

[28] In contrast, Krasa and Villamil (2000) assume that the incentive compatibility constraint *is* satisfied for interest rates that satisfy the borrower's participation constraint. Since I am supposing there is no participation constraint for the borrower, or that it is sufficiently slack, then this result is without loss of generality.

B.2.1 The Lender's Strategy with Many Borrowers

Equilibrium with one borrower and lender required $\hat{R} > R^*$ in order to provide auditing incentives. When the lender can contract with enough borrowers, the equilibrium interest rate is $\bar{R} < \hat{R}$. I will start by supposing that this is so and then state the necessary parameters for this to be true.

Lenders can learn the aggregate state from the fraction of defaulted loans and by enforcing defaulted loans, but neither borrowers or lenders can directly observe the aggregate state. Suppose, given the fraction of defaulted loans, that the lender believes the posterior likelihood that $\omega = \alpha$ is q. Consider the value to holding a portfolio of N defaulted loans with belief $q = \Pr(\omega = \alpha)$ (possibly different from the prior q_0), and suppose that loans default in $s < \bar{s} \leq \hat{s}$. Define

$$a = \mathbb{E}_\alpha [s|s \leq \bar{s}], \quad b = \mathbb{E}_\beta [s|s \leq \bar{s}].$$

We know that $a > \gamma > b$. We can write the value of the portfolio recursively as

$$V(q,N) = \max\{0, \mathbb{E}[s + V(q(s), N-1)] - \gamma\}, \tag{24}$$

where the value to not servicing is 0, the payment received from defaulted loans. Lemma 1 is true in this setup, with identical proof. Proposition 3 holds in this setup as well, with similar proof. Proposition 4 holds. We modify the proof by noting for $V \geq 0$

$$V(q,N) = \mathbb{E}_q [s|s \leq \bar{s}] - \gamma + \mathbb{E}[V(\tilde{q}, N-1)].$$

B.2.2 Equilibrium with Many Borrowers

For a very large but finite number of borrowers, the probability of being serviced after default is q_0. Borrowers will not default for $s > \bar{R}$ if $s(1 - q_0) \leq s - \bar{R}$. Thus, borrowers will default in states $s \leq \bar{s} = \frac{\bar{R}}{q_0}$.[29] As a result, the value to the lender with a large correlated portfolio is

$$W_L(\bar{R}) = q_0 \int_{s_{min}}^{\bar{s}} s f_\alpha(s) ds - \gamma q_0 F_\alpha(\bar{s}) + \bar{R}(1 - F(\bar{s}))$$

Differentiating with respect to $\bar{R}$ yields

$$1 - F(\bar{s}) - \gamma f_\alpha(\bar{s}) + \bar{s} f_\alpha(\bar{s}) - \bar{R} f_\alpha(\bar{s}) = 0.$$

Since $\bar{s} > \bar{R}$ and $f_\alpha(s) \geq f(s)$, it must be that $\bar{R} < \hat{R}$. Additionally, if q_0 is sufficiently high, then $W(\bar{R}) > V(\hat{R})$.[30] Furthermore, in the case when $f_\alpha(s) = f(s)$ over this range, then as $q_0 \to 1$, $\bar{R} \to R^*$. Thus, equilibrium is as follows.

Proposition 8. *Let $N_b \to \infty$. The equilibrium contract is $\bar{R} < \hat{R}$. Borrowers repay $r(s) = 0$ for $s \leq \frac{\bar{R}}{q_0}$ and $r(s) = \bar{R}$ otherwise. The lender enforces payment so long as $q \geq t_N$, enforcing all payments when $\omega = \alpha$ and leaving a fraction approaching 1 of unenforced loans when $\omega = \beta$.*

Proof. As discussed, given the strategy of borrowers, the lender will optimally enforce loans until nearly learning the aggregate state, which implies the probability of a defaulted loan getting enforced is q_0. Since borrowers do not observe the aggregate state, the probability of getting enforced if they default is q_0, which by assumption is high enough to discourage strategic default. Borrowers have no incentive to deviate because doing so would lead to enforcement for sure. □

With a finite portfolio, the fraction defaulting will not perfectly reveal the aggregate state because the fraction defaulting is only asymptotically equal to $F_\omega(\bar{s})$. Even if the fraction defaulting perfectly reveals the state, that means that with probability q_0 every defaulted loan will be monitored. Thus, correlation makes this equilibrium possible.

[29] Borrowers strategically default because they know that lenders will only monitor with probability q_0. This is reminiscent of the allegation by many that after the housing crisis many homeowners did, or would, stop paying their mortgages on the assumption that lenders wouldn't do anything about it anyway. Compare this to the result in baseline model, where borrowers strategically default in the case of direct lending but not when lenders have a large portfolio. The difference here is that the set of default states change with a continuum of states, whereas in the 3-state case the set of default states cannot change, but the default probabilities must.

[30] We derive a more precise condition in the baseline model with 3 states.

B.3 Correlation and Financial Intermediation

The results of the previous model carry over, now recognizing that with a large enough number of agents ($N_b \to \infty$), the per-loan value of the portfolio in each state is given by Table 5.

Table 5: Value-per-loan

State	Value
α	$\bar{R}(1-F_\alpha(\bar{s})) + \mathbb{E}_\alpha[s \mid s \le \bar{s}] - \gamma F_\alpha(\bar{R})$
β	$\bar{R}(1-F_\beta(\bar{s}))$

The intermediary can offer a sure payment of $\bar{R}(1-F_\beta(\bar{s}))$ to investors (assuming that $F_\beta(\bar{s})$ is not so different from $F_\alpha(\bar{s})$). Lenders will prefer to invest in the intermediary whenever this value exceeds $V_L(\hat{R})$, which is what they can get from investing directly themselves. When $F_\alpha(\bar{s}) = F_\beta(\bar{s})$, this is satisfied for sure so long as q_0 is high enough since $\bar{R} < \hat{R}$, as discussed. This is because with direct lending, lenders must offer an inefficiently high interest rate to increase default. Thus, lenders will prefer to invest with the intermediary so long as $F_\beta(\bar{s})$ is not so different from $F_\alpha(\bar{s})$. The emergence of an intermediary is quite natural in this case.[31]

Proposition 9. *Let $N_b \to \infty$. In equilibrium an intermediary emerges. Lenders deposit with the intermediary in exchange for a payment of $\bar{R}(1-F_\beta(\bar{s}))$ in period 3. The intermediary lends to borrowers, who behave as in Proposition 5.*

The proof is immediate. The intermediary contracts with borrowers and the intermediary gets all payments above $\bar{R}(1-F_\beta(\bar{s}))$ when $\omega = \alpha$. Furthermore, when $N_b = \infty$, then the intermediary enforces with probability 1 when the default rates are the same. If the probabilities are the same at R^*, then the intermediary can offer the optimal interest rate R^*.

[31] Actually, even if $\bar{R}(1-F_\beta(\bar{s})) < V(\hat{R})$, an equilibrium with intermediation exists because the intermediary offers a more competitive interest rate than direct lenders can. Thus, borrowers will strictly prefer to borrow from an intermediary—there will be no direct lending possibility. Thus, this is a worse equilibrium for lenders when q_0 is sufficiently low. The intermediaries have priced out direct lending, but in doing so they cannot offer a return that is better than the return to direct lending.

Intermediation and Optimal Contracts

With one borrower and one lender, the optimal contract is debt because of the lack of enforcement commitment. It may not be the case that for a large intermediary contracting with many borrowers with correlated payoffs that the optimal contract is still debt. In general, a standard debt contract is not optimal in cases of multiple investors or multiple risky projects undertaken by the same entrepreneur. As well, stochastic contracts generally dominate when lenders can commit to an enforcement strategy.

While it is true that lenders can in effect commit to enforce by holding a large correlated portfolio, lenders cannot explicitly commit, and thus a time-consistent stochastic contract may not exist. However, it is possible that the optimal contract for a large intermediary includes partial debt forgiveness as a way of minimizing the fraction of loans audited. If so, an important result is that intermediaries will offer contracts that dominate debt, which is the optimal contract for direct lending. Thus, debt leads to intermediation, and intermediation may lead to the development of better contracts, strengthening the results of this section.