Intermediary Financing without Commitment

Yunzhi Hu and Felipe Varas*

September 4, 2020

Abstract

Banks can reduce frictions in the credit market through monitoring. To be a credible monitor, a bank needs to retain a fraction of its loans; we study the credit market dynamics when it cannot commit to do this. Loan prices drop in anticipation of loan sales and reductions in monitoring. With commitment, *intermediation* is irrelevant if the bank *certifies* it will monitor. Without commitment, a bank that only certifies sells its loans over time. By contrast, an intermediating bank that issues short-term deposits internalizes the externalities from monitoring, and therefore retains its loans. While intermediation leads to more lending, an entrepreneur with high net-worth may choose certification.

Keywords: commitment; certification; intermediation; trading; dynamic models

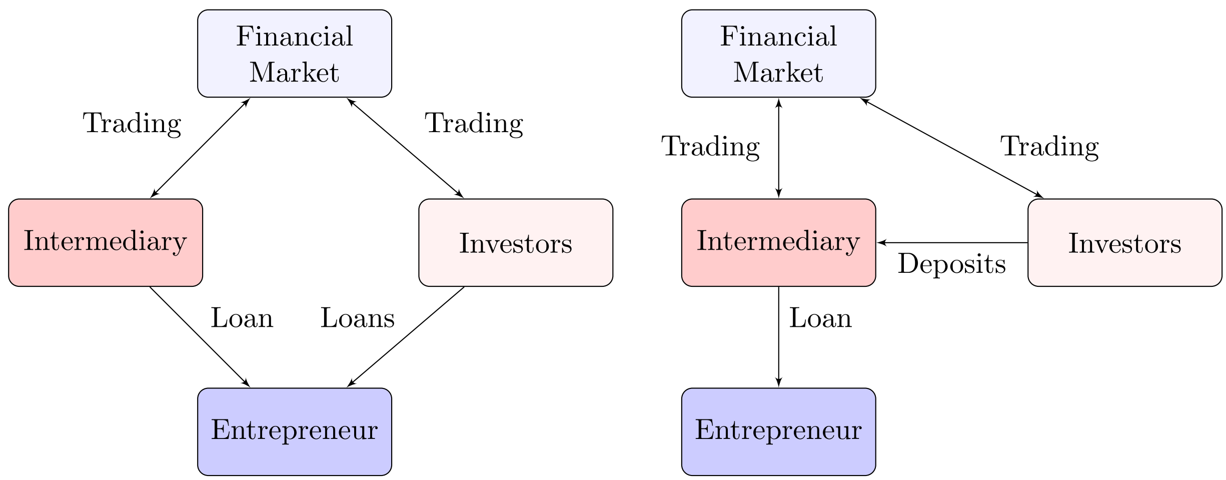
*Hu: Kenan-Flagler Business School, UNC Chapel Hill; Yunzhi_Hu@kenan-flagler.unc.edu. Varas: Fuqua School of Business, Duke University; felipe.varas@duke.edu.

1 Introduction

Financial intermediaries exist to reduce agency frictions in the credit market and therefore benefit the real economy. To conduct valuable services such as screening and monitoring, the optimal arrangement typically involves banks retaining some loans on their balance sheets as the skin in the game; otherwise, the incentives can be misaligned. This paper studies the equilibrium credit market dynamics when banks can *not* commit to retain these loans. By doing so, we develop a dynamic theory of intermediary financing without commitment.

We build on the classic model of Holmstrom and Tirole (1997) in which banks can monitor to reduce a borrower's private benefits. Being a credible monitor requires the bank to retain a sufficient fraction of loans on its balance sheet. However, with the presence of a financial market, the bank has incentives to sell its loans, due to its high cost of capital. The more it sells, the less likely it will monitor, and the price of the loans will drop more as well. We study two types of implementation structures, certification and intermediation. As shown in Figure 1, in certification, both the bank and investors directly invest into the borrower's venture, and the role of the bank is to certify that it will monitor. In intermediation, investors deposit in the bank, which then invests a collection of its own funds and the deposits into the borrower's venture. While the two structures are equivalent in the static framework, they lead to very different credit market dynamics. In certification, the lack of commitment induces the bank to sell its loans gradually, and the bank's monitoring intensity declines over time. In intermediation, the bank is able to issue short-term deposits, which helps it commit to the retention and the decisions to monitor. As a result, an intermediating bank never sells its loans. Therefore, the entrepreneur is always able to borrow more under intermediation. However, if both structures enable the entrepreneur to borrow enough to invest, the entrepreneur may end up choosing certification. In contrast with the static model, we show that the implementation structure that maximizes the initial borrowing may not be the one that maximizes the borrower's expected payoff.

Let us be more specific. An entrepreneur is endowed with an investment opportunity, which requires a fixed-size of investment and pays off some final cash flows at a random time in the future. She has limited personal wealth and needs to borrow to make up the investment shortfall. Due to moral hazard in effort choices, she can only pledge a fraction of the final output to outside creditors, including banks and investors. Banks have higher cost of capital, but, as in Holmstrom and Tirole (1997), only they can monitor to reduce the entrepreneur's private benefits, which increases the project's pledgeable income and enables the entrepreneur to borrow more. However, monitoring also entails a physical cost, so that a credible monitor needs to retain a sufficient fraction of loans as its skin in the game.



(a) Certification
(b) Intermediation

Figure 1: Flow of Funds Certification vs Intermediation

We depart from Holmstrom and Tirole (1997) by introducing a competitive financial market, in which the bank is allowed to trade its loans. Loans are rationally priced, and therefore the prices depend on the bank's incentives to monitor contemporaneously and in the future. If the bank has sold a large fraction of the loans, it will monitor less often and consequently, the price of loans will fall substantially. This price impact deters the bank from selling the loans too fast and too aggressively.

While Holmstrom and Tirole (1997) shows certification and intermediation are equivalent in a two-period model, a first main result of our model is that they lead to different equilibrium dynamics. Specifically, the lack of commitment hurts a certifying bank as in standard durable-goods monopoly models. A certifying bank has incentives to sell, because its marginal valuation is below that of investors. After the initial sell, the bank is willing to sell again since the price impact now only accrues to a smaller number of shares. On the other hand, the price of the loan drops following the expectation of the bank selling and the declining likelihood of monitoring, which in turn reduces the bank's proceeds from selling. Hence, the bank trades off the immediate trading gains versus the drop in its future payoff, including the drop in the loan's valuation as well as the decline in the expected payments it can collect upon the project matures.

By contrast, the lack of commitment does not hurt an intermediating bank, due to its ability to issue short-term deposits. Indeed, when the bank has access to deposit as a source of cheap financing, selling loans will not only depress the valuation of the loan but also increase the rate of deposits. The elevated deposit rate acts as a mechanism that deters the bank from selling. As a result, the bank finds it optimal to retain the entire loan on its balance sheet even though it does not have commitment to do so. This result is related to the literature on the commitment role of short-term debt. Compared to certification, a bank retains the loan in intermediation, which leads to a higher loan price and enables the bank to lend out more. As a result, the entrepreneur is able to borrow a higher amount in intermediation.

This distinction points out an important externality in bank monitoring and how intermediation manages to resolve it. As in Grossman and Hart (1980) and Diamond (1984), monitoring suffers from the free-rider problem because investors enjoy the benefits but do not share the cost with the bank. Therefore, the equilibrium monitoring effort is inefficiently low. In certification, this inefficiency leads to the bank reducing its probability of monitoring over time. By contrast, in intermediation, deposits helps the bank internalize the externality. Indeed, since these deposits are fairly priced, the cost of monitoring is internalized by the bank through the rate of deposits. Therefore, both the benefits and costs of monitoring are shared by the bank and its depositors. Essentially, deposits create a market for the monitoring services offered by the bank to be fairly priced.

Next, we turn to the entrepreneur's initial choice. We show that, if the entrepreneur net worth is high enough, she may end up choosing certification. On the one hand, the entrepreneur is able to borrow less under certification. On the other hand, the lower monitoring under certification (once the bank has sold the loans), allows the entrepreneur to extract higher private benefits. The second effect dominates if private benefits are large enough.

Finally, we explore the impact of common policies designed to address the bank commitment problem. In particular, we look at the impact that lock-up periods and minimum retention levels. While the qualitative nature of the equilibrium outcome remains unchanged, we show that introducing them allows the entrepreneur to borrow more upfront. We also show that, if an intermediating bank's deposits have long maturity or are subsidized by the government, the bank's commitment to retention is also impaired.

Literature Review

While our paper directly builds on Holmstrom and Tirole (1997), the focus of the two papers are different. Holmstrom and Tirole (1997) characterize the bank-lending channel, whereby the aggregate capital of the banking sector affects the aggregate lending and the interest rates. By contrast, our focus is on the intermediaries' ability to commit to its retentions. Holmstrom and Tirole (1997) establish two equivalence results: 1) certification leads to the same amount of lending as intermediation, and 2) maximizing borrowing capacity is equivalent to maximizing the entrepreneur's expected payoff subject to the lenders' participation constraints. We show both equivalence results cease to exist in the dynamic context. Rampini and Viswanathan (2019) also extend Holmstrom and Tirole (1997) to a dynamic context to study how the net worth of the corporate and the financial sector get accumulated following bad aggregate shocks.

Our analysis of certification relates to DeMarzo and Urošević (2006) on how a large shareholder trades off diversification and and monitoring incentives. More broadly, the problem follows the literature of the Coase conjecture on how a monopoly firm sells durable goods over time. Other related papers include DeMarzo and He (2016) and Fuchs and Skrzypacz (2010).

Our analysis of intermediation emphasizes the role of short-term debt as a commitment device, which relates to Diamond and Rajan (2001) and Calomiris and Kahn (1991). While both papers emphasize the demandable feature of debt and the externalities from creditor runs, our result does not rely on the first-come-first-serve constraint. Indeed, we model pari-passu debt, so that if the bank fails, all depositors receive an equal amount. Our mechanism depends on the endogenous deposit rate as the discipline device.

Bizer and DeMarzo (1992) study the problem when the borrower cannot commit to not borrowing from multiple lenders, which leads to higher borrowing cost and more defaults. Also closely related is Myers and Rajan (1998), who show how a firm's asset liquidity reduces the ability to commit to an investment strategy that protects its investors. While both papers focus on the commitment problem of the borrower, we focus on a specific type of lender, i.e., the monitoring financial intermediary.

2 The Model

Our model follows closely the fixed-size investment setup in Holmstrom and Tirole (1997). A key distinction is that, there is a competitive financial market which allows the bank to trade loans. In other words, the bank can not commit to its loan retention and thus the decisions to monitor in the future. The more it sells, the less likely it will monitor, and the loans will be valued lower as well. We will explore how the bank's equilibrium trading behavior interact with its monitoring decisions, in the dynamic context of certification and intermediation.

2.1 Agents and Technology

Time is continuous and goes to infinity: $t \in [0, \infty)$. There are three groups of agents: one entrepreneur (she) – the borrower, competitive intermediaries – banks, and investors. All agents are risk neutral and have limited liabilities. The entrepreneur starts out with cash level A, whereas both banks and investors have deep pockets. We assume that investors do not discount future cash flows, whereas the entrepreneur and intermediaries discount the future at a rate $\rho > 0$.

At time 0, the entrepreneur has access to a project which requires a fixed investment size $I > A$. Thus, she needs to borrow at least $I - A$. The project matures at a random time τ_ϕ, which arrives upon a Poisson event with intensity $\phi > 0$. Let us define $\Phi = \frac{\phi}{\rho+\phi}$ as the effective time discount applied by the entrepreneur and banks to the project's final cash flows. At τ_ϕ, the project generates the final cash flows R in the case of success and 0 in the case of failure. The probability of success is p_H if the entrepreneur works at τ_ϕ and $p_L = p_H - \Delta$ if she shirks. Two options of shirking are available: the high option brings private benefit B, which exceeds b, the private benefit associated with the low option.

2.2 Monitoring, Financial Structures and Contracts

A competitive set of banks are available at $t = 0$, and the entrepreneur signs a contract with one of them. One can think about this bank as the relationship lender. Note that we do not allow for multiple relationship banks to avoid duplication of monitoring efforts and the free-ride problem (Diamond, 1984). At τ_ϕ when the project matures, the bank can monitor to eliminate the high shirking option. To do so, it needs to pay a private monitoring cost $\tilde{\kappa} > 0$, where $\tilde{\kappa} \in [0, \bar{\kappa}]$ has a distribution with $F(\cdot)$ and $f(\cdot)$ being the cumulative distribution function (CDF) and probability density function (PDF), respectively. Note that we differ from Holmstrom and Tirole (1997) by allowing the cost $\tilde{\kappa}$ to vary stochastically. One can interpret this as variations in legal and enforcement costs or simply the fluctuations in costs of hiring loan officers. The stochastic cost assumption makes the bank's equilibrium monitoring decisions smooth and is a continuous function of its loan retention.[1] Figure 2 describes the timing. Note that for simplicity, we assume that both the entrepreneur's effort and the bank's monitoring are needed only at τ_ϕ when the project matures. Introducing long-term effort and monitoring will complicate the model without bringing many new insights.

We study two types of financial structures: certification and intermediation. In certification, the bank puts its own funds in the entrepreneur's venture to ensure that it will monitor, which then attracts investors to *directly* invest into the venture as well. One can think of this type of banks as either venture capitalists or lead investment banks in loan syndications. In certification, the entrepreneur directly signs the initial contract with the bank and investors. Under limited liability, no agent receives anything if the project fails. If the project succeeds, let R_f, R_m, R_u be the scheduled payments to the entrepreneur, the bank, and investors, which sum up to R. We use $R_o = R_m + R_u$ to denote the total claims held by outside creditors, i.e., the bank and investors. For the remainder of this paper, we will also refer to R_o as *loans*.

[1] Holmstrom and Tirole (1997) also introduce an extension with a continuous distribution of private benefits to capture continuous monitoring intensity.

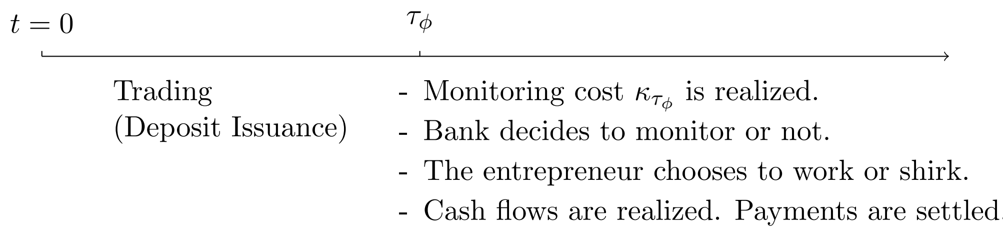


Figure 2: Timing

In intermediation, investors do not directly invest in the entrepreneur's project. Instead, they make deposits in the bank, which in turn lends to the entrepreneur a collection of its own funds and these deposits. This type of banks resemble more of commercial banks. In intermediation, the entrepreneur signs an initial contract with a bank and promises to repay $R_o = R - R_f$ if the project succeeds. The bank, in turn, offers deposit contracts $\{D_t, y_t\}$ to investors over time, where D_t is the amount of deposits and y_t the associated interest payment (henceforward referred to as the deposit rate). We model the deposit contract as debt with *instant* maturity. Whenever the bank fails to honor its deposit payments, the remaining loans are sold to repay the depositors. For simplicity, we assume there is no bankruptcy cost (or loss of charter value) if the bank fails, and a positive bankruptcy cost will only change the results quantitatively.

Given that the entrepreneur has has (weakly) the highest cost of capital among all the agents, she should retain as little stake as possible. Therefore, in both certification and intermediation, it is optimal to let the entrepreneur retain $R_f = b/\Delta$, which guarantees that she will work if the bank monitors. Therefore, $R_o = R - b/\Delta$.

2.3 Trading and Pricing in the Financial Market

A competitive financial market opens after time 0, in which loans can be traded.[2] We normalize the total share of loans outstanding to one and use θ_t to denote the bank's retention at time t. Before trading starts, the bank's initial retention is $\theta_{0-} = R_m/(R_m + R_u)$ in certification, and $\theta_{0-} = 1$ in intermediation. A trading strategy $(\theta_t)_{t\geq 0}$ specifies the bank holdings at each time. We consider trading strategies that admit both smooth and atomic trading, as well as mixed strategies over the time of atomic trades. We restrict the trading strategy to be Markovian, i.e., $\{\theta_t\}_{t\geq 0}$ is a

[2] We assume that the entrepreneur's retention R_f is not tradable, or equivalently the entrepreneur can commit to hold onto R_f on the balance sheet.

Markov process.[3] We will sometimes refer to θ_t as the bank's skin in the game, which is publicly observable. In our model, θ_t will be the payoff-relevant state variable.

The price of loans depends on whether the entrepreneur works or shirks, which in turn depends on the probability of bank monitoring. Conditional on the project maturing, let $p(\theta)$ be the equilibrium probability that the bank with retention θ monitors, investors of the loan receive

$$d(\theta) = p(\theta) R_o. \tag{1}$$

per share. Let $q(\theta)$ be the price of the loan per share when $\theta_t = \theta$. In a competitive financial market, the price is given by the expected present value of the asset:

$$q(\theta) = \mathbb{E}\left[d\left(\theta_{\tau_\phi}\right) \middle| \theta_t = \theta\right], \tag{2}$$

where the expectation operator is taken with respect to the equilibrium path of $\{\theta_s\}_{t \le s \le \tau_\phi}$.

The probability of monitoring $p(\theta)$ will differ in certification and intermediation. Let κ be the realization of the stochastic monitoring cost $\tilde{\kappa}$. In certification, the bank with retention θ chooses to monitor if and only if

$$\kappa \le \kappa_c \coloneqq \Delta R_o \theta, \tag{3}$$

where $\Delta = p_H - p_L$. In intermediation, a bank with retention θ and deposits D monitors if and only if

$$\kappa \le \kappa_i \coloneqq \Delta\left(R_o \theta - D\right). \tag{4}$$

A comparison between (3) and (4) illustrates the static tradeoff between certification and intermediation. For the same retention θ, a higher D reduces an intermediating bank's incentive to monitor: this is the standard debt-overhang effect.

Let $G(\theta)$ be the bank's instant trading gains. In the case of continuous trading, $dG(\theta) = -q(\theta)\dot{\theta}dt$. In the case of atomic trading, the bank's holding jumps to θ^+ and the associated trading gain is $dG(\theta) = q(\theta^+)(\theta - \theta^+)$. Note that trading is settled at price $q(\theta^+)$ to reflect the price impacts and the forward-looking nature of asset prices.

[3] In the context of a durable goods monopoly, Ausubel and Deneckere (1989) show that in the absence of a gap (in durable goods models there is a gap if the lowest consumer valuation is above the marginal cost) there are non-Markov equilibria in which the seller can achieve payoffs close to the commitment solution. The logic behind the construction is similar to the one in the folk theorem for repeated games. In our model, when $\theta = 0$, the marginal valuation of investors coincides with the one of the bank only if $p_L = 0$, in which case, there is no gap.

2.4 Equilibrium Definition

2.4.1 Certification

If the project matures at time t, the bank's expected payoff is

$$\pi_c(\theta) = p_c(\theta) R_o \theta - \int_0^{\kappa_c} \kappa dF(\kappa), \tag{5}$$

where the project succeeds with probability

$$p_c(\theta) := p_L + F(\kappa_c)\Delta, \tag{6}$$

upon which the bank receives $R_o\theta$. The bank maximizes the sum of its payoff upon the project matures $e^{-\rho(\tau_\phi - t)}\pi_c\left(\theta_{\tau_\phi}\right)$ and the cumulative trading gains $\int_0^{\tau_\phi} e^{-\rho(s-t)} dG(\theta_s)$. Since τ_ϕ follows the exponential distribution, the maximization problem is equivalent to

$$\max_{\{\theta_t\}_{t\geq 0}} \mathbb{E}\left[\int_0^\infty e^{-(\rho+\phi)t}\Big(\phi\pi_c(\theta_t)\,dt + dG(\theta_t)\Big)\right], \tag{7}$$

where the expectation operator allows for the mixed strategies in $\{\theta_t\}$.

We consider Markov perfect equilibrium in which the state variable is the bank's retention θ, henceforward referred to as the Certification Equilibrium.

Definition 1. *A* ***Certification Equilibrium*** *is a Markov Perfect Equilibrium consisting of a price function $q\colon [0,1] \to \mathbb{R}_+$ and a trading strategy $(\theta_t)_{t\geq 0}$ that satisfy*

1. *For all $\theta_0 \in [0,1]$, $(\theta_t)_{t\geq 0}$ is a Markov trading strategy with initial value θ_0 that maximizes (7).*
2. *For all $\theta \in [0,1]$, the price $q(\theta)$ satisfies the break-even condition (2).*

2.4.2 Intermediation

If the project matures at time t, the bank's expected payoff is

$$\hat{\pi}_i(\theta, D) = \hat{p}(\theta, D)(\theta R_o - D) - \int_0^{\kappa_i} \kappa dF(\kappa), \tag{8}$$

where the project succeeds with probability

$$\hat{p}_i(\theta, D) := p_L + F(\kappa_i)\Delta, \tag{9}$$

upon which the bank's equity holder receives $\theta R_o - D$ after paying off its depositors. Besides trading gains, an intermediating bank also receives income from deposit issuance. In particular, let D_0 be the value of deposits issued at $t = 0$. The bank's net income from deposit issuance at time t is $dD_t - y_t D_t dt$, where

$$y_t = \hat{y}(\theta, D) = \phi(1 - \hat{p}(\theta, D)) \tag{10}$$

compensates the default risk borne by depositors.[4] In intermediation, the bank trades loans and issues deposits to maximize the the expected payoff upon the project maturing, together with the net income from deposit issuance and trading gains, i.e.,

$$\max_{\{\theta_t,\, D_t\}} \mathbb{E}\left[\int_0^\infty e^{-(\rho+\phi)t}\Big(\phi\hat{\pi}_i(\theta_t, D_t)\,dt + dD_t - \hat{y}(\theta_t, D_t)\,D_t dt + dG(\theta_t)\Big)\right]. \tag{11}$$

The choice of D_t in (11) is restricted by the bank's limited liability, which imposes a borrowing constraint. Lemma 1 shows that the choice of D_t is essentially a static decision. Therefore, we can use θ_t as the state variable and suppress the problem's dependence on D_t.

Lemma 1. *The maximization problem* (11) *is equivalent to solving*

$$\phi\pi_i(\theta) := \max_{D \le \Pi_i(\theta)} \left\{\phi\left[\hat{p}_i(\theta, D)\theta R_o - \int_0^{\kappa_i} \kappa dF(\kappa)\right] + \rho D\right\}, \tag{12}$$

and

$$\Pi_i(\theta_0) = E(\theta_0, D_0) + D_0 = \max_{(\theta_t)_{t\ge 0}} \int_0^\infty \mathbb{E}\left[e^{-(\rho+\phi)t}\Big(\phi\pi_i(\theta_t)dt + dG(\theta_t)\Big)\right]. \tag{13}$$

The bank chooses the trading rate to maximize firm value subject to the endogenous borrowing constrain $D_t \le \Pi_i(\theta_t)$. From (12), it is clear that the choice of deposit D only involves static tradeoffs. A higher D reduces the probability of monitoring $\hat{p}(\theta, D)$ (see Equation (4)), thereby reducing the value of loans. On the other hand, the term ρD makes it clear that deposit is a cheaper cost of funding. Given that all the functions on the right-hand side of (12) are continuous, the optimal deposit issuance is also continuous, i.e, does not admit jumps. Deposit issuance is bounded by the constraint $D \le \Pi_i(\theta)$, which arises from the bank's limited liability. Here, $\Pi_i(\theta)$ is the bank's value function given its retention θ, which implicitly assumes that deposit issuance has been

[4] Implicitly, the bank will honor its debt payments whenever $t < \tau_\phi$, that is, if the project does not mature. In the case of strategic default, depositors acquire the remaining loans and immediately resell them, which necessarily leads to zero payoff to the bank. Thus, strategic default never happens.

chosen at the optimal level. Therefore, this constraint involves a fixed-point for the value function $\Pi_i(\theta)$. In (13), the left-hand side $\Pi_i(\theta_0)$ includes $E(\theta_0, D_0)$, the value to the bank's equity holders and D_0, that to the depositors. For the remainder of this paper, we will sometimes refer to Π_i, E, and D as the bank value, the equity value, and the deposit value, respectively. One implication of Lemma 2 is that, even though the bank's equity holders decide its trading strategy, maximizing the bank's equity value $E(\theta_0, D_0)$ is equivalent to maximizing the total bank value $\Pi_i(\theta_0)$, because deposits D_0 are fairly priced at time 0.

We look for a Markov perfect equilibrium in state variable θ_t, henceforth referred to as the Intermediation Equilibrium.

Definition 2. *An* ***Intermediation Equilibrium*** *is a Markov Perfect Equilibrium consisting on a price function $q\colon [0,1] \to \mathbb{R}_+$, a trading strategy $(\theta_t)_{t\geq 0}$, a deposit issuance policy $D^*\colon [0,1] \to \mathbf{R}_+$, and the deposit rate function $y\colon [0,1] \to \mathbf{R}_+$ that satisfy*

1. *For all $\theta \in [0,1]$, the deposit issuance policy $D^*(\theta)$ solves (12).*
2. *For all $\theta_0 \in [0,1]$, $(\theta_t)_{t\geq 0}$ is Markov trading strategy with initial value θ_0 that maximizes (13).*
3. *For all $\theta \in [0,1]$, the price $q(\theta)$ satisfies the break-even condition (2).*
4. *For all $\theta \in [0,1]$, the deposit rate $y(\theta) := \hat{y}(\theta, D^*(\theta))$ satisfies (10).*

2.5 Parametric Assumptions

To make the problem non-trivial, we impose the following parametric assumptions. The first assumption says the project's expected payoff is always higher if the entrepreneur works.

Assumption 1.

$$p_H R > p_L R + B.$$

Assumption 2 ensures that the problem is not driven by some particular feature of the monitoring cost's distribution function. It is satisfied by most commonly-used distribution functions such as the uniform distribution.

Assumption 2.

$$\frac{f'}{f^2} < \frac{\Phi}{1-\Phi}\Delta, \qquad \forall \kappa \in [0, \bar{\kappa}].$$

Finally, we restrict the (expected) monitoring cost to be sufficiently low.

Assumption 3.

$$\Phi F\left(\Delta R-b\right)\left(\Delta R-b\right)-\int_{0}^{\Delta R-b}\kappa dF\left(\kappa\right)\geq\left(1-\Phi\right)p_{L}\left(\Delta R-b\right).$$

This assumption leads to the following result. If the entrepreneur has the minimum retention $R_f = \frac{b}{\Delta}$ to guarantee she will work if the bank monitors, and the bank always retains the entire loan (i.e., $\theta_t \equiv 1,\ \forall t \leq \tau_\phi$), the bank's payoff exceeds that if it immediately sells the entire loan and never monitors. If this assumption is violated, bank monitoring is never needed in equilibrium.

2.6 Two Benchmarks

We present two benchmark cases in this subsection.

2.6.1 Static Benchmark: Equilibrium without the Financial Market

Let us first solve the static model without the financial market. This is also the static benchmark presented in Holmstrom and Tirole (1997). The goal of this subsection is to establish two equivalence results. First, the maximum amount that the entrepreneur is able to borrow is equivalent under certification and intermediation. Second, maximizing the initial borrowing amount is equivalent to maximizing the entrepreneur's expected payoff.

Under either certification and intermediation, the entrepreneur retains $R_f = b/\Delta$ to maintain incentives to work. With slight abuse of notation, let Φ be the bank's one-period discount rate in this subsection. Following Holmstrom and Tirole (1997), the optimization problem in certification is

$$\begin{gathered}L_c = \max_{\left\{R_f, R_m, R_u\right\}} p_c\left(\Phi R_m + R_u\right) - \Phi\int_0^{R_m\Delta}\kappa dF\left(\kappa\right)\\ s.t.\qquad R_m + R_u = R - \frac{b}{\Delta}.\end{gathered}$$

The optimization problem in intermediation includes two steps. First, the bank chooses R_u,[5] taken $R_o = R_m + R_u$ as given. In the second step, the entrepreneur chooses R_o to maximize L_i:

$$L_i = \max_{R_o}\left\{\max_{R_u} p_i\left(\Phi R_m + R_u\right) - \Phi\int_0^{R_m\Delta}\kappa dF\left(\kappa\right)\right\}.$$

[5] Note that the deposit choice in the static model is $D = R_u$.

Proposition 1 shows that as in Holmstrom and Tirole (1997), certification and intermediation are equivalent in the one-period model.

Proposition 1 (Static Benchmark)**.** *In the static model, $L_c = L_i$ so that the equilibrium outcomes in certification and intermediation are equivalent. Moreover, the optimal contract that leads to maximum borrowing also maximizes the borrower's expected payoff.*

Intuitively, since lenders always break even, and the project's expected payoff on the equilibrium path is always $p_H R$, maximizing borrowing amount is equivalent to maximizing the borrower's expected payoff. We will show in subsection 3.3, this result is no longer true in the dynamic setup.

2.6.2 The Constrained Social Planner's Problem

The social planner chooses the bank's retention $\{\theta_t\}_{t\geq 0}$ to maximize the aggregate social welfare, subject to the constraint that the bank chooses whether to monitor given its retention. Under the full commitment to $\{\theta_t\}_{t\geq 0}$, the choices of deposit D_t are redundant.

Proposition 2 (Constrained Social Planner's Problem)**.** *A social planner subject to constraint* (3) *chooses $\{\theta_t\}_{t\geq 0}$ to maximize the social welfare*

$$\begin{aligned} W = \max_{(\theta_t)_{t\geq 0}} \quad & \int_0^\infty \phi e^{-\phi t} \Bigg\{ \left(1-e^{-\rho t}\right)(1-\theta_t) d\left(\theta_t\right) \\ & \qquad + e^{-\rho t}\left[p(\theta_t)R + \left(1-F(\kappa_c)\right)B - \int_0^{\kappa_c} \kappa dF\left(\kappa\right)\right]\Bigg\} dt. \qquad (14) \end{aligned}$$

The optimal retention always satisfies $\theta_t < 1$.

Due to differences in time discounting, the flow payoff in (14) is a weighted sum of the payoff to investors and that to the bank and the entrepreneur. The project succeeds with probability $p(\theta_t)R$, and the entrepreneur shirks to receive the high private value B with probability $(1 - F(\kappa_c))$. Note that in this case, the optimal retention θ_t is essentially a static choice that balances the benefits and the costs of bank monitoring. Moreover, the difference in time discounting makes this tradeoff time-varying, which implies the optimal retention is in general time-varying as well.

Given that all parties are risk neutral and transfers can be made at the initial date, the constrained social planner's solution is identical to one in which the entrepreneur chooses $\{\theta_t\}_{t\geq 0}$ to maximize her payoff, subject to the lenders' participation constraints as well as the constraint that the bank chooses to monitor or not.

3 Dynamic Model: Equilibrium with the Financial Market

In this section, we solve the model. Subsection 3.1 and 3.2 respectively derive the certification and intermediation equilibrium. Subsection 3.3 presents a special case in which the monitoring cost $\tilde{\kappa}$ follows the uniform distribution and $p_L = 0$, where we obtain closed-form solutions in primitives. The entrepreneur's initial choices therefore follow by comparing the two equilibria. We will also derive the the solution to the social planner's problem with the constraint that the bank needs sufficient retention to monitor.

3.1 Certification Equilibrium

In general, the bank can trade loans smoothly or atomically. We will show that both types of trading can occur in equilibrium. Let $\Pi_c(\theta)$ be the bank's value function with retention θ.[6] In the smooth-trading region, $\Pi_c(\theta)$ satisfies the following Hamilton-Jacobi-Bellman (HJB) equation

$$\rho \Pi_c(\theta) = \max_{\dot{\theta}} \ \phi \Big[\pi_c(\theta) - \Pi_c(\theta) \Big] + \dot{\theta} \Big[\Pi_c'(\theta) - q_c(\theta) \Big]. \tag{15}$$

While the left-hand-side stands for the bank's required return, the first term on the right-hand-side represents the event of the project maturing, in which case the bank receives $\pi_c(\theta)$ defined in (5). The second term captures the overall benefit of trading, which includes the change to the bank's continuation value as well as the trading gain. A necessary condition for smooth trading is

$$\Pi_c'(\theta) = q_c(\theta), \tag{16}$$

so that the bank is indifferent between trading or not. In this case, the per-share trading gain $q_c(\theta)$ is offset by the drop in the bank's continuation value $\Pi'(\theta)$. Substituting the indifference condition (16) into (15), we immediately get that in the region of smooth trading,

$$\Pi_c(\theta) = \Phi \pi_c(\theta). \tag{17}$$

Note that the bank value is equal to the payoff it receives conditional on the project maturing, times the bank's effective discount rate $\Phi = \frac{\phi}{\rho + \phi}$. Surprisingly, the bank does not benefit from its ability to trade these loans in the financial market at all; its payoff is identical to the one that it retains θ until the project finally matures at τ_ϕ. This result on how the lack of commitment fully offsets the trading gains is standard in the literature on the Coase conjecture (Fuchs and Skrzypacz,

[6] A certifying bank does not issue deposits, so there is no distinction between the bank value and the bank equity holder's value.

2010), which has also been applied in corporate finance to study trading by a large shareholder (DeMarzo and Urošević, 2006) and leverage dynamics (DeMarzo and He, 2016).

Even though the bank's equilibrium payoff is identical to the one if it does not trade at all, it doesn't imply the bank will not trade loans on the equilibrium path. In fact, the price of loans following a no-trade strategy will be too high, which gives the bank strict incentives to sell. Indeed, the bank has incentives to sell for two reasons. First, due to the higher cost of capital, the bank's marginal valuation is below that of investors. Second, after the initial sell and the reduction of retention, the bank is willing to sell again because the price impact only accrues to a smaller number of shares. Now, let us characterize the bank's equilibrium trading strategy, which comes from the determination of the equilibrium loan prices. Since investors do not discount future cash flows, $q_c(\theta)$ must satisfy the following asset-pricing equation whenever the bank trades smoothly

$$0 = \phi\left[d_c(\theta) - q_c(\theta)\right] + \dot{\theta} q_c'(\theta), \tag{18}$$

where $\phi\left[d_c(\theta) - q_c(\theta)\right]$ resembles the dividend income and $\dot{\theta} q_c(\theta)$ the capital gain.[7] Combining (16), (17) and (18), and using the relation $d_c(\theta) = \pi_c'(\theta)$, one can derive the following dynamic trading strategies in equilibrium:

$$\dot{\theta} = -\phi \frac{(1-\Phi)\,\pi_c'(\theta)}{\Phi \pi_c''(\theta)} < 0. \tag{19}$$

Clearly, in the smooth-trading region, the bank sells loans over time and its retentions declines continuously, even though it is indifferent between selling the loan or not. Intuitively, the equilibrium loan price is forward-looking and therefore takes into account the bank's decisions of future monitoring. In order to satisfy the bank's indifference condition, the equilibrium price of the loan cannot be too high, implying that the probability of monitoring must decrease over time. Therefore, the only trading strategy consistent with this price requires the bank to sell its loans over time.

So far, we have only focused on the case of smooth trading. Meanwhile, the bank also has the option to sell an atom of loans. In general, the bank can sell either a fraction or all the remaining loans. Lemma 3 in the appendix proves the bank will never sell a fraction. This result follows the intuition in standard Coasian dynamics models. Atomic trading arises whenever the bank has strict incentives to sell. If so, it prefers to sell as much as possible and as early as possible. Therefore, it is always optimal for the bank to sell the remaining loans all at once. Given this result, we are left to check when the bank decides to sell off all its retention at a price $q_c(0)$, where $q_c(0) = p_L R_o$ is

[7]The terms without subscripts c have been defined in equations (1) and (2).

the per-share loan price without monitoring. Indeed, we show there exists a unique θ_* such that $\theta_* q_c(0) = \Phi\pi_c(\theta_*)$, and $\theta_* q_c(0) < \Phi\pi_c(\theta_*)$ if and only if $\theta > \theta_*$. In other words, there exists a unique cutoff θ_* below which the bank finds it optimal to sell of all the remaining loan.

The final step in the equilibrium construction is to derive the trading strategy at $\theta = \theta_*$. First, note that the bank cannot hold onto the loans forever, as the resulting loan price will be too high to induce the bank to sell. It can't be that the bank sells smoothly either; shortly afterwards the bank will have strict incentive to sell. Suppose that the bank sells off immediately after θ_t reaches θ_*. The price of the loan will then experience a deterministic downward jump. Mathematically, let $T_* := \inf\{t > 0 : \theta_t = \theta_*\}$. While q_{T_*-} satisfies (16), $q_{T_*} = q(0) = p_L R_o$. This deterministic downward jump is inconsistent with the asset pricing equation (2). Therefore, the equilibrium necessarily involves some delay before the entire selloff. The only (stationary) trading strategy at θ_* consistent with (2) is for the bank to adopt a mixed strategy:[8] the bank sells off all its remaining loans at some Poisson rate λ that satisfies

$$q_c(\theta_*) = \mathbb{E}\left[d_c(\theta_{\tau_\phi})|\theta_t = \theta_*\right] = \frac{\lambda}{\phi+\lambda} d_c(0) + \frac{\phi}{\phi+\lambda} d_c(\theta_*)$$

Simple derivation shows λ is determined by

$$p_L + \frac{\phi}{\lambda+\phi} F\left(\Delta R_o \theta_*\right)\Delta = \Phi\left[p_L + F\left(\Delta R_o \theta_*\right)\Delta\right]. \tag{20}$$

Proposition 3 summarizes the previous discussion and describes the equilibrium outcome. The formal proof requires to verify that the bank's trading strategy is optimal, which is supplemented in the appendix using results from the theory of optimal control in stratified domains.[9]

Proposition 3 (Certification Equilibrium)**.** *There is a unique **Certification Equilibrium**. Given the bank's initial retention θ_0, the bank sells its loans smoothly at a rate given by equation (19) until T_*, after which it sells off its remaining loans at some Poisson rate λ that satisfies (20). The equilibrium loan price is*

$$q_c(\theta_t) = \begin{cases} \Phi\Big(p_L + F\left(\Delta R_o \theta_t\right)\Delta\Big) R_o & t < T_* \\ \Big(p_L + \frac{\phi}{\lambda+\phi} F\left(\Delta R_o \theta_*\right)\Delta\Big) R_o & T_* \le t < \tau_\lambda \\ p_L R_o & t \ge \tau_\lambda \end{cases} \tag{21}$$

[8] The delay can also be deterministic, but the equilibrium is no longer within the class of Markov Perfect Equilibrium. Time since the bank's retention reaches θ_* will also be a payoff-relevant state variable. The price q_t will also be non-stationary and depends on this time.

[9] Due to the discontinuity in the price function $q_c(\theta)$, the HJB equation (15) is discontinuous at θ_*. This technical problem can be sidestepped using (discontinuous) viscosity solution methods.

The contract that maximizes the initial borrowing amount has $\theta_0 = 1$, with the borrowing amount

$$L_c = \Phi \pi_c (1) . \tag{22}$$

A contract that has the entrepreneur borrow exclusively from the bank at $t = 0$ enables the most upfront borrowing, even though bank capital is costly. Intuitively, under this contract, it takes the longest period for the bank to fully offload its loans. This result is in contrast to that in Holmstrom and Tirole (1997), where the entrepreneur prefers to use as less bank capital as possible. This difference arises because in Holmstrom and Tirole (1997), bank capital is expensive, whereas in the dynamic setup without commitment, there is a second channel whereby higher bank retention slows down the bank's selling process. According to the literature on Coasian dynamics, the bank does not benefit from its ability to sell its loans to investors. Therefore, the longer it takes the bank to sell off the entire loan, the more likely the bank will monitor, and the value of the loan is also higher. As a result, the second channel necessarily dominates.

Let us end this subsection by describing the borrower's expected payoff V_c. Specifically, given that the entrepreneur retains R_f, her payoff is

$$V_c = \mathbb{E}\left[\int_0^\infty \phi e^{-(\rho+\phi)t} \left\{ \mathbb{1}_{\{\kappa \le \kappa_c\}} p_H R_f + \mathbb{1}_{\{\kappa \le \kappa_c\}} (p_L R_f + B) \right\} dt \right], \tag{23}$$

where κ is the realization of $\tilde{\kappa}$ at time t. The expectation operator is taken with respect to the bank's equilibrium trading strategy which involves mixed strategies. Intuitively, if the realized monitoring cost is lower than the threshold κ_c defined in (3), the bank monitors, and the entrepreneur receives $p_H R_f$ in expectation. Otherwise, the bank chooses not to monitor, and the entrepreneur receives the expected return $p_L R_f$ together with the private benefits B.

3.2 Intermediation Equilibrium

The analysis of the intermediation equilibrium has two steps: deposit issuance and loan trading.

3.2.1 Deposit Issuance

Plugging (10) and (8) into (12) and dividing both sides by $\rho + \phi$, we can rewrite the bank's objective function without trading gains as follows:

$$\mathcal{V}(D, \theta) := \Phi \left[\hat{p}_i (\theta, D) \theta R_o - \int_0^{\kappa_i} \kappa dF(\kappa) \right] + (1 - \Phi) D. \tag{24}$$

This objective function includes two terms. The first term is the net payoff to the bank and its depositors: with probability $\hat{p}(\theta, D)$, the project succeeds so that they receive θR_o. $\int_0^{\kappa_i} \kappa dF(\kappa)$ is the expected monitoring cost. The second term in (24) is the value from issuing deposits. An increase in D reduces the bank's monitoring incentive and therefore reduces the first term. Meanwhile, an increase in D also reduces the bank's funding cost and therefore increases the second term. The optimal D is chosen to balance the two effects, which after some simple derivation, solves

$$\phi\left[f(\kappa_i)\Delta^2\right]\tilde{D} = \rho \tag{25}$$

under Assumption 2. For any $\theta \in (0,1)$, let us denote the solution by $\tilde{D}(\theta)$. Meanwhile, note that the bank's equity holders' limited liability constraint requires that $E(\theta, D) = \Pi_i(\theta) - D \geq 0$, so that for any θ, $D \leq \Pi_i(\theta)$. This constraint implicitly defines a maximum value of deposit that the bank can issue for any choice of θ, which we denote as $D^{\max}(\theta)$.[10] Since $\Pi_i(\theta)$, the bank's maximal value given its retention θ has implicitly assumed that deposit issuance is optimally chosen, solving for $D^{\max}(\theta)$ is therefore equivalent to looking for a fixed-point in the constraint. To finish the derivation of optimal deposit choice $D^*(\theta)$, we need one more assumption.

Assumption 4. *The distribution function $f(\tilde{\kappa})$ satisfies*

$$\Delta f'(\kappa)D - f(\kappa) < 0, \qquad \forall D \leq D^{\max}(\theta),\ \forall\theta \in [0,1],\ \forall\kappa \in [0,\bar{\kappa}].$$

Assumption 4 guarantees that for any given θ, the second-order partial derivative of (24) satisfies $\frac{\partial^2 \mathcal{V}}{\partial D^2} < 0$. This condition is always satisfied if f is non-increasing. Under this assumption, we get the following result

Lemma 2. *There exists a unique $\theta_\dagger \in [0,1]$ such that*

$$D^*(\theta) = \begin{cases} \tilde{D}(\theta) & \text{if } \theta \geq \theta_\dagger \\ D^{\max}(\theta) & \text{if } \theta < \theta_\dagger. \end{cases} \tag{26}$$

3.2.2 Trading

Next, we turn to the maximization problem (13) and study how an intermediating bank trades its loans over time. Following the similar logic in the certification equilibrium, the term $\dot{\theta}\left(\Pi_i'(\theta) - q_i(\theta)\right)$ must vanish in the smooth-trading region, which yields to following HJB equa-

[10] Note that in the case that the bank decides to sell off all its remaining loans, $D^{\max}(\theta)$ is defined as 0.

tion:

$$\rho\Pi_i(\theta) = \phi\Big[\pi_i(\theta) - \Pi_i(\theta)\Big] \tag{27}$$

$$\Pi_i'(\theta) = q_i(\theta). \tag{28}$$

Moreover, given an equilibrium price $q_i(\theta)$, the asset-pricing equation

$$0 = \phi\Big[d_i(\theta) - q_i(\theta)\Big] + \dot{\theta}q_i'(\theta), \tag{29}$$

implies the following trading strategy

$$\dot{\theta} = -\phi\frac{d_i(\theta) - q_i(\theta)}{q_i'(\theta)}.$$

where functions $d_i(\theta)$ and $\pi_i(\theta)$ are defined under the optimal deposit issuance $D^*(\theta)$. We can apply the Envelope Theorem in Milgrom and Segal (2002) to get

$$\pi_i'(\theta) = d_i(\theta) + \left[f(\kappa_i)\frac{\partial\kappa_i}{\partial\theta}\Delta\right]\tilde{D} + z(\theta)D^{\max'}(\theta) = d_i(\theta) + \frac{1-\Phi}{\Phi}R_o + z(\theta)D^{\max'}(\theta), \tag{30}$$

where $z(\theta) \geq 0$ is the Lagrange multiplier of the constrain $D \leq D^{\max}(\theta)$. Given so, the equilibrium trading rate is[11]

$$\dot{\theta} = \phi\frac{(1-\Phi)(1-p(\theta))R_o + \Phi z(\theta)D^{\max'}(\theta)}{\Phi\pi_i''(\theta)} > 0. \tag{31}$$

A comparison between (31) and (19) highlights a crucial difference between the two implementation structures. In certification, the bank has incentives to sell because investors have a higher valuation of these loans. In intermediation, this incentive disappears. In fact, the bank has incentives to increase its retentions. This distinction arises because an intermediating bank issues deposits to fund its loans. To see this, let us compare $\pi_c'(\theta) = d_c(\theta)$ with (30), which clearly shows an increase in θ leads to two additional benefits in intermediation. The first benefit is characterized by the term $\left[f(\kappa_i)\frac{\partial\kappa_i}{\partial\theta}\Delta\right]\tilde{D}$. Intuitively, $f(\kappa_i)\frac{\partial\kappa_i}{\partial\theta}$ is the marginal effect of retention θ on the equilibrium probability of bank monitoring, so that $f(\kappa_i)\frac{\partial\kappa_i}{\partial\theta}\Delta$ is the equilibrium probability that deposits will be repaid. Given that deposits are fairly valued, more retention enables the bank to issue *cheaper* deposits. The second benefit is captured by the last term $z(\theta)D^{\max'}(\theta)$, which is only positive

[11] Assumption 2 guarantees $\pi_i''(\theta) > 0$.

if $D^*(\theta) = D^{\max}(\theta)$. Intuitively, an increase in θ also relaxes the constraint on deposit issuance derived from the bank equity holders' limited liability constraint. Therefore, more retention has an add-on benefit by further allowing the bank to issue *more* deposits. Note this second benefit disappears whenever the deposit issuance constraint is slack, i.e., $z(\theta) = 0$. Even though the bank has a higher cost of capital and deposits reduce the bank's incentive to monitor, the availability of cheap deposits as source of financing offers the bank sufficient incentives to increase its retention. By definition, $\theta_0 = 1$ in intermediation. In equilibrium, the bank never sells its loans and $\theta_t \equiv 1$ for all $t \geq 0$. In other words, an intermediating bank retains the entire loan until the project matures.

The idea that short-term deposit (or in general short-term debt) could help with the commitment problem has been introduced in Calomiris and Kahn (1991), Diamond and Rajan (2001), and Diamond (2004). While the mechanism of these papers rely on the demandable feature of deposits and run externalities, the channel in our paper is different. In particular, the bank retention choice directly affects the cost (and also the amount of) of deposits. The rates of deposits enable the bank to internalize the externalities such that monitoring is "awarded" in the form of cheaper deposits. Another difference is the nature of the commitment problem. While these papers study the bank's commitment to enforce claims and collect payments, our focus is on how the bank commits to retain its loans over time.

Finally, to complete the characterization of the equilibrium, we consider the case where the bank trade loans atomically. Following the same logic as in the certification equilibrium, we can show that there exists a unique θ_* such that the bank sells off all the remaining loans at price $q(0) = p_L R_o$ if $\theta < \theta_*$. However, when $\theta = \theta_*$, unlike in the certification equilibrium, the mixed strategy is no longer needed. Instead, the bank chooses to buy the loan smoothly so that θ will increase above θ_*. Therefore, the price of the loan satisfies $q_i(\theta_*) = \Phi\pi'(\theta_*)$.[12] The following proposition summarizes the results.

Proposition 4. *There is a unique* ***Intermediation Equilibrium*** *in which the bank holds $\theta_t = 1$ until the project matures, and keeps a constant debt level $D_t = D^*(1)$. The equilibrium loan price is $q_i(1) = \Phi\pi_i'(1)$.*

Similar to certification, let us use V_i to denote the entrepreneur's expected payoff in intermedi-

[12] Both in the case of certification and intermediation, the price function $q(\theta)$ is discontinuous at θ_*. However, while in the certification case the bank trades towards the discontinuity (that is, $\dot{\theta}(\theta_*+) < 0$), in the intermediation case the bank trades away of the discontinuity (that is, $\dot{\theta}(\theta_*+) > 0$). The construction of the equilibrium, and the analysis of the bank's optimal control problem, is simpler in this latter case because the trajectory θ_t does not "see" the discontinuity.

ation:

$$V_i = \mathbb{E}\left[\int_0^\infty \phi e^{-(\rho+\phi)t}\left\{\mathbb{1}_{\{\kappa\leq\kappa_i\}}p_H R_f + \mathbb{1}_{\{\kappa\leq\kappa_i\}}(p_L R_f + B)\right\}dt\right].$$

The expressions differs from (23) in that the threshold cost for monitoring is replaced by κ_i.

Discussion: Trading Dynamics in Certification and Intermediation. We have seen that the equilibrium trading dynamics are remarkably different in certification and intermediation. Let us offer some further discussions regarding the fundamental mechanism behind. In general, monitoring has the property of a public good in the sense that investors can free ride the bank. The benefits from monitoring are commonly shared by all creditors, whereas the bank exclusively bears the cost. In other words, by monitoring the borrower, the bank brings positive externalities to investors and therefore benefit the entrepreneur indirectly. Therefore, the equilibrium monitoring effort is inefficiently low, and in a dynamic framework with certification, the bank reduces its probability of monitoring over time. Price impacts deter the bank from selling the loan too fast and too aggressively.

In intermediation, the role of deposits is to help the bank internalize the externality from monitoring. Indeed, these deposits are fairly priced, which reflect the probability of monitoring. The income that the bank receives from deposit insurance thus compensates its cost incurred during monitoring. In this case, both the benefits and costs of monitoring are shared by the bank and its deposits. Essentially, deposits create a market for the unique services offered by the bank, i.e., monitoring, to be fairly priced.

The above argument also explains the importance of short-term deposits to align the bank's incentive in monitoring. Indeed, in the case of instantly-maturing deposits, the value and the issuance of deposits are both continuously adjusted without any friction. If instead, depositors have a positive maturity, the externalities from monitoring will not be completely internalized by the bank.

3.3 A Special Case: Uniform Distribution

We present a special case of the model in which the monitoring cost $\tilde{\kappa}$ follows the uniform distribution. While this assumption simplifies the analysis and the expressions, it also helps us focus on the insights that do not rely on the particular features of the distribution function. Under the uniform distribution, the marginal effect of retention θ_t on the probability of monitoring is always a constant. Moreover, we assume the probability of success is $p_L = 0$ if the entrepreneur shirks. This assumption naturally leads to a result that the bank will never sell off all its loans atomically, because the resulting price will be zero.[13] As a result, $\theta_* = 0$ and $T_* \to \infty$ in the certification equilibrium.

Assumption 5. $f(\kappa) = \frac{1}{\bar{\kappa}}$ *and* $p_L = 0$.

The constrained optimal solution turns out very straightforward in this case.

Corollary 1. *Under Assumption 5, the constrained social planner's solution in Proposition 2 becomes*

$$\theta_t = \frac{1}{2 - e^{-\rho t}} \left[\left(1 - e^{-\rho t}\right) + e^{-\rho t} \frac{\Delta R - B}{\Delta R - b} \right].$$

Moreover, $\theta_0 = \frac{\Delta R - B}{\Delta R - b}$*, and* $\lim_{t \to \infty} \theta_t = \frac{1}{2}$.

3.3.1 Certification Equilibrium

Let us first describe the certification equilibrium. Simple calculation shows that $p_c(\theta) = (\kappa_c/\bar{\kappa})\Delta = \Delta^2 R_o \theta / \bar{\kappa}$. Equation (5), the bank's expected payoff when the project matures becomes $\pi_c(\theta) = \frac{1}{2\bar{\kappa}} (\Delta R_o \theta)^2$. The evolution of θ_t in (19) follows $\dot{\theta} = -\rho\theta$, which implies $\theta_t = \theta_0 e^{-\rho t}$.

Corollary 2. *In the certification equilibrium,* $R_o = R - \frac{b}{\Delta}$*. For any given* θ_0*, the firm is able to borrow*

$$L_c(\theta_0) = \Pi_c(\theta_0) + (1 - \theta_0)\, q_c(\theta_0) = \frac{\Phi}{2\bar{\kappa}} (\Delta R_o \theta_0)^2 + \frac{\Phi}{\bar{\kappa}} (\Delta R_o)^2 \theta_0 (1 - \theta_0),$$

and the entrepreneur's payoff is

$$V_c(\theta_0) = \Phi B - \frac{\phi}{2\rho + \phi} \frac{\Delta R_o \theta_0}{\bar{\kappa}} (B - b).$$

Let $\theta_0^* = 1 - \frac{\rho + \phi}{2\rho + \phi} \frac{B - b}{\Delta R_o}$.

1. *If* $I - A \le L_c(\theta_0^*)$*, the entrepreneur chooses* $\theta_0 = \theta_0^*$.
2. *If* $I - A \in (L_c(\theta_0^*), L_c(1)]$*, the entrepreneur chooses* $\theta_0 = \theta_0^{\min}$*, where* $\theta_0^{\min}$ *is the unique solution to* $L_c\left(\theta_0^{\min}\right) = I - A$.
3. *If* $I - A > L_c(1)$*, the entrepreneur is unable to borrow enough.*

[13] In the literature on durable goods monopoly, this is referred to as the "no-gap" case.

The expressions behind $L_c(\theta_0)$ and $V_c(\theta_0)$ have straightforward interpretations. Specifically, $L_c(\theta_0)$ includes $\Pi_c(\theta_0)$, the contribution from the bank and $(1-\theta_0)\,q_c(\theta_0)$, the amount that lent by investors. The entrepreneur's payoff is also clear. While ΦB captures the discounted payoff if she always shirks and obtains the higher private benefit, the second term takes into account bank monitoring, which eliminates the high private benefit B. Instead, under $p_L = 0$ and $R_f = b/\Delta$, the manager effectively receives an expected payoff b. $\Delta R_o \theta_0/\bar{\kappa}$ captures the probability of monitoring under $\theta = \theta_0$, and the term $\frac{\phi}{2\rho+\phi}$ is the effective time discount, which includes random project maturing $(\rho+\phi)$, as well as the bank's equilibrium trading strategy so that its retention θ_t declines exponentially at a rate ρ. Clearly, a higher initial retention θ_0 increases bank monitoring in equilibrium, thereby reducing the entrepreneur's expected payoff $V_c(\theta_0)$ but increasing the amount of total lending $L_c(\theta_0)$. While the borrowing amount $L_c(\theta_0)$ is maximized at $\theta_0 = 1$, the retention that maximizes the firm's overall payoff $V_c(\theta_0) + L_c(\theta_0)$ has an interior solution $\theta_0^* \in (0,1)$. If the borrowing constraint is slack so that $L_c(\theta_0^*) > I - A$, then the entrepreneur chooses θ_0^*. Otherwise, she chooses the minimum θ_0 that enables her to borrow just enough to invest.

3.3.2 Intermediation Equilibrium

Next, let us describe the intermediation equilibrium under Assumption 5. As described in Proposition 4, $\theta_t \equiv \theta_0 = 1$ always holds. Simple calculations show that $\tilde{D}(\theta) = \frac{\rho\bar{\kappa}}{\phi\Delta^2}$ is a constant, whereas

$$D^{\max}(\theta) = -\frac{\bar{\kappa}}{\Delta^2} + \sqrt{\left(\frac{\bar{\kappa}}{\Delta^2}\right)^2 + (R_o\theta)^2}.$$

At $\theta = 1$, the optimal choice of deposit is an interior solution if and only if $\tilde{D}(1) < D^{\max}(1)$, which requires Φ to be sufficiently high or equivalently the bank's cost of capital ρ to be sufficiently low. Intuitively, if the bank's cost of capital is too high compared to investors, it will issue up to the level constrained by its limited liability.

Corollary 3. *In the intermediation equilibrium, the firm is able to borrow*

$$L_i(1) = \Pi_i(1) = \Phi\pi_i(1),$$

where $\pi_i(1) = \frac{\Delta^2\left[R_o^2 - (D^*(1))^2\right]}{2\bar{\kappa}} + \frac{\rho}{\phi}D^*(1)$ *and* $D^*(1) = \min\left\{\frac{\rho\bar{\kappa}}{\phi\Delta^2}, -\frac{\bar{\kappa}}{\Delta^2} + \sqrt{\left(\frac{\bar{\kappa}}{\Delta^2}\right)^2 + (R_o)^2}\right\}$. *The entrepreneur's payoff is*

$$V_i = \Phi B - \Phi\frac{\Delta\left(R_o - D^*(1)\right)}{\bar{\kappa}}(B-b).$$

1. *If* $\Phi \geq \underline{\Phi} := \sqrt{\frac{(\bar{\kappa}/\Delta^2)^2}{(\bar{\kappa}/\Delta^2)^2 + (R_o)^2}}$, *then the bank's deposit choice satisfies* $D^*(1) = \tilde{D}(1) = \frac{\rho\bar{\kappa}}{\phi\Delta^2}$.

2. Otherwise, then the bank's deposit choice satisfies $D^*(1) = D^{\max}(1) = -\frac{\bar{\kappa}}{\Delta^2} + \sqrt{\left(\frac{\bar{\kappa}}{\Delta^2}\right)^2 + (R_o)^2}$.

Similar to V_c, the entrepreneur's payoff V_i also has an intuitive interpretation. While the first term considers the discounted payoff if she always shirks, the second term takes into account bank monitoring, which happens with conditional probability $\frac{\Delta R_o - D^*(1)}{\bar{\kappa}}$. Indeed, this expression is (4) evaluated at $\theta = 1$ and $D = D^*(1)$.

3.3.3 Initial Choice

Let us study the borrower's initial choice between certification and intermediation. Note that the entrepreneur is always able to borrow more in intermediation, regardless of the deposit issuance constraint. Therefore, if $I - A \in (L_c(1), L_i(1))$, the entrepreneur is only able to borrow through intermediation at $t = 0$. The most interesting comparison is under the case that $I - A < L_c(\theta^*)$, so that the borrowing constraint is slack in both certification and intermediation. Proposition 5 shows that certification may dominate intermediation.

Proposition 5. *Under Assumption 5 and $L_c(\theta_0^*) > I - A$, there exists a threshold B^* such that the entrepreneur chooses certification if and only if $B > B^*$.*

Let us explain the mechanism behind Proposition 5. Given $L_i(1) > L_c(\theta_0)$,[14] the entrepreneur may prefer certification only if $V_c(\theta_0^*) > V_i$. Simple calculation shows

$$V_c - V_i = \frac{\phi\Delta}{\bar{\kappa}}(B - b)\left(\frac{R_o - D^*(1)}{\rho + \phi} - \frac{R_o\theta_0}{2\rho + \phi}\right).$$

The term $\frac{R_o\theta_0}{2\rho+\phi}$ captures the fact that in certification, the bank starts with θ_0 but sells at rate ρ over time. By contrast, $\frac{R_o - D^*(1)}{\rho+\phi}$ makes it clear that an intermediating bank always retains $\theta = 1$ and the benefit from issuing deposits $D^*(1)$ will ultimately accrue to the entrepreneur. Clearly, $V_c > V_i$ if and only if

$$\frac{R_o - D^*(1)}{\rho + \phi} > \frac{R_o\theta_0}{2\rho + \phi},$$

which is necessarily the case of ρ, θ_0 and $D^*(1)$ being small. Conditional on $V_c > V_i$, a higher B leads a larger difference in the entrepreneur's payoff. If the difference gets sufficiently large, it can offset the difference in the amount of borrowing. Therefore, the entrepreneur ends up choosing certification.

This result implies that more financially-constrained entrepreneur borrows from intermediating banks such as commercial and community banks, whereas entrepreneurs with more net worth or

[14] The proof follows because $L_i(1) \geq L_c(1) > L_c(\theta_0)$ and $D^*(1) \leq \tilde{D}(1) = \frac{\rho\bar{\kappa}}{\phi\Delta^2}$.

potentially enjoy higher private benefits (or control rents) tend to borrow from certifying banks such as venture capitalists. To the best of our knowledge, this result has not been covered by previous theoretical researches. De Bettignies and Brander (2007) argues the critical distinction between VCs and banks is that only VCs provide value-adding services. Winton and Yerramilli (2008) argues that compared to banks, VCs are more capable monitors but have higher costs of capital. Similarly, Ueda (2004) assumes VCs can screen borrowers better than banks before the investment is made. Her model predicts that entrepreneurs will approach VCs when they require large investment amounts. Landier (2005) introduces two-wide hold-up problems and the stigma associated with a failed entrepreneur. He argues that if the entrepreneur is held up by investors, bank financing is more preferable. By contrast, VC financing is more preferable if investors can be held up by the entrepreneur.

Numerical Example Let us illustrate the comparison using a numerical example. Let $\phi = 1$, $b = 0$, $R = 2$, $p_H = 0.8$, $p_L = 0$, $\rho = 0.01$, and κ follows the distribution on $[0, 2]$. Under given parameters, it is easily calculated that $\tilde{D}(1) = 0.0312$, $D^{\max} = 0.3122$, and $\kappa_i = 1.5750$. We also include the option of directly borrowing from investors in which case the payoffs are $V_d = \Phi B$, $L_d = p_H R - B$, and $W_d = p_H R - (1 - \Phi) B$.

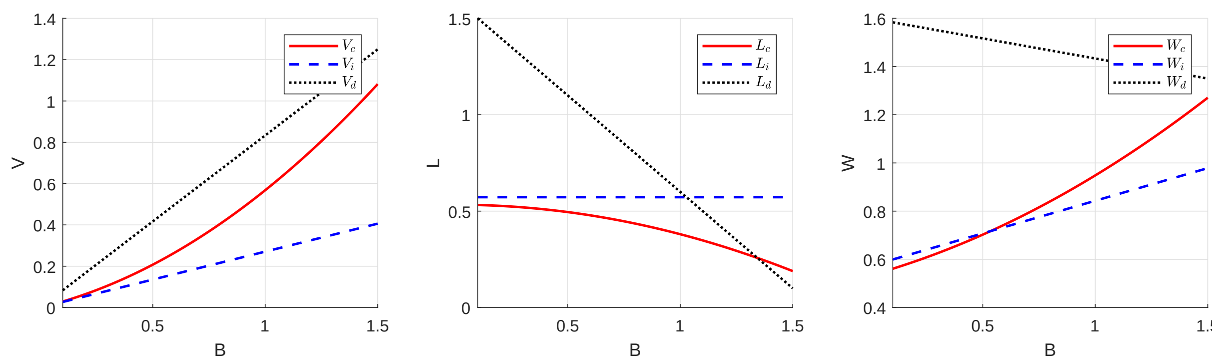


Figure 3: Valuation under Certification, Intermediation and Direct Lending

Figure 3 illustrates the payoffs when the private benefit B varies. The red (solid), blue (dashed), and black (dotted) lines respectively stand for certification, intermediation, and directly lending. The left panel describes the entrepreneur's payoff, which increases with B in all three cases. The entrepreneur always obtains the highest payoff in direct lending. Moreover, her payoff is convex in B in certification but linear in intermediation. The middle panel shows the maximum amount of borrowing, which is independent of B in intermediation but decreases in the other two cases. This is because $\theta_t \equiv 1$ in intermediation. Finally, the right panel compares the overall payoff. Clearly, certification has a higher overall payoff than intermediation once B gets sufficiently high.

One may wonder whether the entrepreneur may benefit by simultaneously borrowing from both an intermediating bank and investors. The answer is yes. Under some circumstances, the constrained-optimal allocation features the bank's retention θ_t increasing over time. This solution can be better approximated by combining intermediation and direct borrowing from outside investors. This arrangement can potentially lead to a higher entrepreneur's payoff than both certification and intermediation.

4 Extensions

4.1 Certification with Lockup and Minimum Retention

Lockup Period

Let us first consider a lockup arrangement which allows the bank to commit to θ_0 for a period $[0, t_\ell]$. Due to the stationary environment, the subgame starting from t_ℓ and the associated equilibrium are unchanged from those in 3.1. Between $[0, t_\ell]$, there is no trading, and the bank's flow payoff is $\phi\pi_c(\theta_0)$. Let L_ℓ be the total lending at $t = 0$. Under Assumption 5, it is easily shown that

$$L_\ell = \Pi_c(\theta_0) + (1-\theta_0)\left[d_c(\theta_0)\left(1-e^{-\phi t_\ell}\right) + e^{-\phi t_\ell} q_c(\theta_0)\right].$$

Again, this expression confirms the earlier result that due to the lack of commitment, the certifying bank does not benefit from its ability to trade loans at all: it is able to lend exactly $\Pi_c(\theta_0)$ regardless of t_ℓ. Meanwhile, investors are willing to lend out more as the lockup period t_L gets longer, since $q_c(\theta_0) = \Phi d_c(\theta_0)$. Therefore, $\frac{\partial L_\ell}{\partial t_\ell} > 0$ so that the lockup period increase the total amount of lending: the incremental lending comes from investors' willingness to lend due to the bank's improved commitment.

Moreover, let us use V_ℓ to denote the entrepreneur's overall payoff. Simple calculation shows that

$$V_\ell = \Phi\left[\frac{\Delta R_o\theta_0}{\bar\kappa}b + \left(1-\frac{\Delta R_o\theta_0}{\bar\kappa}\right)B\right]\left(1-e^{-(\rho+\phi)t_\ell}\right) + e^{-(\rho+\phi)t_\ell}V_c(\theta_0).$$

Intuitively, during the lockup period $[0, t_\ell]$, the entrepreneur gets monitored and receive b with probability $\Delta R_o\theta_0/\bar\kappa$; otherwise, she receives B. It is easily derived that $\frac{\partial V_\ell}{\partial t_\ell} < 0$, so that a longer lockup period leads to a lower payoff to the entrepreneur. Intuitively, the entrepreneur receives b with monitoring but B without monitoring. While a longer lockup period increases monitoring, it also reduces the entrepreneur's equilibrium payoff.

Let $W_\ell = V_\ell + L_\ell$ be the aggregate social welfare, which is also the objective function that the entrepreneur tries to maximize at $t = 0$.

Corollary 4 (Optimal Lockup Period)**.** *For $\theta_0 = \theta_0^*$ as in Corollary 2, the aggregate social welfare attains the maximum as $t_\ell \to \infty$.*

Minimum Retention

According to section 941 of the Dodd-Frank Act, securitizers are required to retain no less than 5% of the credit risks associated with any securitization to perform intermediation services. This rule is commonly known as risk retention. In Feb 2018, the circuit court exempt Collateralized Loan Obligation (CLO) funds, We evaluate such a policy on a minimum retention requirement on the bank.[15]

Suppose the bank must hold at least a fraction $\underline{\theta}$ of the loans on its balance sheet. The equilibrium is qualitatively similar to the one in Proposition 3: There is some threshold $\tilde{\theta}$ such that there is smooth trading for $\theta \in (\tilde{\theta}, 1]$, and there is an atom $\tilde{\theta}$ where the holdings jump to $\underline{\theta}$. For the same reasons as liquidation has to be random in the certification case, the jump from $\tilde{\theta}$ to $\underline{\theta}$ must also happen with a random delay.

The construction of the equilibrium is similar to the one without minimum retention. The only difference is that we need to consider the incentives to jump to the constrain $\underline{\theta}$ rather than the incentives to fully liquidate the loan portfolio (that is, to jump to $\theta = 0$). Because the construction, and the intuition behind the equilibrium conditions, is similar, we only provide a brief sketch of the construction.

4.2 Intermediation with Long-term Deposits and Deposit Subsidies

Long-term Deposits

We solve the intermediation equilibrium in which the bank can only issue long-term deposits at $t = 0$, i.e., deposits that only mature with the project at τ_ϕ. This extension highlights the role of short-term deposits in commitment. Specifically, we are going to show that under long-term deposits, the equilibrium is identical to the certification equilibrium, in which the bank trades loans over time.

Let $D_0 = D$ be the amount of deposits that the bank issues at $t = 0$. Note that after $t = 0$, the bank can no longer raise any deposits. The definitions for κ_i, $\hat{p}(\theta, D)$, and $\hat{\pi}_i(\theta, D)$ follow from those in section 2. Since $D_0 = D$ is only chosen at $t = 0$, we suppress these functions' dependence

[15] Securitization is identical to loan sales in our setup, given the binary outcome of the final cash flows.

on D and therefore refer to $\kappa_i(\theta)$, $p(\theta)$, and $\pi_i(\theta)$, respectively. The bank's maximization problem becomes

$$\max_{\{\theta_t,\, D_0\}} \; (1-y_0)\,D_0 + \mathbb{E}\left[\int_0^\infty e^{-(\rho+\phi)t}\Big(\phi\pi_i(\theta_t)\,dt + dG(\theta_t)\Big)\right],$$

where y_0 compensates the depositors' credit risk. Note that after $t=0$, the bank only chooses its trading strategy, and the solution method as well as the equilibrium outcomes naturally follow the certification equilibrium in subsection 3.1. In fact, they two are identical because depositors are effectively investors who directly invest in the venture.[16]

Deposit Subsidies

Next, we introduce an extension where the deposit rates are partially subsidized by the government. One can think about this as either deposit insurance or the implicit guarantee from government bailout.[17] We will show that once the subsidy gets sufficiently high, an intermediating bank can no longer commit to its retention but instead sells loans over time, just as a certifying bank. This exercise highlights the importance of the deposit rate in aligning the bank's incentives to commit to its retentions.

Specifically, let us assume the bank only needs to pay a fraction ξ of the deposit rate so that Equation (10) becomes

$$y_t = \phi\xi\left(1-\hat{p}(\theta, D)\right),$$

where $\xi \in (0,1)$. The analysis follows that in subsection 3.2, in which deposit issuance and trading are solved sequentially.

Corollary 5. *There exists a $\xi_\dagger$ such that in the intermediation equilibrium, $\dot{\theta} < 0$ if $\xi < \xi_\dagger$ for θ sufficiently large.*

Intuitively, the bank no longer has the incentives to retain its loans if the deposit is mostly subsidized by the government.[18]

[16] There is a subtle difference that long-term depositors cannot trade their deposits, whereas investors in certification can sell the loans. However, given that in the certification equilibrium, the bank sells the loan (and equivalently investors buy the loan) over time, this difference does not affect the result. In other words, in the equilibrium with long-term deposits, depositors buy loans in the secondary market after $t=0$.

[17] In the U.S., deposit insurance takes the form of a maximum guaranteed amount which has been \$250,000 since 2010. There is a one-to-one mapping between the maximum insurance amount and the parameter ξ introduced later on. To see this, note that one can think about the deposit rate is $y_t = 0$ for deposits below $\$250,000$ but follows (10) for deposits above. Our parameter ξ captures the fraction of deposits that are above the limit.

[18] This result is no longer true when the deposit-issuance constraint binds, in which case the bank will have incentives

5 Conclusion and Implications

This paper presents a dynamic theory of indirect financing when intermediaries cannot commit to its retention over time. We show two equivalence results in a static environment fail in the dynamic framework. First, the implementation of the financial structure matters. Certification will in general leads to a lower amount of borrowing compared with intermediation. Second, maximizing borrowing amount is different from maximizing the entrepreneur's expected payoff. In particular, if the entrepreneur can borrow enough under both structures, she may prefer certification over intermediation for a higher expected payoff.

Our result is consistent with the empirical evidence in Berger and Schaeck (2011) that entrepreneurs substitute venture capital for banking relationships.[19] Moreover, their paper argues against the view that firms obtain venture capital when bank financing is difficult to obtain. Instead, firms seem to be aware of which type of financing is more appropriate for them.

In the paper, we only allow for one bank to monitor after $t = 0$. If we allow the entrepreneur to sign contracts with multiple banks that all can monitor after time 0, there will be two new effects. First, if the monitoring costs are *i.i.d.* across different banks, introducing multiple banks naturally leads to a diversification effect, which increases the chances of monitoring. However, there is a pecuniary externality when one bank trades its loans over time in the sense that it only has partial price impact. As a result, banks will sell faster. The overall effect is therefore unclear.

For tractability reasons, we have assumed the entrepreneur is able to commit to her retention over time. Relaxing this assumption and studying the interaction between the commitment problems between the borrower and the bank can be an interesting research for future extensions. In particular, we can study the role of many sophisticated contracts such as covenants in preventing the commitment issues. Another interesting extension is along the lines of Gorton and Pennacchi (1990). Given that a certifying bank is tempted to sell its loans, can it be the case that a more carefully designed security between could mitigate the problem?

to retain loans. The overall effect in this case is unclear.

[19]Their explanation focuses on the business expertise offered by venture capitalists, though.

References

Ausubel, L. M. and R. J. Deneckere (1989). Reputation in bargaining and durable goods monopoly. *Econometrica*, 511–531.

Bardi, M. and I. Capuzzo-Dolcetta (2008). *Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations*. Springer Science.

Barles, G., A. Briani, E. Chasseigne, and C. Imbert (2018). Flux-limited and classical viscosity solutions for regional control problems. *ESAIM: Control, Optimisation and Calculus of Variations 24*(4), 1881–1906.

Berger, A. N. and K. Schaeck (2011). Small and medium-sized enterprises, bank relationship strength, and the use of venture capital. *Journal of Money, Credit and banking 43*(2-3), 461–490.

Bizer, D. S. and P. M. DeMarzo (1992). Sequential banking. *Journal of Political Economy 100*(1), 41–61.

Calomiris, C. W. and C. M. Kahn (1991). The role of demandable debt in structuring optimal banking arrangements. *The American Economic Review*, 497–513.

De Bettignies, J.-E. and J. A. Brander (2007). Financing entrepreneurship: Bank finance versus venture capital. *Journal of Business Venturing 22*(6), 808–832.

DeMarzo, P. and Z. He (2016). Leverage dynamics without commitment. Technical report, National Bureau of Economic Research.

DeMarzo, P. M. and B. Urošević (2006). Ownership dynamics and asset pricing with a large shareholder. *Journal of Political Economy 114*(4), 774–815.

Diamond, D. W. (1984). Financial intermediation and delegated monitoring. *The review of economic studies 51*(3), 393–414.

Diamond, D. W. (2004). Presidential address, committing to commit: short-term debt when enforcement is costly. *The Journal of Finance 59*(4), 1447–1479.

Diamond, D. W. and R. G. Rajan (2001). Liquidity risk, liquidity creation, and financial fragility: A theory of banking. *Journal of Political Economy 109*(2), 287–327.

Fuchs, W. and A. Skrzypacz (2010). Bargaining with arrival of new traders. *American Economic Review 100*(3), 802–36.

Gorton, G. and G. Pennacchi (1990). Financial intermediaries and liquidity creation. *The Journal of Finance 45*(1), 49–71.

Grossman, S. J. and O. D. Hart (1980). Takeover bids, the free-rider problem, and the theory of the corporation. *The Bell Journal of Economics*, 42–64.

Holmstrom, B. and J. Tirole (1997). Financial intermediation, loanable funds, and the real sector. *Quarterly Journal of Economics 112*(3), 663–691.

Landier, A. (2005). Entrepreneurship and the stigma of failure. *Available at SSRN 850446*.

Milgrom, P. and I. Segal (2002). Envelope theorems for arbitrary choice sets. *Econometrica 70*(2), 583–601.

Myers, S. C. and R. G. Rajan (1998). The paradox of liquidity. *The Quarterly Journal of Economics 113*(3), 733–771.

Rampini, A. A. and S. Viswanathan (2019). Financial intermediary capital. *The Review of Economic Studies 86*(1), 413–455.

Ueda, M. (2004). Banks versus venture capital: Project evaluation, screening, and expropriation. *The Journal of Finance 59*(2), 601–621.

Winton, A. and V. Yerramilli (2008). Entrepreneurial finance: Banks versus venture capital. *Journal of financial economics 88*(1), 51–79.

Appendix

A Certification

Lemma 3. *The bank with retention θ never sells a fraction of the loans.*

Proof. Suppose the bank with retention θ sells $\theta^+ - \theta$, where $\theta^+ > 0$, and that after this it continuous trading smoothly. Multiple jumps are ruled out without loss of generality. In this case, the overall trading gains are $dG(\theta) + \Pi_c(\theta^+) - \Pi_c(\theta) = (\theta - \theta^+) q_c(\theta^+) + \Pi_c(\theta^+) - \Pi_c(\theta)$, where $dG(\theta)$ is the instant trading gain and $\Pi(\theta^+) - \Pi(\theta)$ are the gains (negative loss) in its continuation value. Block trading is suboptimal as long as

$$\theta = \arg\max_{\theta^+} \Big\{ \Pi_c(\theta^+) + (\theta - \theta^+) q_c(\theta^+) \Big\}. \tag{32}$$

It is easy to verify that the first order condition is always satisfied at $\theta^+ = \theta$, thus it suffice to show that the second order condition for global optimality is satisfied. $\square$

Verification of Optimality Trading Strategy

In this section, we complete the characterization of the equilibrium by verifying that the equilibrium trading strategy maximizes the bank's payoff given the price function $q(\theta)$. Because the payoff in a mixed strategy equilibrium is given by the payoff of any pure strategy in its support, we can restrict attention to pure strategies in the verification of optimality. A trading strategy for the bank is given by right continuous function with left limits. A trading strategy is admissible if it can be decomposed as

$$\theta_t = \int_0^t \dot{\theta}_t^c dt + \sum_{k \geq 0} (\theta_{t_k}^d - \theta_{t_k -}^d),$$

for some bounded function $\dot{\theta}_t^c$. We denote the set of admissible trading strategies by Θ. The bank's optimization problem is to choose $\theta \in \Theta$ to maximize its payoff

$$\Pi^*(\theta_0) = \sup_{\theta \in \Theta} \int_0^\infty e^{-(\rho+\phi)t} \big(\phi \pi(\theta_t) - \dot{\theta}_t^c q(\theta_t) \big) dt - \sum_{k \geq 0} e^{-(\rho+\phi)t_k} q(\theta_{t_k})(\theta_{t_k}^d - \theta_{t_k -}^d). \tag{33}$$

Due to the discontinuity in the price price function $q(\theta)$, the Hamilton-Jacobi-Bellman (HJB) equation is discontinuous at θ_*, so we need to resort to the theory of viscosity solutions for the analysis of the bank's problem. Our problem is a particular case of the general class of optimal control problems in stratified domains studied by Barles et al. (2018). Our proof relies on their characterization of the value function using viscosity solution methods. The analysis in Barles et al. (2018) does not consider the case in which the trajectory of the state variable can be discontinuous (impulse control). However, as we show below, we can approximate a trading $\theta_t \in \Theta$ by an absolutely continuous trading strategy with derivative $|\dot{\theta}_t| \leq N$ for some N large enough (the approximation is in the sense that it yields a similar payoff). Thus, we can consider a sequence of optimization problems

$$\Pi_N^*(\theta_0) = \sup_{|\dot{\theta}_t| \le N} \int_0^\infty e^{-(\rho+\phi)t} \big(\phi\pi(\theta_t) - \dot{\theta}_t q(\theta_t)\big) dt, \tag{34}$$

and verify that, for any $\theta \in [0,1]$, $\Pi_N^*(\theta) \to \Pi(\theta)$, where

$$\Pi(\theta) = \begin{cases} \Phi\pi(\theta) & \text{if } \theta \in [\theta_*, 1] \\ q(0)\theta & \text{if } \theta \in [0, \theta_*) \end{cases}.$$

The following Lemma establishes that we can indeed consider the limit of bounded absolutely continuous strategies.

Lemma 4. *For any $\theta_0 \in [0,1]$, $\lim_{N\to\infty} \Pi_N^*(\theta_0) = \Pi^*(\theta_0)$.*

Proof. Let $\theta_t^{\epsilon*}$ be and ϵ-optimal policy (at this point in the proof we have not established existence of an optimal policy). For any $k \ge 0$, let $\Delta_k \equiv \inf\{\Delta > 0 : \theta_{t_k-\Delta}^{\epsilon*} + \mathrm{sgn}(\theta_{t_k}^{\epsilon*} - \theta_{t_k-}^{\epsilon*})N\Delta = \theta_{t_k}^{\epsilon*}\}$ (we can find Δ_k if N is large enough as $|\dot{\theta}_t^{\epsilon c*}| \le M$ for some finite M). Consider the policy $\hat{\theta}_t^N = \theta_t^{\epsilon*}$ if $t \notin \cup_{k\ge 0}(t_k - \Delta_k, t_k)$, and $\hat{\theta}_t^N = \theta_{t_k-\Delta_k}^{\epsilon*} + \mathrm{sgn}(\theta_{t_k}^{\epsilon*} - \theta_{t_k-}^{\epsilon*})N(t - t_k + \Delta_k)$ if $t \in \cup_{k\ge0}(t_k - \Delta_k, t_k)$. The difference between the payoff of $\theta_t^{\epsilon*}$ and $\hat{\theta}_t^N$ is

$$\begin{aligned}
\Pi^{\epsilon*}(\theta_0) - \hat{\Pi}_N &= \sum_{k\ge0} \Bigg\{ \int_{t_k-\Delta_k}^{t_k} e^{-(\rho+\phi)t}\big(\phi\pi(\theta_t^{\epsilon*}) - \dot{\theta}_t^{\epsilon c*} q(\theta_t^{\epsilon*}) - \phi\pi(\hat{\theta}_t^N)\big)dt \\
&\quad + \int_{t_k-\Delta_k}^{t_k} e^{-(\rho+\phi)t}\mathrm{sgn}(\theta_{t_k}^{\epsilon*} - \theta_{t_k-}^{\epsilon*})Nq(\hat{\theta}_t^N)dt - e^{-(\rho+\phi)t_k} q(\theta_{t_k}^{\epsilon*})(\theta_{t_k}^{\epsilon d*} - \theta_{t_k-}^{\epsilon d*}) \Bigg\} - \epsilon \\
&= \sum_{k\ge0} \Bigg\{ \int_{t_k-\Delta_k}^{t_k} e^{-(\rho+\phi)t}\big(\phi\pi(\theta_t^{*}) - \dot{\theta}_t^{\epsilon c*} q(\theta_t^{*}) - \phi\pi(\hat{\theta}_t^N)\big)dt \\
&\quad + \frac{\theta_{t_k}^{\epsilon*} - \theta_{t_k-\Delta_k}^{\epsilon*}}{\Delta_k} \int_{t_k-\Delta_k}^{t_k} e^{-(\rho+\phi)t} q(\hat{\theta}_t^N)dt - e^{-(\rho+\phi)t_k} q(\theta_{t_k}^{\epsilon*})(\theta_{t_k}^{\epsilon d*} - \theta_{t_k-}^{\epsilon d*}) \Bigg\} - \epsilon.
\end{aligned}$$

For all $k \ge 0$, we have that $\Delta_k \downarrow 0$ as $N \to \infty$. It follows that

$$\lim_{\Delta_k \downarrow 0} \frac{1}{\Delta_k} \int_{t_k-\Delta_k}^{t_k} e^{-(\rho+\phi)t} q(\hat{\theta}_t^N)dt = \begin{cases} e^{-(\rho+\phi)t_k} q(\theta_{t_k}^{\epsilon*}-) & \text{if } \theta_{t_k}^{\epsilon*} > \theta_{t_k-}^{\epsilon*} \\ e^{-(\rho+\phi)t_k} q(\theta_{t_k}^{\epsilon*}+) & \text{if } \theta_{t_k}^{\epsilon*} < \theta_{t_k-}^{\epsilon*}. \end{cases}$$

The price function is right continuous so $q(\theta_{t_k}^{\epsilon*}+) = q(\theta_{t_k}^{\epsilon*})$. We can conclude that

$$\lim_{N\to\infty}(\Pi^{\epsilon*}(\theta_0) - \hat{\Pi}_N(\theta_0)) = \sum_{k\ge0} e^{-(\rho+\phi)t_k}\big(q(0) - q(\theta_*)\big)(\theta_{t_k}^* - \theta_{t_k-}^*)^+ \mathbf{1}_{\{\theta_{t_k}^* = \theta_*\}} - \epsilon \le 0.$$

Because this holds for any $\epsilon > 0$, we can conclude that $\lim_{N\to\infty}(\Pi^*(\theta_0) - \hat{\Pi}_N(\theta_0)) \le 0$, and given that $\Pi^*(\theta_0) \ge \hat{\Pi}_N(\theta_0)$, we get $\lim_{N\to\infty} \hat{\Pi}_N(\theta_0) = \Pi^*(\theta_0)$. For N large enough, the policy $\hat{\theta}_t^N$ satisfies $|\dot{\hat{\theta}}_t^N| \le N$

(this can be guarantee because for any ϵ there is M such that $|\dot{\theta}_t^{\epsilon c*}| \leq M$), so its payoff, $\hat{\Pi}_N(\theta_0)$ provides a lower bound to $\Pi_N^*(\theta_0)$, which means that $\lim_{N\to\infty} \Pi_N^*(\theta_0) = \Pi^*(\theta_0)$. $\square$

This shows that the value function converges (pointwise) to the one in the equilibrium under consideration. Hence, we can verify the optimality of the bank's strategy by analyzing the control problem (34). For future reference, recall that the price function in the control problem (34) is given by

$$q(\theta) = \begin{cases} \Phi\pi'(\theta) & \text{if } \theta \geq \theta_* \\ p_L R\,(1-\alpha) & \text{if } \theta < \theta_* \end{cases}, \tag{35}$$

where the threshold θ_* is given by $\Phi\pi(\theta_*) = q(0)\theta_*$. Notice that we are not computing the equilibrium in a model in which the bank is restricted to use absolutely continuous trading strategies with bounded derivative $\dot{\theta}_t$, but rather considering the equilibrium price function in the general case, and then considering a sequence of auxiliary optimization problems to construct the value function. Because the expected payoff of the candidate equilibrium strategy is equal to the value function, it is necessarily optimal.

The Hamilton-Jacobi-Bellman equation (HJB) for the optimization problem (34) is

$$(\rho+\phi)\Pi_N(\theta) - H(\theta, \Pi_N'(\theta)) = 0, \tag{36}$$

where H

$$H(\theta, \Pi_N') \equiv \phi\pi(\theta) + \max_{|\dot{\theta}|\leq N} \left\{ \dot{\theta}\big(\Pi_N' - q(\theta)\big) \right\}. \tag{37}$$

We guess and verify that, for N large enough, the solution (in the viscosity sense) of the previous equation is

$$\Pi_N(\theta) = \begin{cases} \Phi\pi(\theta) & \text{if } \theta \in [\theta_*, 1] \\ e^{-\frac{\rho+\phi}{N}(\theta_*-\theta)}\Phi\pi(\theta_*) + \frac{(\rho+\phi)}{N}\int_\theta^{\theta_*} e^{-\frac{\rho+\phi}{N}(y-\theta)}\left(\Phi\pi(y) - \frac{N}{\rho+\phi}q(0)\right)dy & \text{if } \theta \in [\tilde{\theta}_N, \theta_*) \\ \frac{N}{\rho+\phi}\left(1 - e^{-\frac{(\rho+\phi)}{N}\theta}\right)q(0) + \frac{(\rho+\phi)}{N}\int_0^\theta e^{-\frac{(\rho+\phi)}{N}(\theta-y)}\Phi\pi(y)dy & \text{if } \theta \in [0, \tilde{\theta}_N), \end{cases} \tag{38}$$

where $\tilde{\theta}_N$ is the unique solution on $[0,1]$ to the equation

$$\frac{N}{\rho+\phi}\left(1 - e^{-\frac{(\rho+\phi)}{N}\tilde{\theta}_N}\right)q(0) + \frac{(\rho+\phi)}{N}\int_0^{\tilde{\theta}_N} e^{-\frac{(\rho+\phi)}{N}(\tilde{\theta}_N - y)}\Phi\pi(y)dy = \\ e^{-\frac{\rho+\phi}{N}(\theta_*-\tilde{\theta}_N)}\Phi\pi(\theta_*) + \frac{(\rho+\phi)}{N}\int_{\tilde{\theta}_N}^{\theta_*} e^{-\frac{\rho+\phi}{N}(y-\tilde{\theta}_N)}\left(\Phi\pi(y) - \frac{N}{\rho+\phi}q(0)\right)dy \tag{39}$$

A.1 Auxiliary Lemmas

Before proceeding with the verification theorem, we provide several Lemmas providing properties of our candidate value function $\Pi_N(\theta)$ that will be later used in the verification

Lemma 5. *If $\Phi\pi_c(1) > p_L R_o > 0$, then there exists a unique $\theta_* \in (0,1)$ solving the equation*

$$\theta_* q_c(0) = \Phi\pi_c(\theta_*) \tag{40}$$

If $p_L = 0$, then $\theta_* = 0$ is the unique solution to (40) on $[0,1]$.

Proof. As $\Phi\pi_c(0) = 0$, equation (40) is trivially satisfied at $\theta_* = 0$, we want to show that if $\Phi\pi_c(1) > p_L R_o = q_c(0)$, then there is a non trivial solution $\theta_* > 0$ that also satisfies equation (40). First, if $\Phi\pi_c(1) > p_L R_o$, then the right hand side of equation (40) is strictly larger than its left hand side evaluated at $\theta_* = 1$. Second, as $\Phi\pi_c'(0) < q_c(0)$ it follows that for ε small enough $\varepsilon q_c(0) > \Phi\pi_c(\varepsilon)$. Thus, it follows from continuity that a non trivial solution exists on $(0,1)$. Uniqueness follows because

$$\begin{aligned} q_c(0) - \Phi\pi_c'(\theta_*) &= \frac{\Phi\pi_c(\theta_*)}{\theta_*} - \Phi p_c(\theta_*) R_o \\ &= \Phi\left[p_c(\theta_*) R_o - \frac{1}{\theta_*}\int_0^{\kappa_c(\theta_*)} \kappa dF(\kappa)\right] - \Phi p_c(\theta_*) R_o < 0, \end{aligned}$$

so the function $\theta q_c(0) - \Phi\pi_c(\theta)$ single crosses 0 from above, which implies $\theta q_c(0) > \Phi\pi_c(\theta)$ on $\theta \in (0,\theta_*)$ and $\theta q_c(0) < \Phi\pi_c(\theta)$ on $\theta \in (\theta_*, 1]$. Finally, if $p_L = 0$, then $\Phi\pi_c'(0) = q_c(0) = 0$. It follows then from the convexity of $\pi_c(\theta)$ that $\theta_* = 0$ is a global maximum of $\theta q_c(0) - \Phi\pi_c(\theta)$, which means that $\theta q_c(0) < \Phi\pi_c(\theta)$ for all $\theta > 0$. □

Lemma 6. *There is a unique solution $\tilde{\theta}_N \in (0,\theta_*)$ to equation (39).*

Proof. First, we show existence. Given the definition of θ_* and the convexity of $\pi(\theta)$ we have that $\Phi\pi(\theta) < \theta q(0)$ for all $\theta < \theta_*$. Hence,

$$\frac{N}{\rho+\phi}\left(1 - e^{-\frac{(\rho+\phi)}{N}\theta_*}\right) q(0) + \frac{(\rho+\phi)}{N}\int_0^{\theta_*} e^{-\frac{(\rho+\phi)}{N}(\theta-y)}\Phi\pi(y)dy < \Phi\pi(\theta_*).$$

We also have that

$$\begin{aligned} e^{-\frac{\rho+\phi}{N}\theta_*}\Phi\pi(\theta_*) + \frac{(\rho+\phi)}{N}\int_0^{\theta_*} e^{-\frac{\rho+\phi}{N}y}\left(\Phi\pi(y) - \frac{N}{\rho+\phi}q(0)\right)dy &\leq \Phi\pi(\theta_*) - \frac{N}{(\rho+\phi)}\left(1 - e^{-\frac{\rho+\phi}{N}\theta_*}\right) q(0) \\ &\leq \Phi\pi(\theta_*) - \theta_* q(0) = 0. \end{aligned}$$

The existence of a solution follows from the intermediate value theorem. To show uniqueness we consider the derivative of the difference between the left and the right hand sides of equation (39) evaluated at θ, which we denote by $G'(\theta)$.

$$G'(\theta) = e^{-\frac{(\rho+\phi)}{N}\theta}q(0) + \frac{(\rho+\phi)}{N}\Phi\pi(\theta) - \frac{(\rho+\phi)^2}{N^2}\int_0^{\theta} e^{-\frac{(\rho+\phi)}{N}(\theta-y)}\Phi\pi(y)dy$$
$$- \frac{(\rho+\phi)}{N}e^{-\frac{\rho+\phi}{N}(\theta_*-\theta)}\Phi\pi(\theta_*) + \frac{(\rho+\phi)}{N}\left(\Phi\pi(\theta) - \frac{N}{\rho+\phi}q(0)\right)$$
$$- \frac{(\rho+\phi)^2}{N^2}\int_{\theta}^{\theta_*} e^{-\frac{\rho+\phi}{N}(y-\theta)}\left(\Phi\pi(y) - \frac{N}{\rho+\phi}q(0)\right)dy$$

From here we get that

$$\begin{aligned}
G'(\theta_*) &= -\left(1 - e^{-\frac{(\rho+\phi)}{N}\theta_*}\right)q(0) - \frac{(\rho+\phi)^2}{N^2}\int_0^{\theta_*} e^{-\frac{(\rho+\phi)}{N}(\theta_*-y)}\Phi\pi(y)dy + \frac{(\rho+\phi)}{N}\Phi\pi(\theta_*) \\
&\leq \frac{(\rho+\phi)}{N}\left(\Phi\pi(\theta_*) - \theta_* q(0)\right) = 0 \\
G'(0) &= -\frac{(\rho+\phi)}{N}e^{-\frac{\rho+\phi}{N}\theta_*}\Phi\pi(\theta_*) - \frac{(\rho+\phi)^2}{N^2}\int_0^{\theta_*} e^{-\frac{\rho+\phi}{N}y}\left(\Phi\pi(y) - \frac{N}{\rho+\phi}q(0)\right)dy \\
&\geq -\frac{(\rho+\phi)}{N}\Phi\pi(\theta_*) + \left(1 - e^{-\frac{\rho+\phi}{N}\theta_*}\right)q(0) \\
&\geq \frac{(\rho+\phi)}{N}\left(\theta_* q(0) - \Phi\pi(\theta_*)\right) = 0
\end{aligned}$$

Moreover, we get that, for any $\theta \in (0, \theta_*)$, $G(\theta) = 0$ implies

$$\begin{aligned}
G'(\theta) &= \frac{2(\rho+\phi)}{N}\left[\Phi\pi(\theta) - \left(\frac{N}{\rho+\phi}\left(1 - e^{-\frac{(\rho+\phi)}{N}\theta}\right)q(0) + \frac{(\rho+\phi)}{N}\int_0^{\theta} e^{-\frac{(\rho+\phi)}{N}(\theta-y)}\Phi\pi(y)dy\right)\right] \\
&\leq \frac{2(\rho+\phi)}{N}\left[\Phi\pi(\theta) - \theta q(0)\right] < 0.
\end{aligned}$$

It follows that $G(\theta)$ single crosses 0, so there is a unique solution to the equation $G(\theta) = 0$. $\square$

Lemma 7. *There is $\tilde{N}$ such that, for all $N > \tilde{N}$, $\Pi'_N(\theta) < q(0)$ on $(0, \tilde{\theta}_N)$ and $\Pi'_N(\theta) > q(0)$ on $(\tilde{\theta}_N, \theta_*)$.*

Proof. First, we very that $\Pi'_N(\theta) < q(0)$ on $(0, \tilde{\theta}_N)$. The derivative of $\Pi_N(\theta) - \theta q(0)$ on $(0, \tilde{\theta}_N)$ is given by

$$\begin{aligned}
\Pi'_N(\theta) - q(0) &= \frac{\rho+\phi}{N}\Phi\pi(\theta) - \left(1 - e^{-\frac{(\rho+\phi)}{N}\theta}\right)q(0) - \frac{(\rho+\phi)^2}{N^2}\int_0^{\theta} e^{-\frac{(\rho+\phi)}{N}(\theta-y)}\Phi\pi(y)dy \\
&\leq \frac{\rho+\phi}{N}\Phi\pi(\theta) - \left(1 - e^{-\frac{(\rho+\phi)}{N}\theta}\right)q(0) \leq \frac{\rho+\phi}{N}\left(\Phi\pi(\theta) - \theta q(0)\right) < 0.
\end{aligned}$$

The derivative of $\Pi_N(\theta) - \theta q(0)$ on $(\tilde{\theta}_N, \theta_*)$ is given by

$$\Pi'_N(\theta) - q(0) = \frac{\rho+\phi}{N}\left[e^{-\frac{\rho+\phi}{N}(\theta_*-\theta)}\Phi\pi(\theta_*) - \Phi\pi(\theta) + \frac{(\rho+\phi)}{N}\int_{\theta}^{\theta_*} e^{-\frac{\rho+\phi}{N}(y-\theta)}\left(\Phi\pi(y) - \frac{N}{\rho+\phi}q(0)\right)dy\right]$$

Differentiating the HJB equation we get that

$$\Pi_N''(\theta) = \frac{(\rho+\phi)}{N}\left(\Pi_N'(\theta) - \Phi\pi'(\theta)\right)$$
$$\Pi_N'''(\theta) = \frac{(\rho+\phi)}{N}\left(\Pi_N''(\theta) - \Phi\pi''(\theta)\right).$$

From here we get that $\Pi_N''(\theta) = 0 \implies \Pi_N'''(\theta) < 0$, so we the function $\Pi_N'(\theta)$ is quasi-concave on $(\tilde{\theta}_N, \theta_*)$. Moreover, $\Pi_N'(\theta_*-) = q(0)$, and

$$\Pi_N'(\tilde{\theta}_N+) = q(0) + \frac{(\rho+\phi)}{N}\left(\Pi_N(\tilde{\theta}_N+) - \Phi\pi(\tilde{\theta}_N)\right) > q(0),$$

so we can conclude that $\Pi_N'(\theta) > q(0)$ on $(\tilde{\theta}_N, \theta_*)$ as long as $\Pi_N(\tilde{\theta}_N+) > \Phi\pi(\tilde{\theta}_N)$, which follows from

$$\begin{aligned}
\Pi_N(\tilde{\theta}_N+) - \Phi\pi(\tilde{\theta}_N) &= \Pi_N(\tilde{\theta}_N-) - \Phi\pi(\tilde{\theta}_N) \\
&= \frac{N}{\rho+\phi}\left(1 - e^{-\frac{(\rho+\phi)}{N}\tilde{\theta}_N}\right)q(0) - \Phi\pi(\tilde{\theta}_N) + \frac{(\rho+\phi)}{N}\int_0^{\tilde{\theta}_N} e^{-\frac{(\rho+\phi)}{N}(\tilde{\theta}_N - y)}\Phi\pi(y)dy \\
&\geq \frac{N}{\rho+\phi}\left(1 - e^{-\frac{(\rho+\phi)}{N}\tilde{\theta}_N}\right)q(0) - \Phi\pi(\tilde{\theta}_N) \\
&= \tilde{\theta}_N q_0 - \frac{(\rho+\phi)\tilde{\theta}_N^2}{N} - \Phi\pi(\tilde{\theta}_N) + O\big(1/N^2\big).
\end{aligned}$$

$\tilde{\theta}_N q_0 > \Phi\pi(\tilde{\theta}_N)$ because θq_0 single crosses $\Phi\pi(\theta)$ at $\theta_* \geq \tilde{\theta}_N$. Hence, there is $\tilde{N}$ such that, for all $N \geq \tilde{N}$, we have $\Pi_N(\tilde{\theta}_N+) > \Phi\pi(\tilde{\theta}_N)$. □

Lemma 8. *Let*

$$\Pi(\theta) = \begin{cases} \Phi\pi(\theta) & \text{if } \theta \in [\theta_*, 1] \\ q(0)\theta & \text{if } \theta \in [0, \theta_*) \end{cases}.$$

Then, for any $\theta \in [0,1]$

$$\lim_{N\to 0} \Pi_N(\theta) = \Pi(\theta).$$

Proof. For all $\theta \geq \theta_*$, $\Pi_N(\theta) = \Pi(\theta)$, and, for any $\theta < \theta_*$, $\lim_{N\to\infty} \Pi_N(\theta) = \theta q(0) = \Pi(\theta)$ by L'Hopital's rule. □

A.2 Verification of Optimality

We start providing the necessary definitions from the theory of viscosity solutions, together with the relevant results from the theory of optimal control in stratified domains in Barles et al. (2018). We make some changes in notation to make it consistent with our setting, and to translate their minimization problem into a maximization one. While Barles et al. (2018) considers the state space to be the complete real line, the state space in our case is $[0,1]$. However, we can extend the state space by letting the payoff on the complement of $[0,1]$ to be sufficiently low. This can be achieved by adding a penalization term, and setting the flow payoff equal to $\phi\pi(1) - \dot{\theta}q(1) - k|\theta - 1|$ for $\theta > 1$, and $\phi\pi(0) - \dot{\theta}q(0) - k|\theta|$ for $\theta < 0$. By choosing k large enough, we can ensure that the optimal solution never exits the interval $[0, 1]$. Due to the discontinuity in the Hamiltonian at θ_*, a viscosity solution might fail to be unique. In order to fully characterize the value function we need to specify its behavior at θ_*. This is done in Barles et al. (2018) by considering the concept of *Flux-limited* sub- and supersolutions. Letting $\Omega_0 = (-\infty, \theta_*)$ and $\Omega_1 = (\theta_*, \infty)$, we consider the equation

$$\begin{cases} (\rho+\phi)\Pi - H_0(\theta, \Pi) = 0 & \text{in } \Omega_0 \\ (\rho+\phi)\Pi - H_1(\theta, \Pi) = 0 & \text{in } \Omega_1 \\ (\rho+\phi)\Pi - \phi\pi(\theta_*) = 0 & \text{in } \{\theta_*\}, \end{cases} \tag{41}$$

where

$$H_0(\theta, \Pi') = \phi\pi(\theta) + k\min\{0, \theta\} + \max_{|\dot{\theta}|\le N} \Big\{\dot{\theta}\big(\Pi' - q(0)\big)\Big\}$$
$$H_1(\theta, \Pi') = \phi\pi(\theta) - k\max\{0, \theta - 1\} + \max_{|\dot{\theta}|\le N} \Big\{\dot{\theta}\big(\Pi' - q(\theta)\big)\Big\}$$

In $\Omega_0 \cup \Omega_1$, the definitions are just classical viscosity sub- and supersolution, which we provide next for completeness.

Definition 3 (Bardi and Capuzzo-Dolcetta (2008), Definition 1.1). *A function $u \in C(\mathbb{R})$ is a viscosity subsolution of (36) if, for any $\varphi \in C^1(\mathbb{R})$,*

$$(\rho+\phi)u(\theta_0) - H(\theta_0, \varphi'(\theta_0)) \le 0, \tag{42}$$

at any local maximum point $\theta_0 \in \mathbb{R}$ of $u - \varphi$. Similarly, $u \in C(\mathbb{R})$ is a viscosity supersolution of (36) if, for any $\varphi \in C^1(\mathbb{R})$,

$$(\rho+\phi)u(\theta_1) - H(\theta_1, \varphi'(\theta_1)) \ge 0, \tag{43}$$

at any local minimum point $\theta_1 \in \mathbb{R}$ of $u - \varphi$. Finally, u is a viscosity solution of (36) if it is simultaneously a viscosity sub- and supersolution.

Before providing the definition of sub- and supersolution on $\{\theta_*\}$, we introduce the following space $\Im$ of real valued test functions: $\varphi \in \Im$ if $\varphi \in C(\mathbb{R})$ and there exist $\varphi_0 \in C^1(\overline{\Omega}_0)$ and $\varphi_1 \in C^1(\overline{\Omega}_1)$ such that $\varphi = \varphi_0$ in $\overline{\Omega}_0$, and $\varphi = \varphi_1$ in $\overline{\Omega}_1$. Next, we introduce two Hamiltonians that are needed to define a flux-limited sub- and supersolution at $\{\theta_*\}$.

$$H_1^+(\theta_*, \Pi') \equiv \phi\pi(\theta) + \sup_{0<\dot{\theta}\le N} \Big\{\dot{\theta}\big(\Pi' - q(\theta_*)\big)\Big\}$$
$$H_0^-(\theta_*, \Pi') \equiv \phi\pi(\theta) + \sup_{0>\dot{\theta}\ge -N} \Big\{\dot{\theta}\big(\Pi' - q(0)\big)\Big\}.$$

Definition 4 (Barles et al. (2018), Definition 2.1). *An upper semi-continuous, bounded function $u : \mathbb{R} \to \mathbb{R}$ is a flux-limited subsolution on $\{\theta_*\}$ if for any test function $\varphi \in \Im$ such that $u - \varphi$ has a local maximum at*

θ_*, *we have*

$$(\rho+\phi)u(\theta_*) - \max\{\phi\pi(\theta_*), H_0^-(\theta_*, \varphi_0'(\theta_*)), H_1^+(\theta_*, \varphi_1'(\theta_*))\} \leq 0. \tag{44}$$

A lower semi-continuous, bounded function $v : \mathbb{R} \to \mathbb{R}$ is a flux-limited supersolution on $\{\theta_*\}$ if for any test function $\varphi \in \Im$ such that $u - \varphi$ has a local minimum at θ_*, we have

$$(\rho+\phi)v(\theta_*) - \max\{\phi\pi(\theta_*), H_0^-(\theta_*, \varphi_0'(\theta_*)), H_1^+(\theta_*, \varphi_1'(\theta_*))\} \geq 0. \tag{45}$$

The Hamiltonians H_0^- and H_1^+ are needed to express the optimality conditions at the discontinuity θ_*. H_1^+ consider controls that starting at θ_* take θ_t towards the interior of $[\theta_*, 1]$, and H_0^- considers controls that starting at θ_* take θ_t towards the interior of $[0, \theta_*]$. The use of the Hamiltonians H_0^- and H_1^+ at $\{\theta_*\}$, instead of H_0 and H_1, distinguishes *flux-limited* viscosity solutions from the traditional (discontinuous) viscosity solutions.

We consider the following control problem, equivalent to the one defined in (34),

$$\begin{aligned}
\tilde{\Pi}_N^*(\theta_0) &= \sup_{|\theta_t|\leq N} \int_0^\infty e^{-(\rho+\phi)t}\Big(\phi\tilde{\pi}(\theta_t) - \theta_t\tilde{q}(\theta_t)\mathbf{1}_{\{\theta_t\neq\theta_*\}} - k\big(\max\{0, \theta_t - 1\} - \min\{0, \theta_t\}\big)\Big)dt \\
\tilde{\pi}(\theta) &= \pi(0)\mathbf{1}_{\{\theta<0\}} + \pi(\theta)\mathbf{1}_{\{\theta\in[0,1]\}} + \pi(1)\mathbf{1}_{\{\theta>1\}} \\
\tilde{q}(\theta) &= q(0)\mathbf{1}_{\{\theta<0\}} + q(\theta)\mathbf{1}_{\{\theta\in[0,1]\}} + q(1)\mathbf{1}_{\{\theta>1\}}.
\end{aligned}$$

The following Theorem characterizes the value function $\tilde{\Pi}_N^*$ in terms of flux-limited viscosity solutions.

Theorem 1 (Barles et al. (2018), Theorem 2.9)**.** *The value function $\tilde{\Pi}_N^*$ is the unique flux-limited viscosity solution to equation* (41).

We can now proceed to apply Theorem 1 to verify that Π_N defined in (38) is the value function of the control problem (34).

Verification Lemmas 6 and 7 imply that Π_N is a classical solution on $\Omega \setminus \{\tilde{\theta}_N, \theta_*\}$ so we only need to verify the conditions for a viscosity solution on $\{\tilde{\theta}_N, \theta_*\}$. Π_N defined in (38) is a classical solution on $(\theta_*, 1)$. At $\theta = \theta_*$, Π_N has a convex kink so we only need to verify the supersolution property. That is, that for any $\varphi'(\theta_*)$ in the subdifferential of $\Pi_N(\theta)$ at θ_*, which is $[\Pi_N'(\theta_*-), \Pi_N'(\theta_*+)]$, inequality (45) is satisfied. $H_1^+\big(\theta_*, \varphi'(\theta_*)\big)$ is nondecreasing in $\varphi'(\theta_*)$ and $H_0^-\big(\theta_*, \varphi'(\theta_*)\big)$ is nonincreasing in $\varphi'(\theta_*)$; thus, the supersolution property follows from

$$\begin{aligned}
(\rho+\phi)\Pi_N(\theta_*) - H_1^+\big(\theta_*, \Pi_N'(\theta_*+)\big) &= (\rho+\phi)\Pi_N(\theta_*) - \phi\pi(\theta_*) = 0 \\
(\rho+\phi)\Pi_N(\theta_*) - H_0^-\big(\theta_*, \Pi_N'(\theta_*-)\big) &= (\rho+\phi)\Pi_N(\theta_*) - \phi\pi(\theta_*) = 0.
\end{aligned}$$

As $\Pi_N'(\tilde{\theta}_N-) < q(0) < \Pi_N'(\tilde{\theta}_N+)$, $\Pi_N(\theta)$ has a convex kink at $\tilde{\theta}_N$, we only need to verify the property for a supersolution. Thus, we need to verify that for any $\varphi'(\tilde{\theta}_N) \in [\Pi_N'(\tilde{\theta}_N-), \Pi_N'(\tilde{\theta}_N+)]$, inequality (43) is satisfied. This amounts to verify that

$$(\rho+\phi)\Pi_N(\tilde{\theta}_N) - \phi\pi(\tilde{\theta}_N) \geq N \max\left\{\left|\Pi_N'(\tilde{\theta}_N-) - q(0)\right|, \left|\Pi_N'(\tilde{\theta}_N+) - q(0)\right|\right\}$$

By definition of $\tilde{\theta}_N$, we have

$$(\rho+\phi)\Pi_N(\tilde{\theta}_N) - \phi\pi(\tilde{\theta}_N) = N\big(\Pi_N'(\tilde{\theta}_N+) - q(0)\big) = N\big(q(0) - \Pi_N'(\tilde{\theta}_N-)\big),$$

so it follows that

$$(\rho+\phi)\Pi_N(\tilde{\theta}_N) - \phi\pi(\tilde{\theta}_N) = N\big|\Pi_N'(\tilde{\theta}_N-) - q(0)\big| = N\big|\Pi_N'(\tilde{\theta}_N+) - q(0)\big|.$$

Finally, at $\theta = 1$, by choosing k large enough, we have that the solution of the HJB equation on $\{\theta > 1\}$ entails $\dot{\theta}(\theta) = -N$. Moreover, $\Pi'(1-) = q(1)$ implies that the value function is differentiable at $\theta = 1$ (in the extended problem) and that $\dot{\theta}(1) \leq 0$ is optimal. Thus, the state constraint is satisfied at $\theta = 1$. A similar argument applies at $\theta = 0$. Thus, we can conclude that $\Pi_N(\theta)$ is a flux-limited viscosity solution, so, by Theorem 1, it is the value function of the optimal control problem.

Proof of Corollary 2

Proof. Let us define

$$v(\theta) = F(\kappa_c)\, p_H (R - R_o) + (1 - F(\kappa_c))\, B \tag{46}$$

as the borrower's expected payoff if the asset matures. Therefore, the borrower's expected payoff at $t = 0$ is

$$V_c(\theta_0) = \int_0^\infty e^{-(\rho+\phi)t}\phi v(\theta_t)\, dt = \Phi B + \frac{\phi(b - B)}{\bar{\kappa}} \frac{\Delta R_o \theta_0}{2\rho + \phi}, \tag{47}$$

where we have substituted $R_o = R - \frac{b}{\Delta}$ and therefore $p_H (R - R_o) = b$. Finally, the borrower's problem at time 0 is

$$W_c = \max_{\{\theta_0, R_o\}} \quad V_c(\theta_0) + L_c(\theta_0) \tag{48}$$

$$s.t. \quad L_c(\theta_0) \geq I - A. \tag{49}$$

It is easily verified that, $L_c(\theta_0)$ is maximized at $\theta_0 = 1$. However, $V_c(\theta_0) + L_c(\theta_0)$ is maximized at

$$\theta_0 = 1 - \frac{\rho+\phi}{2\rho+\phi} \frac{B - b}{\Delta R_o}. \tag{50}$$

□

B Intermediation

Proof of Lemma 1

Proof. Using integration by parts and a Transversality condition $\lim_{t\to\infty} e^{-(\rho+\phi)t} D_t = 0$, we have

$$\int_0^\infty e^{-(\rho+\phi)s} dD_s = \int_0^\infty e^{-(\rho+\phi)s} (\rho+\phi) D_s ds - D_0.$$

(11) can be rewritten as

$$\max_{(D_t,\theta_t)_{t\geq 0}} \mathbb{E}\left[\int_0^\infty e^{-(\rho+\phi)s}\left[\left(\phi\hat{\pi}(\theta_s, D_s) - y\left(\theta_s, D_s\right) D_s + (\rho+\phi) D_s\right) ds + dG\left(\theta_s\right)\right]\right] - D_0$$

The optimization in the lemma follows directly from the definition of $\phi\pi(\theta)$ □

Proof of Lemma 2

Proof. The bank's optimal deposit is given by

$$D(\theta) = \arg\max_D \mathcal{V}(D,\theta) \tag{51}$$

$$\Phi\pi_i(\theta) = \max_D \mathcal{V}(D,\theta). \tag{52}$$

The first order condition follows from (25), and the second order condition is

$$\mathcal{V}_{DD}(D,\theta) \propto \Delta f'\left(\kappa_i(\theta, D)\right) D - f\left(\kappa_i(\theta, D)\right) < 0. \tag{53}$$

□

Unique $\theta_\dagger$

We show that there is a unique solution to the equation

$$\Phi\pi_i'(\theta) = q\left(0\right)\theta. \tag{54}$$

The existence proof naturally follows. The uniqueness follows if we take the *F.O.C.* of the two sides

$$\Phi\pi_i'(\theta) - q\left(0\right) = \Phi d_i(\theta) + (1-\Phi) R_o - p_L R_o = \Phi R_o\left(p\left(\theta\right) - p_L\right) + (1-\Phi) R_o\left(1 - p_L\right) > 0. \tag{55}$$

Proof of Lemma 2

Proof. We need to show that $\mathcal{B}(\tilde{D}(\theta),\theta) - p_L R_o \theta$ is monotonic in θ. Using the envelope theorem, we have that

$$\frac{d\mathcal{B}\left(\tilde{D}(\theta),\theta\right)}{d\theta} - p_L R_o = \left[\Phi\Delta^2 f(\kappa_i)\tilde{D}(\theta) + \Phi\Delta F\left(\kappa_i\right) - (1-\Phi)p_L\right]R_o.$$

If the solution to the first order condition $\tilde{D}(\theta)$ is interior (that is, less than $D(\theta)^{\dagger}$) we have

$$\Phi\Delta^2 f\left(\kappa_i(\theta,\tilde{D}(\theta))\right)\tilde{D}(\theta) = 1-\Phi,$$

substituting in the previous equation, we get

$$\frac{d\mathcal{B}(\tilde{D}(\theta),\theta}{d\theta} - p_L R_o = \left[(1-\Phi)(1-p_L) + \Phi\Delta F\left(\kappa_i\right)\right]R_o > 0.$$

□

Proof of Corollary 3

Proof. In the uniform case with $p_L = 0$ we have that $D^{\max}(\theta)$ solves

$$\frac{\Delta^2}{2\bar{\kappa}}\left[(R_o\theta)^2 - D^2\right] = D.$$

The previous equation has two roots, and the positive root is

$$D^{\max}\left(\theta\right) = -\frac{\bar{\kappa}}{\Delta^2} + \sqrt{\left(\frac{\bar{\kappa}}{\Delta^2}\right)^2 + (R_o\theta)^2}.$$

Next, let us look for conditions where

$$\tilde{D}\left(1\right) = \frac{\rho\bar{\kappa}^2}{\phi\Delta^2} \leq D^{\max}\left(1\right) = -\frac{\bar{\kappa}}{\Delta^2} + \sqrt{\left(\frac{\bar{\kappa}}{\Delta^2}\right)^2 + R_o^2}.$$

Simple derivation shows this is satisfied if and only if

$$\Phi > \sqrt{\frac{\left(\bar{\kappa}/\Delta^2\right)^2}{R_o^2 + \left(\bar{\kappa}/\Delta^2\right)^2}},$$

which holds if and only if ρ is sufficiently low. If $\tilde{D}\left(1\right) < D^{\max}\left(1\right)$, then we can plug in $D^*(1)$ and get $\pi_i\left(1\right) = \frac{\Delta^2}{2\bar{\kappa}}\left(R_o - \frac{\rho\bar{\kappa}}{\phi\Delta^2}\right)^2 + \frac{\rho + \phi\frac{\Delta^2}{\bar{\kappa}}\left(R_o - \frac{\rho\bar{\kappa}}{\phi\Delta^2}\right)}{\phi}\frac{\rho\bar{\kappa}}{\phi\Delta^2}$.

According to (4), $\kappa_i = \Delta\left(R_o - \frac{\rho\bar{\kappa}}{\phi\Delta^2}\right)$ and $p_i\left(1\right) = \frac{\kappa_i}{\bar{\kappa}}\Delta = \frac{\Delta^2}{\bar{\kappa}}\left(R_o - \frac{\rho\bar{\kappa}}{\phi\Delta^2}\right)$. Consequently, $\hat{\pi}_i\left(1, D^*\right)$ and

$\pi_i(1)$ defined in (8) and (12) become

$$\hat{\pi}_i(1, D^*) = \frac{\Delta^2}{2\bar{\kappa}}\left(R_o - \frac{\rho\bar{\kappa}}{\phi\Delta^2}\right)^2 \tag{56}$$

$$\pi_i(1) = \frac{\Delta^2}{2\bar{\kappa}}\left(R_o - \frac{\rho\bar{\kappa}}{\phi\Delta^2}\right)^2 + \frac{\rho + \phi\frac{\Delta^2}{\bar{\kappa}}\left(R_o - \frac{\rho\bar{\kappa}}{\phi\Delta^2}\right)}{\phi}\frac{\rho\bar{\kappa}}{\phi\Delta^2}. \tag{57}$$

As in the certification case, we can similarly define

$$v(1) = F(\kappa_i)\, p_H(R - R_o) + (1 - F(\kappa_i))\, B = B + \frac{\Delta\left(R_o - \frac{\rho\bar{\kappa}}{\phi\Delta^2}\right)}{\bar{\kappa}}(b - B) \tag{58}$$

$$V_i(1) = \int_0^\infty e^{-(\rho+\phi)t}\phi v(\theta_t)\, dt = \Phi v(1) \tag{59}$$

$$L_i(1) = \Pi_i(1) = \Phi\pi_i(1). \tag{60}$$

Finally, the borrower's payoff at time 0 is

$$W_i = V_i(1) + L_i(1). \tag{61}$$

□

Proof of Proposition 5

Proof. In the proposition, $L_c(\theta_0^*) > I - A$ guarantees that the borrowing constraints are slack in both certification and intermediation. Under certification

$$W_c = \Phi B + \frac{\phi(b - B)}{\bar{\kappa}}\frac{\Delta R_o}{2\rho + \phi}\theta_0 + \frac{\Phi}{\bar{\kappa}}(\Delta R_o)^2\theta_0\left(1 - \frac{\theta_0}{2}\right),$$

where θ_0 is evaluated at

$$\theta_0^* = 1 - \frac{\rho + \phi}{2\rho + \phi}\frac{B - b}{\Delta R_o} = 1 - \frac{1}{2 - \Phi}\frac{B - b}{\Delta R_o},$$

Under intermediation,

$$W_i = \Phi\left[B + \frac{\Delta\left(R_o - \frac{\rho\bar{\kappa}}{\phi\Delta^2}\right)}{\bar{\kappa}}(b - B)\right] + \frac{\Phi}{2\bar{\kappa}}\Delta^2\left(R_o - \frac{\rho\bar{\kappa}}{\phi\Delta^2}\right)^2 + \frac{\rho + \phi\frac{\Delta^2}{\bar{\kappa}}\left(R_o - \frac{\rho\bar{\kappa}}{\phi\Delta^2}\right)}{\rho + \phi}\frac{\rho\bar{\kappa}}{\phi\Delta^2}.$$

Certification dominates intermediation if $W_c > W_i$, letting $\Delta W \equiv W_c - W_i$,

$$\begin{aligned}\Delta W =&\Phi\left[\frac{1}{2-\Phi}\frac{(b-B)}{\bar{\kappa}}\left(\Delta R_o - \frac{1}{2-\Phi}(B-b)\right) + \frac{1}{2\bar{\kappa}}\left(\Delta R_o - \frac{1}{2-\Phi}(B-b)\right)\left(\Delta R_o + \frac{1}{2-\Phi}(B-b)\right)\right.\\ &\left.-\frac{(b-B)}{\bar{\kappa}}\left(\Delta R_o - \frac{\rho\bar{\kappa}}{\phi\Delta}\right) - \frac{1}{2\bar{\kappa}}\left(\Delta R_o - \frac{\rho\bar{\kappa}}{\phi\Delta}\right)^2 - \left(\frac{1-\Phi}{\Phi} + \frac{\Delta}{\bar{\kappa}}\left(\Delta R_o - \frac{\rho\bar{\kappa}}{\phi\Delta}\right)\right)\frac{\rho\bar{\kappa}}{\phi\Delta^2}\right]\\ =&\frac{\Phi}{2\bar{\kappa}}\left[\left(\frac{B-b}{2-\Phi}\right)^2 + 2\left(\frac{1-\Phi}{2-\Phi}\Delta R_o - \frac{\rho\kappa}{\phi\Delta}\right)(B-b) - \left(\frac{\rho\bar{\kappa}}{\phi\Delta}\right)^2\right]\end{aligned}$$

The previous equation is negative for B on $(\underline{B}, \bar{B})$, where

$$\underline{B} = b - (2-\Phi)^2\left[\sqrt{\left(\frac{1-\Phi}{2-\Phi}\Delta R_o + \frac{\rho\kappa}{\phi\Delta}\right)^2 + \left(\frac{\rho\bar{\kappa}}{\phi\Delta(2-\Phi)}\right)^2} - \left(\frac{1-\Phi}{2-\Phi}\Delta R_o - \frac{\rho\kappa}{\phi\Delta}\right)\right]$$
$$\bar{B} = b + (2-\Phi)^2\left[\sqrt{\left(\frac{1-\Phi}{2-\Phi}\Delta R_o - \frac{\rho\kappa}{\phi\Delta}\right)^2 + \left(\frac{\rho\bar{\kappa}}{\phi\Delta(2-\Phi)}\right)^2} - \left(\frac{1-\Phi}{2-\Phi}\Delta R_o - \frac{\rho\kappa}{\phi\Delta}\right)\right]$$

Because, $\underline{B} < b$, we only need to consider the upper bound, so, after substituting R_o, we get that the expression is negative as long as B satisfies

$$b < B < b + (2-\Phi)^2\left[\sqrt{\left(\frac{1-\Phi}{2-\Phi}(\Delta R - b) - \frac{\rho\kappa}{\phi\Delta}\right)^2 + \left(\frac{\rho\bar{\kappa}}{\phi\Delta(2-\Phi)}\right)^2} - \left(\frac{1-\Phi}{2-\Phi}(\Delta R - b) - \frac{\rho\kappa}{\phi\Delta}\right)\right].$$

Thus, we can conclude that certification dominates if

$$B > b + (2-\Phi)^2\left[\sqrt{\left(\frac{1-\Phi}{2-\Phi}(\Delta R - b) - \frac{\rho\kappa}{\phi\Delta}\right)^2 + \left(\frac{\rho\bar{\kappa}}{\phi\Delta(2-\Phi)}\right)^2} - \left(\frac{1-\Phi}{2-\Phi}(\Delta R - b) - \frac{\rho\kappa}{\phi\Delta}\right)\right].$$

We need to compare now W_c and W_d.

$$W_c - W_d = B + \frac{\phi\,(b-B)}{\bar{\kappa}}\frac{\Delta R_o}{2\rho+\phi}\theta_0 + \frac{\Phi}{\bar{\kappa}}\left(\Delta R_o\right)^2\theta_0\left(1 - \frac{\theta_0}{2}\right) - p_H R$$

□

Proof of Proposition 2 and Corollary 1

Proof. The expression for the welfare function is straightforward. The optimal retention satisfies the first order condition

$$\left(1 - e^{-\rho t}\right)\left[(1-\theta_t) f\left(\kappa_c\right)\left(\Delta R_o\right)^2 - p\left(\theta_t\right) R_o\right] + e^{-\rho t} f\left(\kappa_c\right)\left(\Delta R - B - \kappa_c\right)\Delta R_o = 0.$$

Plugging in $\theta_t = 1$, we can show that the FOC satisfies

$$\left(1 - e^{-\rho t}\right)\left[-p\left(1\right) R_o\right] + e^{-\rho t} f\left(\kappa_c\right)\left(b - B\right)\Delta R_o < 0.$$

The solution under Assumption 5 is straightforward, as well as the proof that the second-order condition is negative. $\square$

Proof of Corollary 4

Proof. For any θ_0, it is easily shown that

$$\left.\frac{\partial^2 W_\ell}{\partial t_\ell^2}\right|_{\frac{\partial W_\ell}{\partial t_\ell}=0} > 0$$

so that the local extremum is always a local minimum in the welfare function. Let the solution to $\frac{\partial W_\ell}{\partial t_\ell} = 0$ be $t_\ell^{\min}$. Finally, it is easily shown that

$$W_\ell\left(t_\ell \to \infty\right) - W_\ell\left(t_\ell = 0\right) = \left(1 - \theta_0^*\right)\left(1 - \Phi\right)\frac{\left(\Delta R_o\right)^2 \theta_0^*}{\bar{\kappa}} - \Phi\frac{\rho}{2\rho + \phi}\frac{\Delta R_o \theta_0^*}{\bar{\kappa}}\left(B - b\right) > 0$$

$\square$

Proof of Proposition 5

Proof. We can similarly define $\mathcal{V}$, the bank's objective function without trading gains as

$$\mathcal{V}(D,\theta) := \Phi\left[\hat{p}_i\left(\theta, D\right)\theta R_o - \int_0^{\kappa_i}\kappa dF\left(\kappa\right)\right] + \left(1 - \Phi\right)D + \Phi\left(1 - \hat{p}_i\left(\theta, D\right)\right)\left(1 - \xi\right)D,$$

where the new term $\Phi\left(1 - \hat{p}_i\left(\theta, D\right)\right)\left(1 - \xi\right)D$ stands for the benefit of the government subsidy.

If θ is sufficiently large such that the deposit issuance constraint is slack, simple derivation shows that the optimal deposit issuance satisfies

$$\tilde{D}\left(\theta\right) = \frac{\left(1 - \Phi\right) - \Phi\left(1 - p(\theta)\right)\left(1 - \xi\right)}{\Phi\left[1 - \left(1 - p(\theta)\right)\left(1 - \xi\right)\right] f\left(\kappa_i\right)\Delta^2}.$$

Recall that in Lemma 2, the constraint is slack whenever θ is sufficiently high.

In the region that the deposit-issuance constraint is slack, the HJB equation implies

$$\dot{\theta} = \phi\frac{R_o\xi\frac{(1-\Phi)-\Phi(1-p(\theta))(1-\xi)}{1-(1-p(\theta))(1-\xi)} - \left(1 - \Phi\right)p(\theta)R_o}{\Phi\pi_i''\left(\theta\right)}.$$

Clearly, when $\xi = 1$, we get the results in subsection that 3.2 $\dot{\theta} = \phi\frac{R_o(1-\Phi)(1-p(\theta))}{\Phi\pi_i''(\theta)} > 0$. Moreover, when $\xi = 0$ so that the entire deposit rate is subsidized by the government, $\dot{\theta} = \phi\frac{-(1-\Phi)p(\theta)R_o}{\Phi\pi_i''(\theta)} < 0$, implying the bank sells loans over time. In general, there exists a $\xi_\dagger$ and $\dot{\theta} < 0$ if $\xi < \xi_\dagger$. □

Proof of Proposition 2

Proof. Given any $\{\theta_t\}_{t\geq 0}$, the entrepreneur's payoff is $\int_0^\infty e^{-(\rho+\phi)t}\phi v(\theta_t)dt$, whereas the bank receives $\int_0^\infty e^{-(\rho+\phi)t}\phi\pi(\theta_t)dt$ and investors receive $\int_0^\infty e^{-\phi t}\phi(1-\theta_t)d(\theta_t)dt$.

$$\begin{aligned} v(\theta) &= p(\theta)(R - R_o) + (1 - F(\kappa_c))B \\ \pi(\theta) &= p(\theta) R_o \theta - \int_0^{\kappa_c} \kappa dF(\kappa) \\ d(\theta) &= p(\theta) R_o. \end{aligned}$$

From here, we get that the planners problem is

$$\begin{aligned} W &= \max_{(\theta_t)_{t\geq 0}} \int_0^\infty \phi e^{-\phi t}\Big[(1-\theta_t)d(\theta_t) + e^{-\rho t}\left(v_c(\theta) + \pi_c(\theta)\right)\Big]dt \\ &= \max_{(\theta_t)_{t\geq 0}} \int_0^\infty \phi e^{-\phi t}\left\{\left(1 - e^{-\rho t}\right)\left[(1-\theta_t)p(\theta_t)R_o\right] + e^{-\rho t}\left[p(\theta_t)R + (1 - F(\kappa_c))B - \int_0^{\kappa_c}\kappa dF(\kappa)\right]\right\}dt. \end{aligned}$$

We can maximize the previous expression pointwise to get the first order condition

$$(1-\theta_t)f(\kappa_c)\Delta R_o\left(1-e^{-\rho t}\right)\frac{\partial \kappa_c}{\partial\theta} - p(\theta_t)\left(1-e^{-\rho t}\right)R_o + e^{-\rho t}f(\kappa_c)(\Delta R - B - \kappa_c)\frac{\partial\kappa_c}{\partial\theta} = 0$$

Substituting $\kappa_c = \Delta R_o\theta$ we get

$$(1-\theta_t)f(\kappa_c)\left(1-e^{-\rho t}\right)(\Delta R_o)^2 - (p_L + F(\kappa_c)\Delta)\left(1-e^{-\rho t}\right)R_o + e^{-\rho t}f(\kappa_c)(\Delta R - B + \kappa_c)\Delta R_o = 0$$

Under Assumption 5, this simplifies to

$$\theta_t = \frac{\Delta R - b - e^{-\rho t}(B-b)}{(\Delta R - b)\left(2 - e^{-\rho t}\right)}.$$

Differentiating we get

$$\dot{\theta}_t = \frac{\rho e^{\rho t}(2B - b - \Delta R)}{(\Delta R - b)\left(2 - e^{\rho t}\right)^2}.$$

Therefore,

$$\dot{\theta}_t < 0 \Leftrightarrow \frac{\Delta R - B}{B - b} > 1.$$

At time zero we have

$$\theta_0 = \frac{\Delta R - B}{\Delta R - b},$$

and as $t \to \infty$

$$\lim_{t\to\infty} \theta_t = \frac{\Delta R - b}{2(\Delta R - b)} = \frac{1}{2}.$$

□

C Minimum Retention

Next, we construct the equilibrium in the case of certification with minimum retention. First, let's consider the price of the loan once the bank hits the minimum retention level $\underline{\theta}$. Because $\theta = \underline{\theta}$ is an absorbing state, the price of the loan satisfies $q_c(\underline{\theta}) = d_c(\underline{\theta}) = \pi_c'(\underline{\theta})$. For any $\theta > \underline{\theta}$ with smooth trading, the HJB is (15) the loan price satisfies equation (18) are unchanged. Therefore, in any region with smooth-trading, the bank's value function, the equilibrium trading strategy, and the price of loans is unchanged. It remains to check if at some θ, the bank has incentives to trade atomically to $\theta^+ = \underline{\theta}$. The payoff of jumping from θ to $\underline{\theta}$ is $\Phi\pi_c(\underline{\theta}) + q_c(\underline{\theta})(\theta - \underline{\theta})$ while the payoff in the smooth trading region is $\Pi_c(\theta) = \Phi\pi_c(\theta)$. If $\Phi\pi_c(\underline{\theta}) + q_c(\underline{\theta})(1 - \underline{\theta}) > \Phi\pi_c(1)$, the certifying bank always sells immediately to $\theta^+ = \underline{\theta}$. Otherwise, as in the proof of the case without minimum requirements, there is a unique $\tilde{\theta}$ satisfying

$$\Phi\pi_c(\tilde{\theta}) = \Phi\pi_c(\underline{\theta}) + q(\underline{\theta})(\tilde{\theta} - \underline{\theta}) = \Phi\pi_c(\underline{\theta}) + \pi_c'(\underline{\theta})(\tilde{\theta} - \underline{\theta}). \tag{62}$$

By no-arbitrage, the price function must be upper semi-continuous (that is $q(\tilde{\theta}) = q(\tilde{\theta}+)$), which means that $q(\tilde{\theta}) = \Phi\pi_c'(\tilde{\theta})$.[20] At the same time, no-arbitrage also requires to be equal to the expected dividend $\mathbb{E}\left[d_c(\theta_{\tau_\phi})|\theta_t = \tilde{\theta}\right]$. If the bank stops trading at $\tilde{\theta}$, and remains there for an exponential time with mean arrival rate λ, at which time sells $\tilde{\theta} - \underline{\theta}$, the expected dividend is

$$\mathbb{E}\left[d_c(\theta_{\tau_\phi})|\theta_t = \tilde{\theta}\right] = \frac{\lambda}{\phi + \lambda} d_c(\underline{\theta}) + \frac{\phi}{\phi + \lambda} d_c(\tilde{\theta}).$$

Combining the previous conditions, we get that λ is implicitly given bu

$$\Phi\pi_c'(\tilde{\theta}) = \frac{\lambda}{\phi + \lambda}\pi_c'(\underline{\theta}) + \frac{\phi}{\phi + \lambda}\pi_c'(\tilde{\theta}) \tag{63}$$

The next proposition, summarize the previous discussion and describe the equilibrium in the certification case with minimum retention requirements.

Proposition 6 (Equilibrium with Minimum Retention)**.** *There is a unique **Certification Equilibrium with Minimum Retention** . Given the bank's initial retention θ_0, the bank sells its loans smoothly at a rate given by equation (19) until $\tilde{T} = \min\{t > 0 : \theta_t = \tilde{\theta}\}$, were $\tilde{\theta}$ is the unique solution to equation (62). After time $\tilde{T}$, the bank holds $\theta_t = \tilde{\theta}$ until an exponentially distributed random time τ_λ, at which time it sells off $\tilde{\theta} - \underline{\theta}$. The exponential time τ_λ has a mean arrival rate $\lambda \mathbf{1}_{\{\theta_t = \tilde{\theta}\}}$, where λ satisfies (63). After time τ_λ,*

[20] Otherwise, there would be a deterministic downward jump in the price at the time θ_t reaches $\tilde{\theta}$ which would be inconsistent with no-arbitrage.

the bank holds $\underline{\theta}$ until the projects maturity. The equilibrium loan price is

$$q_c(\theta_t) = \begin{cases} \Phi\Big(p_L + F\left(\Delta R_o \theta_t\right)\Delta\Big) R_o & t < \tau_\lambda \\ \left(p_L + \frac{\phi}{\lambda+\phi} F\left(\Delta R_o \tilde{\theta}\right)\Delta\right) R_o & T_* \le t < \tau_\lambda \\ p_L + F\left(\Delta R_o \underline{\theta}\right)\Delta\Big) R_o & t \ge \tau_\lambda \end{cases} \tag{64}$$

We omit a formal proof of this proposition as it follows the step of the proof of Proposition 3.