Optimal Banking and Deposit Insurance with Delegated Monitoring

Nemanja Antic[*] and Tai-Wei Hu[†]

March 2, 2022
Preliminary

Abstract

A large number of firms with risky projects requires external funding from lenders. The realized return of the firms is private information. We adopt the costly-state-verification model by Townsend (1979) to elaborate the delegated monitoring model (Diamond, 1984) of financial intermediation, but with three novelties. First, agents cannot commit to their verification strategy. Second, lenders may be risk averse. Third, we allow aggregate uncertainty and introduce a charter system. Static bank contracts are socially beneficial without aggregate uncertainty, but they may not be so if there is significant aggregate uncertainty. The beneficial role is retained with the charter system and dynamic bank contracts, even with aggregate uncertainty. They also provide more financial stability than static ones. Two banking regulations are shown to be optimal. Regulation Q can be optimal if the regulator wants to minimize dead weight loss from verification when a monopoly bank competes with potential direct contracts. Deposit insurance can be welfare-improving if the dynamic bank contract is not financially stable.

[*]MEDS, Northwestern University, email: nemanja.antic@kellogg.northwestern.edu
[†]Corresponding author; University of Bristol, email: taiwei.hu@bristol.ac.uk

1 Introduction

Financial intermediaries played a prominent role in the Global Financial Crisis of 2007-2008 and are the central focus of many new government regulations. Financial intermediaries, such as banks, provide liquidity to depositors by making loans. In particular, bank loans are the predominant source of external funding for firms (Gorton and Winton, 2003) and, since the 1980s, the financial services sector comprises an increasing share of GDP: it accounted for around 7% of US GDP in 2012 (Philippon, 2015). To better design appropriate policy responses, it is essential to understand their role as an economic institution and to examine the effect of some long-lasting banking regulations. In this paper we study banks as intermediaries between lenders and borrowers and ask how to best regulate them. As middlemen between depositors (who seek returns on savings) and firms (who need external financing), banks exist to intermediate borrowing and lending. To explain the role of financial intermediaries some frictions are necessary, since otherwise complete contracts between the borrower and lender would achieve the first-best outcome. Freixas and Rochet (2008), among others, argue that private information is the key friction that gives financial intermediaries a useful role.

There are at least two leading explanations for the existence of financial intermediaries based on asymmetric information. Diamond and Dybvig (1983) argue that depositors have private information about their liquidity needs while entrepreneurs' projects are illiquid, and this friction leaves room for financial intermediation with a specific liability structure, namely, demand deposits under a fractional-reserve banking system. This idea, which is influential to policy-making, is the main motivation behind regulations such as deposit insurance in the literature. Another strand of literature focuses on the asset side of banks. Diamond (1984) argues that entrepreneurs know their returns as private information, and lenders may monitor them at a cost. Banks can save monitoring cost through diversification and hence increase efficiency. Based on the same idea, Williamson (1986) and Krasa and Villamil (1992) show the efficiency of financial intermediation in the context of the costly-state-verification (CSV) model initiated by Townsend (1979) and Gale and Hellwig (1985).

We follow the second strand mentioned above and develop a model of banks as intermediaries who alleviate a private information friction, which is rich enough to allow us to examine three long-lasting regulations in the banking industry: the interest-ceiling (or Regulation Q), the charter system and deposit insurance. The motivation for Regulation Q is to deter unhealthy competition between banks and encourage financial stability (see Hellman et al., 2000). In contrast to Hellman et al. (2000), in our model defaults are endogenously determined by contracts between the firms, banks and depositors. In our model, monopolistic banks have an incentive to give a higher interest rate to depositors and a higher debt level to firms. This induces more defaults by firms in equilibrium and this is socially costly because it requires more monitoring; we show that an interest rate ceiling limits this incentive and therefore reduces defaults. The second regulation, the bank charter system, makes the operation of a bank subject to supervision from a regulator. Under the charter system, banking is regarded as a privilege and the banks' behavior is subject to inspection. As such, the charter system is a dynamic scheme that can discipline bank monitoring behavior by threatening future profits. In our model, we show that this incentive scheme can improve welfare when there are aggregate shocks to firms' returns and hence to a bank's available funds. Finally, we consider deposit insurance where a regulator is both responsible for collecting the insurance premium, as well as for guaranteeing the banks' liabilities. We show how these regulations may promote the stability of the financial system when aggregate shocks are more volatile than can be handled by a simple charter system.

We introduce three novel features to the baseline CSV model to analyze the regulations. First, in contrast to Townsend (1979) and Gale and Hellwig (1985), in our model lenders (as well as banks) cannot commit to their monitoring strategy ex-ante and hence incentives for monitoring become a crucial aspect of our model. Second, in contrast to most of literature on CSV, we allow for risk aversion of lenders (or depositors). Third, we introduce aggregate uncertainty about the return to the entrepreneurs. Our main focus is to understand optimal regulations and the organization of financial intermediaries in the presence of limited commitment from the bank, and demonstrate that a charter system with dynamic contracts can improve welfare and enhance financial stability. The bank's

lack of commitment to a monitoring strategy is crucial for these results, as otherwise there is no role for dynamic contracts.

We obtain three sets of results regarding the welfare role of financial intermediaries. In order to show that financial intermediaries can provide welfare-improving contracts, we characterize the optimal direct contracts between firms and potential lenders. In our model each firm needs many lenders to finance a risky project and we show that simple debt contracts are Pareto optimal. This result generalizes the main theorem in Krasa and Villamil (2000): they assume risk neutrality and one lender, while we allow for multiple risk averse lenders. Moreover, the optimal contract we find features random priority, as in Winton (1995).[1] This result serves as a useful benchmark and provides meaningful bounds on the payoffs that agents can attain under optimal direct contracts.

Our second result shows that the bank can provide a two-sided contract that dominates the optimal direct contract without aggregate uncertainty. While Williamson (1986) and Krasa and Villamil (1992) obtain similar results when ex-ante commitment to a monitoring strategy is assumed, we extend this to a setting without such commitments. In an environment with aggregate uncertainty the fact that we need to provide the bank with ex-post incentives to monitor, requires that the bank's profit, net of the monitoring cost, be nonnegative in every state. This implies that, if the firms' aggregate returns are worse than the depositors' outside options in some states, then the bank contract cannot be financially stable in the sense that the depositors get the same return in all states.

Our third result shows that a bank charter system can be welfare-improving as it provides dynamic incentives to the bank. Under the charter system, the bank may pay a cost and self-verify its income, as well as the state of the world, to the banking authority. Depending on the result, the authority may terminate the bank's charter privilege; the authority can then use this threat to the bank's future profits as incentive for the bank to perform costly monitoring of the firms. We show that, under some mild conditions,

---

[1]It is well known that lack of commitment to the monitoring strategy is crucial for simple debt contract to be optimal. Indeed, Mookherjee and Png (1989) show that, with commitment power, it can be optimal to have stochastic monitoring and debt contract may not be optimal. See also related results in Krasa and Villamil (2000).

such a dynamic bank contract can always dominate any optimal direct contract, and, unless the bank can provide a financially stable static contract, one can find a dynamic contract that Pareto dominates the static one. Both risk aversion and monitoring cost play a role for those results. If the static contract is not financially stable, the depositors need to monitor the bank occasionally and that is costly. In contrast, in the dynamic setting, because of future profits, the bank has incentive to self-verify even if it is costly, and that is more efficient. Moreover, since under dynamic contract the bank can make a negative profit at some states, the bank can provide a deposit contract that second order stochastically dominates the static one, and this is welfare-improving for risk averse depositors. These results also show the socially beneficial role of bank profits, which are essential for the provision of dynamic incentives.

We derive two further results regarding banking regulations. In our model, the number of defaulting firms is endogenously determined by the bank contract. In general, society will face higher deadweight loss from monitoring if the bank asks for a higher debt level from the firms; competition requires higher debt levels translate into higher returns to depositors. Under the static setting without aggregate uncertainty, we show that an interest-rate ceiling, or Regulation Q, can be a useful policy to reduce the social cost of monitoring and to reduce defaults; this applies when there is a monopoly bank who faces competition from direct contracting between the firms and potential depositors.

In the dynamic setting with aggregate uncertainty, we show that deposit insurance is welfare improving if the bank contract is not financially stable. Under the deposit insurance scheme, the bank has to pay a premium in good states but may request a bail-out in bad sates. However, to request a bail-out, costly self-verification is necessary for the authority to ascertain the state and the amount of funds already available to the bank. Depending on the result, the bank charter may be terminated. In this case, the scheme needs to provide an incentive for the bank to pay the premium in good states, and to monitor firms as well as to self-verify in bad states. We show that, for sufficiently high discount factors and sufficiently low self-verification cost, a financially stable bank contract is always incentive compatible and is welfare-improving under deposit insurance.

We also consider the case with oligopolistic banks, each of which faces uncertainty, where the economy as a whole has no aggregate uncertainty, and show that the optimal deposit insurance scheme can be self-financing, ex post. Again, risk aversion is crucial for these results: deposit insurance allows for the bank to engineer a contract that second order stochastically dominates the original one, and that makes risk averse depositors strictly better off.

Our results then provide new rationales for bank charters, deposit insurance, and regulation Q based on Pareto efficiency. They show that, in an environment in which the bank cannot commit to its verification behavior, bank profits can be crucial to implement these welfare-improving banking regulations, which also help restore financial stability. In particular, our rationale for deposit insurance is orthogonal to the one provided in Diamond and Dybvig (1983), which is motivated by equilibrium selection. In contrast, in our model deposit insurance does pay out in equilibrium and is welfare-improving as such, but regulator has to overcome a moral hazard problem for banks who are bailed out by the insurance scheme.

2   Model

Consider an economy with a continuum of firms of measure one and a continuum of lenders of measure M. For simplicity, we assume that M is a natural number. Each lender is endowed with one unit of fund that may be stored at no cost with a gross return $r > 1$. Each firm has an investment project that requires M units of funds from the lenders to operate, and the project has a stochastic return $w \in [0, \bar{w}]$ distributed according to a distribution function $F : [0, \bar{w}] \to [0, 1]$. We assume that F is absolutely continuous with respect to the Lebsgue measure, that F has a full support,[2] and that $\int_w w \, \mathrm{d}F(w) > Mr$.

---

[2]Both the assumptions about absolute continuity and full support are mainly for expositional reasons than substantial. None of our results rely on absolute continuity. Full support assumption, however, guarantees that the asymmetric-information problem is substantial and avoids cumbersome discussion of uninteresting cases.

The firms are risk-neutral and short-lived, but the lenders are long-lived and may be risk averse. The instantaneous utility function of the lender for consuming c is $u(c)$, which is strictly increasing and concave, and the lender has discounted factor $\beta \in (0,1)$. The return w_n to firm n is private information to the firm. The lenders, however, may verify the return by paying a cost γ, which is additive to his utilities of consumption. The verified return is private information to the verifier that cannot be credibly shared with other lenders.

2.1 Optimal direct contracts

As a benchmark, we consider direct contracts where one firm borrows directly from M lenders. Since firms last for just one period, only static contracts are feasible. We characterize Pareto optimal contracts. A contract is a pair, (a, b), both with domain $W = [0, \bar{w}]$, and is implemented with the following trading procedure. After the return w realizes, the firm makes a report to each of his lenders. Let $\tilde{w}_m$ be the report to lender m. Each lender m then decides to verify or not, based on $\tilde{w}_m$. If he decides not to verify, lender m receives a payment of $b(\tilde{w}_m)$. Otherwise, he pays a cost γ, and receives a payment of $a(v, M')$, where v is the amount of funds available *after* paying those who did not verify and M' is the number of verifiers. As mentioned, both the message $\tilde{w}_m$ and the verified output v are private information to the specific lender that cannot be (credibly) shared with other lenders.

A strategy for a lender is then a measurable function $s_l : W \to \{0, 1\}$, where 0 indicates no verification and 1 indicates verification. A strategy for the firm is a measurable function $s_f : W \to \Delta(W^M)$, his reporting strategy. Note that the firm is allowed to randomize over possible massages to the lenders. We have the following feasibility requirements the contract and the strategies:

(F1) Feasibility of a: $a(v, k)k \leq v$ for all $v \in W$.

(F2) Feasibility of s_f under b: $\sum_{m=1}^{M} b(\tilde{w}_m) \leq w$ for all $(\tilde{w}_1, ...\tilde{w}_M)$ in the support of $s_f(w)$ and for all $w \in W$.

We focus on PBE that are symmetric in the following sense: first, all lenders use the same strategy s_l, and, for the firm, s_f has a finite support,[3] and, if a profile of message $(\tilde{w}_1, ..., \tilde{w}_M)$ is in the support of the equilibrium strategy s_f, then all permutations of that message occur with equal probabilities.

We say that a direct contract (a, b) is **implementable** if the lender is willing to participate. We let W_m be the subset of W in which m lenders verify the firm with positive probability. Note that if, in equilibrium, upon receiving a message $\tilde{w}$ the lender does not verify the firm's return, the firm simply pays $b(\tilde{w})$, it must be the case that the payment is the same under b across all such messages that are sent with positive probability, which we will denote by $\bar{b}$. Hence, we may assume that there is only one message corresponding to the subset of W where nobody verifies the firm, W_0. Furthermore, because lenders cannot commit to their verification strategies upon signing the contract, they need to be incentivised to verify after the report from the firm comes in. This requires that γ is smaller than the expected payment when receiving a message that implies verification in equilibrium. Formally, let $\tilde{w}$ be a message that in equilibrium the lender verifies, it must be the case that

$$\mathbb{E}\left[a(w, w - (M-m)\bar{b}, m) \mid \tilde{w}\right] \geq \gamma,$$

where within the conditional expectation both w and m are random variables. Since this condition has to hold for all messages $\tilde{w}$ which are sent with a positive probability and for which the lender verifies in equilibrium, it implies that

$$\mathbb{E}\left[a(w, w - (M-m)\bar{b}, m) \mid \text{ verification}\right] \geq \gamma.$$

As a result, it is then without loss of generality to assume that in equilibrium the firm only sends two messages, $\tilde{w}_0$ or $\tilde{w}_1$, where $\tilde{w}_0$ indicates no verification while $\tilde{w}_1$ indicates verification.

We make one assumption regarding the cost γ:

(A0) $\gamma F(Mr) < \int_0^r u\left(\frac{w}{M}\right) \mathrm{d}F(w)$.

When $M = 1$, assumption (A0) is essentially identical to assumption (A.2) in Krasa

---

[3]This assumption is mainly for technical convenience than substantial.

and Villamil (2000), and hence it is a generalization of that assumption for general M and it plays a similar here. In particular, note that we only consider equilibria where the lenders use a pure strategy, which can be justified along the renegotiation-proof requirement introduced by Krasa and Villamil (2000) under (A0).[4]

The following theorem characterizes (Pareto) optimal contracts, which generalizes the main result in Krasa and Villamil (2000) to multiple lenders and risk aversion.

Theorem 2.1. *Assume (A0). Any optimal contract, $(W_0, ..., W_M, a, \bar{b})$, takes the form $W_0 = [M\bar{b}, \bar{w}]$, and $a(w, n) = w/n$.*

According to Theorem 2.1, the optimal contract is a simple debt contract from the firm's perspective: it is characterized by a debt level $M\bar{b}$, and if the return is above that the firm pays that to all lenders equally outright; otherwise, at least some of the lenders verify the firm but all the returns are divided by the lenders. What is slightly delicate is the structure of verification. In general there is random seniority as in Winton (1995): some lenders are chosen to be paid outright while others are called to verify the firm to divide whatever is left. From the perspective of risk sharing, it is best for all lenders to share the return whenever the firm cannot meet its obligations to all lenders, but this is very costly as all lenders have to pay the verification cost. From the perspective of saving verification cost, it is optimal for the firm to pay the debt level to as many lenders as possible, and ask as few lenders to verify as possible. Indeed, if $u(c) = c$, this will be the optimal contract. When there is risk aversion, we can in fact fully characterize the optimal number of lenders asked to verify. The details can be found in the Supplemental Appendix, A.1. Finally, although Theorem 2.1 assumes (A0), what is needed for that theorem to hold may be much weaker than (A0), depending on the fundamentals. See also the Appendix for more discussions.

Theorem 2.1 will be useful for our later analysis as it gives welfare bounds to both the

---

[4]Krasa and Villamil (2000) assume risk neutrality and that the firm also occurs a small cost when it is verified, and show that equilibria with random verification is not sustainable under (A0) with $M = 1$. Our result will not be affected by the introduction of a small cost to the firm and the logic to exclude random verification in Krasa and Villamil (2000) does not seem to depend on M nor risk neutrality.

firms and to the lenders under direct contracting. Indeed, under optimal direct contracts with debt level $B = M\bar{b}$, the firm's payoff is

$$\int_{w \geq B} (w - B) \, \mathrm{d}F(w).$$

Each lender's payoff is bounded from above by

$$\int_{w < B} u(w/M) \, \mathrm{d}F(w) + u(B/M)[1 - F(B)] - (1/M) \sum_{n=1}^{M} F(nB/M)\gamma. \qquad (1)$$

The first two terms reflect the expected payoff to each lender, assuming that they share the returns whenever the firm fails to pay all lenders the debt level, which is the upper bound for that payoff, the third term reflects the cost of monitoring, assuming the fewest number of lenders monitoring and all have equal chance to do so, which is a lower bound of the monitoring cost.

3 Banking: static and dynamic

Here we show the beneficial role of banks. We begin with the static case where there is no aggregate shocks, and the returns to the firms are i.i.d. In this case we show that a static bank contract is sufficient to bring in Pareto improvement against direct contracting. We then introduce aggregate shocks under which the firms' returns are correlated, and show that these shocks can limit the bank's ability to improve direct contracts. Finally, a charter system is introduced and is shown to bring back the beneficial role of the bank.

3.1 Essentiality of financial intermediation: no aggregate shocks

We begin with the static case, where, as in Section 2, there is a continuum of depositors of measure M and a continuum of firms of measure 1 with i.i.d. returns. We assume that the LLN holds exactly when we aggregate the firms' returns and hence there is no aggregate shock.[5] Here we introduce the bank, which is a risk-neutral agent capable of

---

[5]The literature has recognized issues related to i.i.d. returns in a continuum economy; see, e.g. Sun (2006). Our use of this assumption, however, is mainly for expositional convenience. All our results can

taking deposits and then lending to the firms. When the bank decide to verify a firm, he has to pay the cost c_E. The cost of monitoring for the bank is not subtracted from its funds taken from the firms; instead, we interpret c_E as labor cost to the banker.[6] Again, the verified returns are private information to the bank and cannot be credibly shared with the depositors.

Since debt contracts are optimal without financial intermediation, to show that banks are *essential*, we only need to consider such contracts. We call a two-sided contract between the depositors and the bank and between the firms and the bank, a *bank contract*. Assuming that it is a debt contract for both sides, we may denote it by (B, d), where B is the debt level with each firm and d is the promised repayment to each depositor. Later on we will establish that such contracts are in fact optimal. One of the main insights from Section 2 is that it is optimal to restrict the message space to $\{0, 1\}$, where 0 indicates repayment of the debt and 1 indicates verification, and, if the lender, which can be either the bank or the depositor, fails to obey (fails to verify) does not happen, he or she gets nothing. Thus, from here on we restrict attention to the message space $\{0, 1\}$. The two-sided nature of bank contracts introduces new incentive issues that are not present in direct contracting. To formalize those issues we give a precise description of the game as follows. The game is played among the firms, the bank, and the depositors. We assume that all agents have agreed upon the contract (B, d) and discuss agents' participation decisions later.

1. Firms' returns are realized and firms simultaneously report (either 0 or 1) to the bank.

2. The bank decides, based on the reports, whether or not to verify each firm.

3. After the bank receives all the payments from the firms, the bank sends a massage (again, either 0 or 1) to each depositor simultaneously.

be generalized to a setting with a finite but large number of depositors and firms, as will be discussed in concluding remarks.

[6]Note that the cost is c_E per firm. Obviously the bank here will correspond to a large entity in reality with potential monitoring issues within itself; here we abstract away from those complications.

4. After seeing the individual message, each depositor decides whether or not to verify the bank.

Given a bank contract (B, d), we formulate the agents' strategies below. The strategies for the firms and for the depositors are straightforward extensions from the ones under direct contracting. Each firm decides, as a function of its realized profit, to either repay the debt in full or request verification. The bank's strategy comes in two stages. The first is a verification strategy for each firm, as a function of all the information it has available, i.e., the bank knows which firms have repaid the debt and which requested verification. Second, the bank makes a recommendation to each of its depositors, asking them to verify it or repaying the depositors in full. Each depositor's strategy is then a function which takes the bank's recommendation and maps it to a binary decision (to verify or not). The following is the formal definitions.

Definition 3.1. (a) For the firm, the strategy is a measurable function $s_f : [0, \bar{w}] \to \{0, 1\}$ that maps its return to its report to the bank.

(b) The bank's strategy has two components. The first is a measurable function $s_B^1 : \mathcal{M}([0, 1]) \times [0, 1] \to \Delta(\{0, 1\})$. The first argument is the subset of firms sending 1, and the second argument is the identity of the firm.[7] Thus, for any given received reports, s_B^1 specifies a (randomized) verification decision for each firm. Then, after the verifications, the bank's strategy has a second component, a measurable function $s_B^2 : [0, \bar{w}] \times [0, M] \to \Delta(\{0, 1\})$, in which the first argument is the amount of available funds to the bank (in per depositor terms) and the second is the identity of the depositor, and hence it maps the available funds to the bank's (randomized) report to each depositor.[8]

(c) A depositor's strategy is a function $s_l : \{0, 1\} \to \{0, 1\}$, which maps the report from the bank to the verification strategy.

As before, symmetric PBE is our solution concept, i.e., all firms use the same strategy s_f and all depositors use the same strategy s_l in equilibrium, and, for the bank, although

---

[7] $\mathcal{M}([0, 1])$ is the set of all measurable subsets of $[0, 1]$ (endowed with the sup norm).

[8] In general, s_2^B may also depend on the history that leads up to the available funds.

we allow randomization, we only focus on equilibrium strategies in which s_B^1 verifies all firms that cannot pay B.[9]

Here we discuss participation decisions by our agents. We say that a bank contract (B, d) is **implementable** if there is a symmetric PBE in which the depositors are willing to participate and the bank is willing to verify the firms' returns whenever they cannot meet their debt, and the bank's expected payoff (net of monitoring costs) is nonnegative.

In contrast to Krasa and Vilamil (1992), because of the lack of commitment, we need to discuss the bank' incentive to verify firms after the returns realize and requests come in. In particular, there are two types of deviation to worry about. First, the bank can simply verify none of the requesting firms; no matter how the depositors react, the bank can make at least a zero profit. Second, the bank has to decide whether to verify each requesting firm, and it will do so unless the expected gain (net of what is expected to pay to the depositors at the margin) from doing so exceeds the cost. Under the contract B, $F(B)$ fraction of the firms require verification. We have the following lemma.

Lemma 3.1. *Let (B, d) be a bank contract. Bank verification of the firms is incentive compatible if and only if*

$$\frac{1}{M}\left[\int_0^B w\mathrm{d}F(w) + [1 - F(B)]B\right] - d \geq \frac{1}{M}F(B)c_E, \tag{2}$$

$$\int_0^B w\mathrm{d}F(w) \geq F(B)c_E. \tag{3}$$

To understand the two conditions above, first note that under the bank contract (B, d), the bank's expected profit (per depositor per firm) is given by

$$\frac{1}{M}\left[\int_0^B w\,\mathrm{d}F(w) + [1 - F(B)]B - F(B)c_E\right] - d. \tag{4}$$

Hence, condition (2) simply says that the bank is making a nonnegative profit and this takes care the first type of deviation. Conditional on the firm sending a message that recommends verification under contract (B, d), the bank's expected payment to be received

---

[9]This assumption is in line with the observation from Krasa and Villamil (2000). In particular, we suspect that if we introduce renegotiation then randomized verification is not sustainable in equilibrium.

is given by

$$\frac{\int_0^B w \, \mathrm{d}F(w)}{F(B)},$$

and hence condition (3) requires that conditional expected return to at least cover the cost of verification, c_E, and this takes care of the second type of deviation. The proof essentially shows that these two types are the only relevant deviations.

For bank contract to be more efficient, we need the following assumptions.

(A1) Efficient bank verification: $c_E < u^{-1}(\gamma) \sum_{m=1}^{M} \frac{F\left(\frac{mB}{M}\right)}{F(B)}$.

(A2) $\int_{w \in [0,B]} w \mathrm{d}F(w) + (1 - F(B))B - F(B)c_E > Mr$.

Assumption (A1) is necessary: if c_E is large relative to γ, then the bank contract cannot save on monitoring costs. However, the precise comparison is related to the curvature of u, since γ enters the depositors' payoff in comparison with u. M also plays a role, as the bank monitor each firm for M depositors. Without (A2), bank is better off to use the storage. The following theorem shows that these two assumptions are all we need to have a beneficial role for the bank.

Theorem 3.1. *Assume (A0). For any implementable direct contract with debt level B that satisfies (A1) and (A2), there is an implementable (static) bank contract that Pareto dominates the direct contract.*

Theorem 3.1 is proved by constructing a bank contract that Pareto dominates the direct contract. The bank contract is given by (B, d), where the debt level is left unchanged, and d is given by

$$d = \frac{1}{M} \left\{ \int_0^B w \mathrm{d}F(w) + [1 - F(B)]B - F(B)c_E - \varepsilon \right\}, \tag{5}$$

where the main proof is to show that one can choose $\varepsilon > 0$ such that

$$u(d) > \int_{w<B} u(w/M) \, \mathrm{d}F(w) + u(B/M)[1 - F(B)] - (1/M) \sum_{n=1}^{M} F(nB/M)\gamma.$$

Indeed, if such $\varepsilon > 0$ exists, then, by (4), the bank has profit at least ε, and, by (1), the depositors are better off. To show such ε exists, first notice that $d + \frac{1}{M} F(B)c_E + \varepsilon$ is the

14

expected payment from the firm to each lender, and hence its utility value is higher than the expected value of that payment. Assumption (A1) then ensures that the monitoring cost is lower under the bank contract as well. This then illustrates the advantage of the bank contract through diversification: first, it can reduce risk by pooling the funds, and, secondly, it can reduce monitoring, as the bank can offer a deposit contract that the bank can almost surely pays to the depositors and hence only monitoring of the bank on behalf of M depositors per firm is required.

Assumption (A0) ensures that the optimal direct contracts are simple debt contracts and the lenders' payoffs are bounded by (1). The assumptions (A1) and (A2) are tight for Theorem 3.1. We have seen the necessity of (A2), which is in fact implied by implementability of the direct contract under risk neutrality. When u is linear, (A1) is also necessary for the bank to be able to provide a better contract. When u is strictly concave, however, (A1) can be relaxed, since, as mentioned earlier, the constructed bank contract is strictly better because it is able to provide the certainty equivalence of what the direct contract can provide to the lender.

3.2   Aggregate shock and financial stability

Now we consider aggregate shocks. Suppose that, instead of independence, the returns to the firms are correlated. Specifically, suppose that the aggregate returns depend on an aggregate state, s, which can be either h (high) or ℓ (low), that s is i.i.d. across time according to the distribution π (π_h denotes the probability of state h and π_ℓ state ℓ), and that the firms' returns are i.i.d. according to $F_s(w)$ conditional on s. Moreover, F_h first-order stochastically dominates F_ℓ and that for any $B \in (r, \bar{w})$, $F_h(B) < F_\ell(B)$. The realization of the state is observable to both the bank and the firms, and is a verifiable information for the depositors if they monitor the bank.[10] We remark here that the results regarding direct contracting are not affected by this aggregate shock at all: for

---

[10]Note that even if the bank does not receive this information, it can learn of the state s based on the reports made from the firms. This assumption amounts to eliminate the potential asymmetric information between the bank and the firms about the state.

each specific firm, its return is characterized by the distribution function $F = \pi_h F_h + \pi_\ell F_\ell$.

A general bank contract can be quite complicated. In particular, the contract should in general take into account the fraction of depositors that verify the bank; this complication then requires a new proof to show that debt contract is optimal with the firms. We show in the Supplemental Appendix (see, in particular, Lemma A.4) that, under the following assumption, debt contract with the firms is in general optimal.

(A3) $\int_0^{Mr} w dF_s(w) > c_E F_s(Mr)$ for both $s = h, \ell$.

Assumption (A3) plays a similar role to (A0) in direct contracting for this result. In particular, it implies that, in equilibrium, each firm sending a message 1 is worth of verification if the bank can keep the verified return. However, depending the contract with the depositors, this may or may not be the case and additional considerations about bank's incentive to verify the firms are necessary.

Given Lemma A.4, we can focus on bank contracts in which the contract between the bank and the firms is a simple debt contract (that may be state-dependent), and the bank verifies the firms iff they do not pay in full. Such a contract may be denoted $\mathbf{B} = (B_h, B_\ell)$. Given $\mathbf{B}$, we define

$$\eta_s(\mathbf{B}) = \frac{1}{M} \left\{ \int_0^{B_s} w dF_s(w) - F_s(B_s) c_E + [1 - F_s(B_s)] B_s \right\}, \qquad (6)$$

and

$$\zeta_s(\mathbf{B}) = \frac{1}{M} \left\{ \int_0^{B_s} w dF_s(w) + [1 - F_s(B_s)] B_s \right\}. \qquad (7)$$

Here $\eta_s(\mathbf{B})$ is the revenue (net of monitoring cost) from the firms to the bank and $\zeta_s(\mathbf{B})$ is the available funds at the bank at state s in equilibrium, both in per depositor's term.

While Lemma A.4 allows us to focus on debt contracts with firms, the depositor side of the bank contract is more complicated. We are able to obtain a partial characterization, and report the results in Lemma A.2 in the Appendix. In particular, because of the continuum assumption, we may characterize the depositor side of the contract by three numbers, (d, d_h, d_ℓ), where d is the promised payment without verification to the depositors, and d_s is the *average* payment to the depositors under state s, for $s = h, \ell$. From this information one can then infer the fraction of verifying depositors at each state

16

by symmetry. Thus, to characterize its outcome, any optimal contract can then be represented by $(\mathbf{B}, \mathbf{d})$, where $\mathbf{d} = (d, d_h, d_\ell)$. For our welfare consideration, however, a specific kind of contract, if implementable, will be Pareto efficient, in which $d = d_h = d_\ell$.

Definition 3.2. A bank contract, $(\mathbf{B}, \mathbf{d})$, is **financially stable** if there is a PBE in which the bank's expected payoff is nonnegative, and all depositors obtain $u(d) \geq u(r)$ with probability one in equilibrium.

In contrast to our results in Section 3.1, static contracts have limited ability to implement financially stable contracts under aggregate uncertainty. Our next theorem gives a sufficient condition for this to happen, but the key incentive problem is reported in Lemma A.2. Lemma A.2 gives a key necessary condition for implementation: $\eta_s(\mathbf{B}) \geq d_s$ for both $s = h, \ell$, that is, the bank has to make a nonnegative *ex post* profit in both states. This result is a direct consequence of the fact that the bank cannot commit to its verification strategy; indeed, if the bank were to make a negative profit at some state, it may well simply verify no firms, which would give at least a zero profit. This also implies that the contract with the depositors may not be a simple debt contract; it is possible that upon verification, the bank may still keeps some of its available funds. Indeed, when $d_\ell < d_h$, it is necessary that a fraction of depositors have to verify the bank at state ℓ, but $d_\ell \leq \eta_\ell(\mathbf{B})$ implies that the bank has to keep a fraction of its available funds, $\zeta_\ell(\mathbf{B})$.[11] As a direct corollary of Lemma A.2, the following theorem gives a sufficient condition for this to be the case.

Theorem 3.2. *Let* $(\mathbf{B}, \mathbf{d})$ *with* $d_\ell \leq d_h$ *be an optimal implementable bank contract. If* $B_\ell \leq r$, *then it is not financially stable.*

Compared to financially stable contracts, nonstable contracts can be costly to the depositors for two reasons. Consider such a contract with $d_\ell < d_h$. First, since incentive compatibility requires the bank to make a nonnegative profit, at state ℓ the available fund to the bank, $\zeta_\ell(\mathbf{B})$, is strictly higher than $\eta_\ell(\mathbf{B})$, which is in turn weakly higher

---

[11]Lemma A.2 only focuses on the case where $d_\ell \leq d_h$. This is with no loss of generality as we do not make use of the assumption that F_h first-order-stochastically dominates F_ℓ in the proof.

than d_ℓ. Thus, for risk-averse depositors, if one could ask the bank to pay out more in state ℓ and pay less in state h to keep its ex ante profit constant, this would be a Pareto improvement. Second, a low d_ℓ also entails that a larger fraction of depositors have to verify the bank at state ℓ, and that is costly. Under static contracts, however, this may be the best thing to do for a given division of surplus between the firms and the depositors. These considerations also imply that the static bank contract may not be able to improve upon the direct contract under aggregate uncertainty. In next section we show that both inefficiencies can be reduced by a dynamic contract, and, under mild conditions, one can device a dynamic bank contract that Pareto dominates the direct contract, even under aggregate uncertainty.

3.3   Dynamic contracts

In this section we introduce dynamic bank contracts. Specifically, the dynamic nature is implemented by a bank charter system with a regulator. Banking charter is a privilege in the sense that once removed, the bank can no longer operate and can no longer receive future profits. Under this system the regulator can implement dynamic contracts with the bank through threats of removing the charter privilege, and by requiring self-verification from the bank. As seen in the last section, under a static contract incentive compatibility requires the bank to make a nonnegative profit state by state, and this requirement limits the bank's ability to offer a financially stable contract. In contrast, a dynamic contract can overcome this incentive issue and improve social welfare.

We first describe the charter system. Under this system, the bank has to be chartered by a regulator; the main benefit from this is that the regulator can provide dynamic incentives to the bank. We assume that the bank is long-lived and has discount factor β. The fact that this discount factor coincides with the depositors' discount factor play no role in our analysis, but simplifies notation.

The regulator has the power to terminate the bank's charter and hence uses future profits as a threat to implement contracts.[12] However, the regulator does not observe the

---

[12]The key differences between our work and the large literature on dynamic principal-agent problems

funds collected by the bank, nor the realized state, s, but has to rely on self-reporting from the bank. We assume a technology with which the bank can self-verify and make both the state and available funds in the bank credibly known to the regulator at a per depositor cost of c_B.[13] This self-verification can be done by hiring an external auditor, for example.

The charter system implemented by the regulator operates as follows. If the bank fails to pay the deposits in full and fails to self-verify, its charter is automatically terminated: in this case, the depositors may verify the bank and receive whatever funds are left. Otherwise, the termination is determined by the following dynamic contract. First the contract specifies a level of promise to pay to the depositors, d. If the bank does not pay d and engages in self-verification, the contract then specifies a termination policy, $T(s, y) \in \{0, 1\}$ (here 0 indicates termination), and a payment to each depositor, $D(s, y) \in \mathbb{R}_+$, where s is the state and y is the available funds per depositor revealed through verification. After termination, the bank's continuation payoff is zero and hence, if designed appropriately, the termination policy can provide the bank with an incentive to verify the firms. Moreover, these dynamic considerations can sometimes make financially stable contracts feasible in equilibrium, even though a static contract cannot.

Given the dynamic contract, the sequence of actions needs some modification. In particular, items 5 and 6 from Section 3.1 require a change. The timing in each period is as follows (after all parties agree to participate):

1. After the firms' returns realize, all firms simultaneously make reports (either 0 or 1) to the bank.

(e.g., Thomas and Worrall, 1990) is that we have to attend to the agent's (the bank, in our case) incentive to pay third-parties (depositors), which the principal (regulator) also cares about. So, while we use future utility promises to incentivise truthful reporting of private information, our construction is quite different to those papers; in particular, the agent (bank) will necessarily need to make a profit (in order to ensure it has an incentive to pay depositors) and thus the agent cannot be held down to his participation constraint, as in Thomas and Worrall (1990).

[13]While we assume a fixed cost per depositor, our analysis goes through with a cost which varies with the size of the bank's balance sheet.

2. The bank decides, based on the reports from the firms, whether or not to verify each firm.

3. After the bank receives all the payments from the firms it can take one of the following actions:

   (a) pay all depositors d (which may not be feasible);

   (b) engage in self-verification—and is thus subject to termination policy, $T(s, y)$.

   (c) do nothing—in which case the bank's charter is automatically terminated.

4. If the bank fails to pay all depositors d and fails to sell-verify, each depositor decides whether or not to verify the bank.

Given the dynamic nature and the modified timing of the game, the agents' strategies can be defined in an analogous manner to those in Definition 3.1. We emphasize a few key differences. First, strategies have to be indexed by time and would depend on the observed histories from previous periods as well as the state s. Furthermore, s_B^2, the banks reporting strategy has to be amended to allow for the possibilities in item 5 above (we'll denote paying d by 0, self-verification by 1, and doing nothing by 2). Finally, s_ℓ, the strategy of the depositor, chooses whether to verify the bank or not only when the bank decided to do nothing. Note that while all this seems complicated, we will see that when discount rates are high and the cost of self-verification is low, the depositors have a relatively simple optimal strategy—they will decide to verify if given the opportunity.

Note that the dynamic contract is more efficient than the static ones only if the cost c_B is small, and that will be the case we focus on. Moreover, when c_B is small, the optimal dynamic contract will use depositors' verification only as an off-equilibrium threat. In that sense, under the dynamic contract, our focus is on the interaction between the bank against the firms and against the regulator.

As before, we focus on symmetric equilibria in which the interactions between the bank and the firms in a PBE with the property that the bank verifies all firms that send message 1, and we say that a dynamic contract is **implementable** if there is such a PBE

in which the depositors are willing to participate and the bank is making a nonnegative (expected) profit at the beginning of each period.[14] The following theorem shows that the dynamic bank contracts can be welfare improving.

Theorem 3.3. *Assume that u is strictly concave and that c_E satisfies (A3).*

(a) Assume (A0). For any implementable direct contract with debt level B that satisfies (A1) and (A2) for $F = \sum_{s \in S} \pi(s) F_s$, there is an implementable dynamic bank contract that Pareto dominates the direct contract for sufficiently high β and sufficiently low c_B.

(b) Let $(\mathbf{B}, \mathbf{d})$ be an implementable static bank contract that is not financially stable. Then, for sufficiently high β and sufficiently low c_B, there is an implementable dynamic bank contract that Pareto dominates it.

The main ingredient of the proof, which is also the main contribution of the above theorem, is in the design of a dynamic contract which allows for the usual repeated-games arguments and, thus, a Pareto improvement on the original contract. While the details differ for parts (a) and (b), the main ideas can be outlined as follows. We use expected future profits to incentivise the bank to suffer a short-term loss, which allow the contract to increase the depositors' returns at state ℓ in the expense of returns at state h; when designed properly, this is welfare-improving given that depositors are risk averse. Once we obtain a contract that both improves depositors' welfare as well as gives the bank a positive profit, we can apply the usual repeated-game argument that a sufficiently high β and a sufficiently low c_B implies that the dynamic incentives are powerful enough to discipline the bank. These requirements are indispensable; for low β's, the dynamic incentives have no bite, and for high c_B's, the self-monitoring technology of the bank is not sufficiently efficient to be useful.

According to Theorem 3.3 (a), one can always devise a dynamic contract to dominate the direct contract, even under aggregate uncertainty, as long as (A1), (A2), and (A3) are satisfied. This result does not hold for static contracts; indeed, for F_ℓ sufficiently concentrated on returns that are close to zero, the depositors have no incentive to monitor

---

[14]Hence, implicitly we allow the bank to leave the charter system at any point.

the bank at state ℓ. Instead, in the dynamic bank contract, the bank is motivated to self-verify because of concerns about future profits.

Part (b) of Theorem 3.3 shows that a dynamic bank contract can improve upon any implementable static bank contract that is not financially stable. Recall that under the static contract, the bank's incentive to verify defaulting firms requires that the bank's profit is nonnegative across states, and hence $d_\ell \leq \eta_\ell(\mathbf{B}) < \zeta_\ell(\mathbf{B})$, where d_ℓ is the average payment to depositors, $\eta_\ell(\mathbf{B})$ is the revenue to the bank (net of monitoring cost) and $\zeta_\ell(\mathbf{B})$ are the funds available to the bank in state ℓ. Given this condition, we construct a dynamic contract in which the bank takes a short-term loss at state ℓ and pays $\zeta_\ell(\mathbf{B})$ to each depositor. There is a loss at state ℓ because of the monitoring cost. This dynamic contract is incentive compatible because of future profits: we can decrease d_h, the payment to depositors in state h, so that the bank is making a positive expected profit. This change improves the welfare of the depositors as they are strictly risk averse and prefer the higher payment in the low state.

We remark here that while bank profit is crucial to use the dynamic incentives, it also limits the benefits to the depositors. In fact, for any given static contract, there is a maximal (average) profit that can give to the bank to ensure that the depositor is better off, which depends the depositors' risk aversion. Given the profit, there is then a cut-off discount factor that makes the dynamic contract incentive feasible.

Finally, although the charter system can provide better contracts, it may not guarantee financial stability as defined earlier. In particular, the ability to increase d_ℓ depends on both the available funds $\zeta_\ell(\mathbf{B})$ and the incentive compatibility condition for the bank to suffer losses. In the next section we show that the first issue can be solved by the introduction of a deposit insurance scheme.

4   Applications to banking policies

4.1   Interest rate ceiling

Here we consider a monopoly bank and its pricing decision. Last section shows that the bank has an advantage in providing a better contract by being large and hence it seems reasonable to assume that the bank enjoys some monopoly power. The firms, however, may still issue their debts directly to the depositors without going through the bank, and, hence, the bank contract has to be sufficiently attractive in the sense that there cannot be any direct contract that Pareto dominates the bank contract from the perspective of the depositors and the firms. This motivates the following definition.

Definition 4.1. A bank contract, (B, d), is **competitive** if there is no direct contract that Pareto dominates it.

We are able to characterize the set of competitive contracts for risk neutral depositors.

Lemma 4.1. *Suppose that $u(c) = c$. Bank contract (B, d) is stable iff*

$$d \geq \max \left\{ r, \int_{w<B} \frac{w}{M} \, \mathrm{d}F(w) + [1 - F(B)] \frac{B}{M} - \gamma \frac{1}{M} \sum_{n=1}^{M} F\left(\frac{nB}{M}\right) \right\}. \qquad (8)$$

The bank's problem is then to choose a contract that maximizes its profit subject to being competitive. Assume (A0) and (A1), and hence $c_E \leq \gamma$, the contract between the bank and the firm is such that the bank verifies iff $w < B$. Then, given Lemma 4.1, bank profit under B, subject to being competitive, is given by

$$\gamma \frac{1}{M} \sum_{n=1}^{M-1} F\left(\frac{nB}{M}\right) \qquad \text{if } U^{direct}(B) \geq r$$

$$\int_{0}^{B} \frac{w}{M} \mathrm{d}F(w) + (1 - F(B))(B/M) - F(B)c_E/M - r \qquad \text{otherwise,}$$

where $U^{direct}(B)$ is the expected payoff to the lenders under the direct contract B.

Let B^* maximize $\int_{0}^{B} \frac{w}{M}\mathrm{d}F(w) + (1 - F(B))(B/M) - F(B)c_E/M$. The bank's profit is maximized either at $B = B^*$, or at $B = \bar{w}$. Suppose that we are in the latter case.

Without interventions, the bank contract induces the maximum monitoring cost to the society, which may not be desirable. One way to encounter this is to introduce interest ceiling, or regulation Q, that limits the amount of interest rate that the bank can give to the depositors, and, to be competitive, it also requires B to be lower than the optimal contract under the monopoly bank. Such interventions will then reduce the monitoring cost for the society.

4.2 Deposit insurance

Here we introduce deposit insurance. We assume that there is a monopoly bank. The insurance scheme sets an amount τ of transfers from the regulator to the depositors (in terms of per depositor) and an amount ϕ of premium paid by the bank to the regulator (in terms of per depositor). Consistent with the previous section, to claim for the transfer, the bank has to pay a cost c_B (in terms of per depositor) to self-verify to make its available funds and the state known to the regulator. The transfer τ may depend on the funds y and the state s. Alternatively, the bank may simply pay the premium. If the bank fails to pay the premium and fails to self-verify at the same time, the bank is declared bankrupt (with depositors rushing in to claim their d) and banned from future businesses. Thus, under deposit insurance, the dynamic contract has three components: the promised payment d, the termination and payment rules, $T(s, y)$ and $D(s, y)$, set out by the regulator and the insurance aspect: the premium ϕ and transfer $\tau(s, y)$.

Given the deposit insurance policy, the sequence of events is as described in Section 3.3, except that one needs to modify item 5, describing the bank's options after it receives payments from the firms, as follows:

(a) pay all depositors d and the premium ϕ (which may not be feasible);

(b) engage in self-verification (subject to payment and termination policy, $D(s, y)$ and $T(s, y)$, and deposit insurance pays $\tau(s, y)$ in addition to depositors);

(c) do nothing—in which case the bank's charter is automatically terminated.

We say that a deposit insurance scheme is **implementable** if the depositors are willing to participate, the bank has a nonnegative (expected) profit every period, and that the scheme is *ex ante* budget balanced. Almost by definition, with one single bank, any deposit insurance scheme cannot be ex post budget balanced if it is used in equilibrium. More precisely, let ζ_s be the expected funds available at the bank for state s in equilibrium. Then,

$$-\sum_{s\in L}\pi_s\tau(s,\zeta_s) + \sum_{s\notin L}\pi_s\phi \geq 0, \tag{9}$$

where L is the set of states at which the bank self-verifies. We have the following theorem.

Theorem 4.1. *Assume (A3) and assume that u is strictly concave. Let $(\mathbf{B},\mathbf{d})$ be an implementable dynamic bank contract that is not financially stable. Then, for β sufficiently high, there exists an implementable deposit insurance scheme and a bank contract that Pareto dominates it and is financially stable.*

Theorem 4.1 shows that, with the deposit insurance scheme, one can always achieve financial stability, and the scheme is Pareto optimal. We emphasize that the rationale for deposit insurance in our model is drastically different from that in Diamond and Dybvig (1983). In particular, the deposit insurance has to pay out its funds *in equilibrium*, while in Diamond-Dybvig model the scheme is used to restore good equilibria but should not be paid out in equilibrium.

Theorem 4.1 assumes ex ante budget balancedness. This amounts to assume that the external lender to the deposit insurance scheme can commit to future lending. In the Supplemental Appendix A2, we show that this commitment is not necessary. In the case of a monopoly bank, the deposit insurance scheme has to be funded externally (outside the banking system). However, if we extend our model to allow for multiple industries and assume that each industry has a specialized bank (which has a comparative advantage in monitoring that industry), we can get a similarly functioning deposit insurance policy. If we further assume that, while each industry may be subject to an idiosyncratic shock, the economy as a whole has no aggregate uncertainty, then the deposit insurance scheme can again restore financial stability for each bank without resorting to external funding. For details refer to Supplemental Appendix A2.

5 Concluding remarks

We investigated how financial intermediaries may improve welfare when there is asymmetric information between firms who require external funds and potential lenders about the firms' outputs and when such asymmetric information can only be resolved by costly monitoring. We showed that a large bank, which can take deposits from a large number of depositors and then lend to the firms, may reduce the cost of monitoring relative to direct contracts and hence improve welfare. In our setup the number of firms defaulting is then endogenously determined by the contracts. We showed that regulation Q could help reduce such deadweight loss of monitoring by requiring the bank not to offer a high interest rate to the depositors, which, in many cases, is actually profitable to do without interventions.

In our setup where the lender cannot commit to ex post monitoring, a new incentive issue emerges: with bank contract, the incentive to induce the bank to monitor the firms is a nontrivial issue. That issue is even more severe when we introduce aggregate uncertainty, and we showed that dynamic bank contracts under a charter system can be welfare improving. When such dynamic contract cannot reach financial stability, a deposit insurance scheme is proved to be welfare improving. This gives a new role for deposit insurance than the typically assumed panic-preventing function.

A Appendix: Proofs

Proof of Theorem 2.1

Before proving the theorem, we have the following useful lemma.

Lemma A.1. *An equilibrium outcome is characterized by a tuple,* $(W_0, W_1, ..., W_M, a, \bar{b})$, *where* $W_m \subset [0, \bar{w}]$ *is the set of returns under which* m *lenders verify with a positive probability that satisfies* $[(m-1)\bar{b}, m\bar{b}) \subset W_{M-m+1} \cup ... \cup W_M$ *and* $\bar{b}$ *is the repayment without verification. Moreover, for all* $m > 0$ *and for all* $w \in W_0 \cap W_m$, $a(w - (M-m)\bar{b}, m) = \bar{b}$.

Proof: Let (a, b) be a given contract and let (s_l, s_f) be a PBE. Let $A = \{\tilde{w} \in W : s_l(\tilde{w}) = 1\}$. Now, for each $m = 0, ..., M$, define

$$A_m = \{(\tilde{w}_1, ..., \tilde{w}_M) \in W^M : (\tilde{w}_{i_1}, ..., \tilde{w}_{i_m}) \in A^m, (\tilde{w}_{i_{m+1}}, ..., \tilde{w}_{i_M}) \in A^{M-m}\}, \qquad (10)$$

the set of messages for which m lenders verify in equilibrium, and let

$$W_m = \{w \in W : s_f(w) \text{ assigns a positive probability to } A^m\}.$$

Here we assume that $W_0 \neq \emptyset$; the proof for the other case is similar. Now, we show that, for each m and each $(\tilde{w}_1, ..., \tilde{w}_M) \in A_m$ that is in the support of s_f, if $\tilde{w}_i \notin A$, then

$$\bar{b} \equiv \min\{b(\tilde{w}_1) : (\tilde{w}_1, ..., \tilde{w}_M) \in A_0 \text{ and is assigned a positive probability under } s_f\}$$

$$= b(\tilde{w}_i). \qquad (11)$$

To see this, if this inequality doe snot hold, or if the minimum does not exist, then the firm always have a positive deviation to send another message to replace the message $\tilde{w}_i$ so that a lower payment without verification is feasible.

By (11) and feasibility, it is easy to see that

$$[(m-1)\bar{b}, m\bar{b}) \subset W_{M-m+1} \cup ... \cup W_M.$$

∎

We are now ready to prove theorem 2.1.

Let (a, b) be a given contract and let (s_l, s_f) be a PBE in which the lender's expected payoff is at least $u(r)$. For each $w \in W$ and each $m = 1, ...M$, we use $\kappa_m(w)$ to denote the probability that the firm sends a message to m lenders to coming to verify; note that by symmetry of s_f, conditional on having m lenders coming, s_f induces a uniform distribution over the identities of lenders to come. Obviously, $\kappa_m(w) > 0$ only if $w \in W_m$.

Note that by Lemma A.1, for all $w \in W_0$, the firm's payment is a constant, $M\bar{b}$. Now, let $\bar{b}'$ solve:

$$\int_0^{M\bar{b}'} w \, dF(w) + [1 - F(M\bar{b}')]M\bar{b}' \qquad (12)$$

$$= \int_{W_0} M\bar{b} \, dF(w) + \sum_{m=1}^{M} \int_{W_m} \left[ma\left(w - (M-m)\bar{b}, m\right) + (M-m)\bar{b} \right] \kappa_m(w) \, dF(w).$$

Note that since by feasibility, for each $w \in W_m$ with $m \geq 1$,

$$ma \left(w - (M - m)\bar{b}, m \right) + (M - m)\bar{b} \leq w, \tag{13}$$

it follows that $\bar{b}' < \bar{b}$, unless the above holds as an equality for almost every w below $M\bar{b}'$, in which case (a, b) is already a debt contract. Now, we construct a new contract, (a', b') and a new PBE (s'_f, s'_l) as follows. The strategy s'_f only sends two possible messages, $\tilde{w}_0$ and $\tilde{w}_1$, to each lender, and $s'_l(\tilde{w}_0) = 0$ and $s'_l(\tilde{w}_1) = 1$. We set $b'(\tilde{w}_0) = \bar{b}'$ and $b'(\tilde{w}) > \bar{b}'$ for any other $\tilde{w}$, and $a'(v, m) = v/m$ if $v \leq M\bar{b}'$ and $a'(v, m) = \bar{b}$ otherwise.

Now, for each $w \geq M\bar{b}'$, $s'_f(w)$ sends $\tilde{w}_0$ to all lenders with probability 1. Since $\bar{b}' \leq \bar{b}$, for any $w < M\bar{b}'$, $w \notin W_0$ by feasibility. For each $w < M\bar{b}'$, note that $\sum_{m=1}^{M} \kappa_m(w) = 1$ by feasibility. Thus, for each m, $s'_f(w)$ is a two-stage lottery: first it chooses m with probability $\kappa_m(w)$; second, conditional on choosing m, it sends exactly $\binom{M}{m}$ messages in the support of this second lottery, each of which is sent with equal probability and designates the m people verifying (and $M - m$ people getting fully repaid).

Now consider the two lotteries, X and Y, that describe each lender's payment to be received from the borrower, denoted p, induced by contracts $(a, \bar{b})$ and $(a', \bar{b}')$ and the above PBE's, respectively. Note that by (12) the two lotteries have the same expectation, i.e., $E_X[p] = E_Y[p]$. Let F_X and F_Y denote the distribution function for the two lotteries. We claim that for each $p < \bar{b}'$,

$$F_X[X < p] \geq F_Y[Y < p].$$

To see this, fix some $p < \bar{b}'$. Note that the realization of X and Y depend on the realization of w and the resolution of the randomization in s_f and s'_f. Given w, in both s_f and s'_f, there are two stages of randomization: the first involves the number of lenders to verify and the second involves the identities of lenders to verify.

Now, since $Y < p$ implies that $w < M\bar{b}'$, by construction of how s'_f assigns number of verifiers using κ_m's and by symmetry of the two strategies, the two-stage lotteries in the two strategies have exactly the same distributions, conditional on such w. As a result, for any such w that leads to $Y < p$, it corresponds to the event in which the lender receives $\tilde{w}_1$ at that w under $s'_f(w)$. As argued above, conditional on at w, that probability is exactly

28

the same as under $s_f(w)$, and by (13), at that event $X \leq Y < p$. This then implies that F_X and F_Y cross exactly once at the point $p = \bar{b}'$. Thus we have single crossing of CDFs and hence $(a', \bar{b}')$ second order stochastically dominates $(a, \bar{b})$.

The overall probability of verification is strictly smaller under $(a', \bar{b}')$ than under $(a, \bar{b})$, since under $(a', \bar{b}')$, the corresponding verification regions, $W'_m = W_m \cap [0, M\bar{b}')$ for all $m \geq 1$, $\kappa_m(w)$ is the same under s'_f for all $w < M\bar{b}'$, and since $\bar{b}' < \bar{b}$. Thus the contract $(a', \bar{b}')$ strictly improves the lenders' expected utilities, while leaving the borrower's expected utility unchanged. This also implies $\bar{b}' > r$ as the original contract is implementable.

Note that under the debt contract $(a', \bar{b}')$, the payoffs of the two parties are continuous in $\bar{b}'$ (decrease the full repayment level, keep all W_m regions the same). This follows since the contract and utility functions of all agents are continuous. Since $\bar{b}' > r$, we can find $\bar{b}'' < \bar{b}'$ such that the contract $(a', \bar{b}'')$ still strictly improves the lenders' expected utilities, while the borrower now makes a strictly smaller expected payment and thus we have a strict Pareto improvement.

Finally, we need to show that (s'_f, s'_l) does constitute a PBE. It is straightforward to see that s'_f is optimal. Now, to show that s'_l is optimal, we need to show that the lender has the incentive to verify whenever seeing message $\tilde{w}_1$ and has incentive to take $\bar{b}'$ whenever seeing message $\tilde{w}_0$. The latter follows since the lender can never get more than $\bar{b}'$ by verifying. For the former, one needs to show that the expected payoff to the lender is higher than γ when seeing $\tilde{w}_1$. Now (A0) guarantees all lenders have an incentive to verify, since $\bar{b}' > r$ and $\frac{1}{F(M\bar{b}')} \int_0^{\bar{b}'} u\left(\frac{w}{M}\right) dF(w) > \frac{1}{F(Mr)} \int_0^r u\left(\frac{w}{M}\right) dF(w) > \gamma$; where the second inequality holds because u is increasing. ∎

The optimal verfication intervals, W_m's, are derived in appendix A.1.

Proof of Lemma 3.1

We first show necessity. Suppose that (2) does not hold. Then, the bank is making a negative profit in equilibrium. But by not verifying any firm at all, the bank can make a nonnegative profit and hence verification is not optimal. Similarly, suppose that (3) does

not hold. We have shown that (2) is necessary, and that, together with the assumption that (3) does not hold, implies

$$[1 - F(B)]B > d.$$

Now, this implies that the bank would have sufficient funds to pay to each depositor without verifying any firm that sent message 1. Since verifying each individual firm results in a net loss for the bank, it is better off not to do it.

Here we show sufficiency. We show that in equilibrium the bank sends message 0 to all depositors and no depositor verifies the bank, and if the bank sends message 1, the verifying depositors share all the remaining funds. We separate two cases.

(a) Suppose that $[1 - F(B)]B \geq Md$. Then, the bank can pay off d to each depositor without verifying any firm that sent message 1. However, by (3), the bank is making a positive expected profit by verifying such a firm, and hence the bank is willing to do that.

(b) Suppose that $[1 - F(B)]B < Md$. Thus, the bank cannot meet its obligations to all depositors unless it verifies a positive fraction of firms that sent message 1. Let f^* be the minimum fraction that the bank can meet that, and let f be the fraction of such firms that the bank actually verifies. Since the bank will pay out all its available funds when being verified, if $0 < f < f^*$, then the bank's profit is negative; if $f^* \leq f < 1$, then it is optimal for the bank to pay off d to all depositors, and its payoff (in per depositor terms) is given by

$$\frac{1}{M}\left[f \int_0^B [w - c_E]\mathrm{d}F(w) + [1 - F(B)]B\right] - d,$$

which is maximized at $f = 1$ by (3). Finally, having $f = 1$ is better than having $f = 0$ because of (2).

Proof of Theorem 3.1

Concavity of u implies that, together with (A1),

$$u'(r)c_E < \sum_{m=1}^{M} \frac{F\left(\frac{mB}{M}\right)}{F(B)}\gamma. \tag{14}$$

To see this, note that (A0) implies that $\gamma < u(r)$. Now,

$$c_E < u^{-1}(\gamma) \sum_{m=1}^{M} \frac{F\left(\frac{mB}{M}\right)}{F(B)} \leq \frac{\sum_{m=1}^{M} \frac{F\left(\frac{mB}{M}\right)}{F(B)} \gamma}{u'\left[u^{-1}(\gamma)\right]} < \frac{\sum_{m=1}^{M} \frac{F\left(\frac{mB}{M}\right)}{F(B)} \gamma}{u'(r)}.$$

Moreover, we show that

$$c_E F(B) < \int_0^B w \mathrm{d}F(w). \tag{15}$$

Note that implementability implies that

$$\gamma < \frac{\sum_{m=0}^{M-1} \int_{mB/M}^{(m+1)B/M} u((w-mB)/(M-m))\mathrm{d}F(w)}{\sum_{m=0}^{M-1} F((m+1)B/M)} < \frac{\int_0^B u(w/M)}{F(B)} < u\left(\frac{\int_0^B \frac{w}{M}\mathrm{d}F(w)}{F(B)}\right).$$

Therefore,

$$c_E < u^{-1}(\gamma) \sum_{m=1}^{M} \frac{F\left(\frac{mB}{M}\right)}{F(B)} < \frac{\int_0^B \frac{w}{M}\mathrm{d}F(w)}{F(B)} \sum_{m=1}^{M} \frac{F\left(\frac{mB}{M}\right)}{F(B)} < \frac{\int_0^B w\mathrm{d}F(w)}{F(B)}.$$

Consider the following bank contract, (B, d). The bank has a simple debt contract with each firm with debt level B. For the deposit side, the depositor receives d that will be specified below without verification, and, in case the bank refuses to pay d, all depositors verify and receive equal payments from the remaining funds.

Let

$$d = \frac{1}{M}\left\{\int_{w\in[0,B]} w\mathrm{d}F(w) + (1 - F(B))B - F(B)c_E - \varepsilon\right\}. \tag{16}$$

For any $\varepsilon > 0$, (2) is satisfied with a strict inequality and, recalling that (15) implies (3), the bank has strict incentive to verify firms' returns. It also implies that the bank has a strictly positive payoff.

Now, let U be the expected payoff for each lender from the direct contract. Then,

$$U \leq \int_0^B u\left(\frac{w}{M}\right)\mathrm{d}F(w) + (1 - F(B))u\left(\frac{B}{M}\right) - \frac{1}{M}\sum_{m=1}^{M} F\left(\frac{mB}{M}\right)\gamma. \tag{17}$$

We claim that $u(d + \varepsilon/M) > U$ for any $\varepsilon > 0$ (note that d is defined by (16)) and hence $u(d) > U$ for ε small. To see this, since u is concave, we have:

$$u\left[\frac{1}{M}\left\{\int_0^B w\mathrm{d}F(w) + (1 - F(B))B\right\}\right] \geq \int_{w\in[0,B)} u\left(\frac{w}{M}\right)\mathrm{d}F(w) + (1 - F(B))\frac{u(B)}{M},$$

31

that is,

$$u\left[d + \varepsilon/M + \frac{1}{M}F(B)c_E\right] \geq \int_{w \in [0,B)} u\left(\frac{w}{M}\right) dF(w) + (1 - F(B))\frac{u(B)}{M}. \tag{18}$$

Now, (A1) implies that $d + \varepsilon/M > r$ and hence, by concavity of u,

$$\begin{aligned} u(d + \varepsilon/M) &\geq u\left[d + \frac{\varepsilon}{M} + \frac{1}{M}F(B)c_E\right] - u'(d)\frac{1}{M}F(B)c_E \\ &\geq u\left[d + \frac{1}{M}F(B)c_E\right] - \frac{1}{M}\sum_{m=1}^{M}F\left(\frac{mB}{M}\right)\gamma \geq U, \end{aligned}$$

where the second inequality follows from (14) and hence

$$c_E < \sum_{m=1}^{M}\frac{F\left(\frac{mB}{M}\right)}{F(B)}\gamma/u'(r) < \sum_{m=1}^{M}\frac{F\left(\frac{mB}{M}\right)}{F(B)}\gamma/u'(d + \varepsilon/M)$$

as $d + \varepsilon/M > r$, and the last inequality follows from (18).

Finally, we need to show that the depositors have incentive to verify the bank's return when the bank does not pay off d. For the continuum model this is off-equilibrium path and we assume that, when receiving the message 1, the depositors believe that the bank has d. For finite but large number of depositors, this follows from the CLT, and one can show that

$$\lim_{N \to \infty} \mathbb{E}[u(y_N)|y_N < \mathbb{E}(y_N) - F(B)c_E] = u(d) > U \geq u(r) > \gamma,$$

where y_N is the random variable that represents the average funds available in the bank to each depositor when there are N firms. $\square$

Proof of Theorem 3.2

Let d_s denote the average amount paid to the depositors at state s, given a contract (d, e). The proof of theorem3.2 is a corollary of the following lemma.

Lemma A.2. *Let $(\mathbf{B}, \mathbf{d})$ be an implementable bank contract. Then, $d_s \leq \eta_s(\mathbf{B})$ for both $s = h, \ell$, and it has to satisfy the following conditions.*

(a) Suppose that $d = d_h = d_\ell$. Then, $u(d) \geq u(r)$, and

$$\int_0^{B_s} w dF_s(w) > F_s(B_s)c_E \text{ for both } s = h, \ell. \tag{19}$$

(b) If $d_\ell < d_h$, there are two subcases:

(b.1) $d_h = d$ and letting m_ℓ be the fraction of depositors monitoring at state ℓ,

$$\int_0^{B_h} w \mathrm{d}F_h(w) > F_h(B_h)c_E; \tag{20}$$

$$u\left(\frac{Md_\ell - (M - m_\ell)d}{m_\ell}\right) \geq \gamma; \tag{21}$$

$$\pi_h u(d) + \pi_\ell \left[\frac{M - m_\ell}{M}u(d) + \frac{m_\ell}{M}u\left(\frac{Md_\ell - (M - m_\ell)d_h}{m}\right) - \frac{m_\ell}{M}\gamma\right] \geq u(r). \tag{22}$$

(b.2) $d_h < d$ and letting m_s be the fraction of depositors monitoring at state $s = h, \ell$,

$$\sum_{s = h, \ell} \frac{\pi_s m_s}{\pi_h m_h + \pi_\ell m_\ell} u\left(\frac{Md_s - (M - m_s)d}{m_s}\right) \geq \gamma; \tag{23}$$

$$\sum_{s = h, \ell} \pi_s \left[\frac{M - m_s}{M}u(d) + \frac{m_s}{M}u\left(\frac{Md_s - (M - m_s)d}{m_s}\right) - \frac{m_s}{M}\gamma\right] \geq u(r). \tag{24}$$

Proof. The necessity in (a) uses the same arguments as in Lemma 3.1. Consider (b). Conditions (22) and (24) are obvious IR conditions. (21) and (23) are necessary for depositors to verify the bank. For (b.1), since $d < \zeta_h(\mathbf{B})$, for the bank to verify in full at state h, (20) is necessary. $\square$

Now, suppose, by contradiction, that the bank can offer a financially stable contract with debt level with depositors d. By Lemma A.2, $B_\ell \leq r$ implies that

$$d = d_\ell \leq \eta_\ell(\mathbf{B}) < r,$$

which leads to a contradiction with implementability. $\square$

Proof of Theorem 3.3

(a) Let B be a direct contract between the firm and the lenders that satisfies (A1) and (A2) for $F = \sum_{s \in S} \pi(s)F_s$. Concavity of u implies that, together with (A1),

$$u'(r)c_E < \frac{\sum_{s = h, \ell} \pi_s \left[\sum_{m=1}^{M} F_s\left(\frac{mB}{M}\right)\right]}{\sum_{s = h, \ell} \pi_s F_s(B)}\gamma. \tag{25}$$

To see this, note that (A0) implies that $\gamma < u(r)$. Now,

$$
\begin{aligned}
c_E &< u^{-1}(\gamma) \frac{\sum_{s=h,\ell} \pi_s \left[\sum_{m=1}^{M} F_s \left(\frac{mB}{M} \right) \right]}{\sum_{s=h,\ell} \pi_s F_s(B)} < \frac{\sum_{s=h,\ell} \pi_s \left[\sum_{m=1}^{M} F_s \left(\frac{mB}{M} \right) \right]}{\sum_{s=h,\ell} \pi_s F_s(B)} \frac{\gamma}{u' \left[u^{-1}(\gamma) \right]} \\
&< \frac{\sum_{s=h,\ell} \pi_s \left[\sum_{m=1}^{M} F_s \left(\frac{mB}{M} \right) \right]}{\sum_{s=h,\ell} \pi_s F_s(B)} \frac{\gamma}{u'(r)},
\end{aligned}
$$

where the last inequality follows from the fact that $u^{-1}(\gamma) < r$ and that u is strictly concave.

Now we construct the following bank contract, (B, d), and show that it dominates the original direct contract, B. The bank has a simple debt contract with each firm with debt level B (and hence the firms are indifferent between the original contract and the constructed contract), and let $\eta_s(B)$ and $\zeta_s(B)$ be defined as in (6) and (7) with $B_h = B_\ell = B$ for both $s = h, \ell$. Now, let U be the expected payoff for each lender from the direct contract. Then,

$$
U \leq \sum_{s=h,\ell} \pi_s \left\{ \int_0^B u \left(\frac{w}{M} \right) dF_s(w) + (1 - F_s(B)) u \left(\frac{B}{M} \right) - \frac{1}{M} \sum_{m=1}^{M} F_s \left(\frac{mB}{M} \right) \gamma \right\}. \quad (26)
$$

We consider two cases.

(a.1) Suppose that $\zeta_\ell(B) \geq \sum_{s=h,\ell} \pi_s \eta_s(B)$. Then, set $\hat{d} = \frac{1}{M} \sum_{s=h,\ell} \pi_s \eta_s(B)$. We claim that $u(\hat{d}) > U$ and hence $u(\hat{d} - \varepsilon) > U$ for ε small. This then implies that the depositors strictly prefer the constructed contract. To see this, since u is strictly concave, we have

$$
u \left[\hat{d} + \sum_{s=h,\ell} \pi_s \frac{1}{M} F_s(B) c_E \right] > \sum_{s=h,\ell} \pi_s \left\{ \int_{w \in [0,B]} u \left(\frac{w}{M} \right) dF_s(w) + (1 - F_s(B)) u \left(\frac{B}{M} \right) \right\}. \quad (27)
$$

Now, (A2) implies that $\hat{d} > r$ and hence, by strict concavity of u,

$$
\begin{aligned}
u(\hat{d}) &> u \left[\hat{d} + \sum_{s=h,\ell} \pi_s \frac{1}{M} F_s(B) c_E \right] - u'(\hat{d}) \left[\sum_{s=h,\ell} \pi_s \frac{1}{M} F_s(B) c_E \right] \\
&> u \left[\hat{d} + \sum_{s=h,\ell} \pi_s \frac{1}{M} F_s(B) c_E \right] - \frac{1}{M} \sum_{s=h,\ell} \pi_s \left[\sum_{m=1}^{M} F_s \left(\frac{mB}{M} \right) \right] \gamma \geq U,
\end{aligned}
$$

where the second inequality follows from (25) and $\hat{d} > r$, and the last inequality follows from (26) and (27).

Thus, for $\varepsilon > 0$ small, $u(\hat{d} - \varepsilon) > U$. Now, set the contract with the depositors as $d = \hat{d} - \varepsilon$, and there is no self-verification on the equilibrium path. On the off-equilibrium path, we set $T(s, y) = 0$ and $D(s, y) = y$ for all y and for both s. The depositors are then better off against the direct contract. The bank makes a strictly positive expected profit, which (in per depositor terms) equals

$$\frac{1}{M} \sum_{s=h,\ell} \pi_s \eta_s(B) - d = \hat{d} - (\hat{d} - \varepsilon) = \varepsilon > 0.$$

Given the contract (T, D), for sufficiently high β, the bank has the incentive to repay d to all depositors as failure to do so will result in losing all future profits, and $\zeta_\ell(B) \geq \sum_{s=h,\ell} \pi_s \eta_s(B)$ ensures that the bank always have sufficient fund to do so if they verify all firms that sent message 1. Using the same logic as before, the bank has incentive to verify such firms because of (A3).

(a.2) Suppose that $\zeta_\ell(B) < \sum_{s=h,\ell} \pi_s \eta_s(B)$. Then, set

$$d_\ell = \frac{1}{M} \zeta_\ell(B) \quad \text{and} \quad \hat{d}_h = \eta_h(B) - \frac{1}{M} \frac{\pi_\ell}{\pi_h} F_\ell(B) c_E.$$

Here d_ℓ will be the average repayment to the depositors at state ℓ and $\hat{d}_h - \varepsilon$ will be the average repayment to the depositors at state h with $\varepsilon > 0$ to be determined below. We claim that

$$\pi_h u(\hat{d}_h) + \pi_\ell u(d_\ell) > U. \tag{28}$$

To see this, since u is strictly concave, we have

$$\pi_h u \left[\hat{d}_h + \frac{1}{M} \left(F_h(B) c_E + \frac{\pi_\ell}{\pi_h} F_\ell(B) c_E \right) \right] + \pi_\ell u(d_\ell) \tag{29}$$
$$> \sum_{s=h,\ell} \pi_s \left\{ \int_{w \in [0,B]} u \left(\frac{w}{M} \right) \mathrm{d}F_s(w) + (1 - F_s(B)) u \left(\frac{B}{M} \right) \right\}.$$

Now, (A2) and $\zeta_\ell(B) < \sum_{s=h,\ell} \pi_s \eta_s(B)$ imply that $\hat{d}_h > r$ and hence, by strict concavity of u,

$$u(\hat{d}_h) > u \left[\hat{d} + \frac{c_E}{M} \left(F_h(B) c_E + \frac{\pi_\ell}{\pi_h} F_\ell(B) \right) \right] - u'(\hat{d}_h) \left[\frac{c_E}{M} \left(F_h(B) c_E + \frac{\pi_\ell}{\pi_h} F_\ell(B) \right) \right]$$
$$\geq u \left[\hat{d} + \frac{c_E}{M} \left(F_h(B) c_E + \frac{\pi_\ell}{\pi_h} F_\ell(B) \right) \right] - \frac{1}{M \pi_h} \sum_{s=h,\ell} \pi_s \left[\sum_{m=1}^{M} F_s \left(\frac{mB}{M} \right) \right],$$

where the second inequality follows from (25) and $\hat{d} > r$, and hence,

$$\pi_h u(\hat{d}_h) + \pi_\ell u(d_\ell)$$
$$> \sum_{s=h,\ell} \pi_s \left\{ \int_{w \in [0,B]} u\left(\frac{w}{M}\right) dF_s(w) + (1 - F_s(B)) u\left(\frac{B}{M}\right) \right\} - \frac{1}{M} \sum_{s=h,\ell} \pi_s \left[\sum_{m=1}^{M} F_s\left(\frac{mB}{M}\right) \right] \geq U.$$

Thus, for $\varepsilon > 0$ small, $\pi_h u(\hat{d}_h - \varepsilon) + \pi u(d_\ell) > U$. Now, set the contract with the depositors as $d = \hat{d}_h - \varepsilon$, which is paid to the depositors when $s = h$ without self-verification, and, at $s = \ell$, there is self-verification and the bank pays d_ℓ equally to all depositors; i.e., $D(\ell, \zeta_\ell(B)) = d_\ell$. The depositors are then strictly better off against the direct contract. We set $T(h, y) = 0$ and $D(h, y) = y$ for all y, and $T(\ell, y) = 0$ iff $y < \zeta_\ell(B)$ and $D(\ell, y) = y$ for all y.

The bank has a positive expected profit provided that $c_B < \varepsilon/\pi_\ell$. Given this positive profit, the bank has incentive to repay d at state h and to self-verify at ℓ, as failure to do so results in losing all future profits, provided that β is sufficiently high, if it has verified all firms that sent message 1. Using the same logic as before, the bank has incentive to verify all such firms because of (A3).

(b) Let $(\mathbf{B}, d)$ be an implementable static contract that is not stable. With no loss of generality we assume $d_\ell < d_h$ and $\zeta_\ell(\mathbf{B}) \leq \zeta_h(\mathbf{B})$ (the fact that the order of the d's align with the order of the ζ's does not matter either). Incentive compatibility requires $d_\ell \leq \eta_\ell(\mathbf{B})$.

We consider two cases.

(b.1) Suppose that $\hat{d} = \pi_h d_h + \pi_\ell d_\ell \leq \zeta_\ell(\mathbf{B})$. We can choose $\varepsilon > 0$ such that

$$\pi_h u(d_h) + \pi_\ell u(d_\ell) - \frac{d - d_\ell}{d} \gamma < u(\hat{d} - \varepsilon)$$

because u is strictly concave. Then, set the new contract as $(\mathbf{B}, \mathbf{d}')$ with $d'_h = d'_\ell = d' = \hat{d} - \varepsilon$, and the bank never engages in self-verification. Note that since the static contract is implementable, the bank's profit is nonnegative. This implies that

$$\hat{d} \leq \sum_{s=h,\ell} \pi_s \left\{ [1 - F_s(B_s)] B_s + \int_0^{B_s} w dd F_\ell(w) - F_s(B_s) c_E \right\}.$$

36

Thus, the bank's profit is at least $\varepsilon > 0$ under the new contract. Moreover, the depositors are strictly better off against the static contract.

The dynamic contract is given as follows. $T(s, y) = 0$ and $D(s, y) = y$ for all s, y. Clearly $d' < \zeta_h(\mathbf{B})$. Using the same logic as in (a.1), we can show that for β sufficiently high, the bank has incentive to repay d' at both states because of future profits and has incentive to verify all firms that send message 1 because of (A3).

(b.2) Suppose that $\pi_h d_h + \pi_\ell d_\ell > \zeta_\ell(\mathbf{B})$. Let $\hat{d}_h$ be such that

$$\pi_h \hat{d}_h + \pi_\ell \zeta_\ell(\mathbf{B}) = \pi_h d_h + \pi_\ell d_\ell.$$

Since $d_\ell \leq \eta_\ell(\mathbf{B})$, $\hat{d}_h < d_h$. Let $d'_\ell = \zeta_\ell(\mathbf{B})$, and let

$$T(\ell, y) = 1 \text{ iff } y \geq d'_\ell \text{ and } D(\ell, y) = y \text{ for all } y.$$

Then, since u is strictly concave,

$$\pi_h u(\hat{d}_h) + \pi_\ell u(d'_\ell) > \pi_h u(d_h) + \pi_\ell u(d_\ell) - \frac{d - d_\ell}{d}\gamma,$$

and hence for $\varepsilon > 0$ small,

$$\pi_h u(\hat{d}_h - \varepsilon) + \pi_\ell u(d'_\ell) > \pi_h u(d_h) + \pi_\ell u(d_\ell) - \frac{d - d_\ell}{d}\gamma.$$

The bank then promise to repay depositors $d' = \hat{d}_h - \varepsilon$. The contract is such that $T(h, y) = 0$ and $D(h, y) = y$ for all y. On the equilibrium path, the bank pays $d' = \hat{d}_h - \varepsilon$ to all depositors at $s = h$ without self-verification, and self-verifies when $s = \ell$. Thus, if $c_B < \frac{\pi_h}{2\pi_\ell}\varepsilon$, the bank is making a strictly positive expected profit. Moreover, the depositors are strictly better off against the static contract.

Finally, the fact that the bank is willing to repay d' at $s = h$ and to self-verify at $s = \ell$ for $c_B < \frac{\pi_h}{2\pi_\ell}\varepsilon$ and for sufficiently high β follows exactly the same logic as in (a.2) because of (A3). $\square$

Proof of Lemma 4.1

Since $u(c) = c$, under direct contracting and debt level B, the depositor's payoff under the optimal contract is given by

$$U(B) \equiv \int_{w<B} \frac{w}{M} \, dF(w) + [1 - F(B)]\frac{B}{M} - \gamma\frac{1}{M}\sum_{n=1}^{M} F\left(\frac{nB}{M}\right).$$

Such a contract is implementable if and only if $U(B) \geq r$. The result then follows immediately. $\square$

Proof of Theorem 4.1

Let d_s be the payment to the depositors under the original dynamic contract at state s. We may assume that $d_h = d$, the promised payment. Since it is not financially stable, $d_\ell < d$ and the bank has to engage self-verification at $s = \ell$.

Let $\hat{d} = \pi_h d_h + \pi_\ell d_\ell$. If $\hat{d} \leq \zeta_\ell(\mathbf{B})$, then, using the same arguments as in Theorem 3.3, we can have a financially stable contract that dominates the original one without using deposit insurance. So suppose that $\hat{d} > \zeta_\ell(\mathbf{B})$ and hence, for some $\varepsilon > 0$,

$$u(\hat{d} - \varepsilon) > \sum_{s=h,\ell} \pi_s u(d_s). \tag{30}$$

Then, take $d' = \hat{d} - \varepsilon$, $L = \{\ell\}$ and let $\tau(\ell, y) = \max\{d - y, 0\}$ if $y \geq \zeta_\ell(\mathbf{B})$, $\tau(\ell, y) = 0$ otherwise, $T(\ell, y) = 1$ iff $y \geq \zeta_\ell(\mathbf{B})$ and $D(\ell, y) = y$ for all y. For $s = h$, $\tau(h, y) = 0$, $T(h, y) = 0$, and $D(h, y) = y$ for all y. Let:

$$\phi = \frac{\pi_\ell \tau(\ell, \zeta_\ell(\mathbf{B}))}{\pi_h}.$$

On the proposed equilibrium path, the bank self-verifies only in state ℓ, and, in that state, the bank pays off all its available funds, $\zeta_\ell(\mathbf{B})$, and the deposit insurance scheme pays out additional funds so that all depositors receive d'; at state h, the bank pays d' to all depositors and pays ϕ to the deposit insurance. Note that the equilibrium τ is given by $d' - \zeta_\ell(\mathbf{B})$ and that on the proposed equilibrium path the budget constraint (9) is satisfied.

The bank's expected profit (in per depositor term) on the proposed equilibrium path is then given by

$$\pi_h \left[\eta_h(\mathbf{B}) - d' - \phi\right] + \pi_\ell \left[\eta_\ell(\mathbf{B}) - \zeta_\ell(\mathbf{B}) - c_B\right]$$

$$= \left[\sum_{s=h,\ell} \pi_s \eta_s(\mathbf{B})\right] - \pi_h d' - \pi_h \frac{\pi_\ell[d' - \zeta_\ell(\mathbf{B})]}{\pi_h} - \pi_\ell \zeta_\ell(\mathbf{B}) - \pi_\ell c_B$$

$$= \left[\sum_{s=h,\ell} \pi_s \eta_s(\mathbf{B})\right] - d' - \pi_\ell c_B = \left\{\sum_{s=h,\ell} \pi_s[\eta_s(\mathbf{B}) - d_s]\right\} + \varepsilon - \pi_\ell c_B,$$

which is then ε above to the original dynamic contract in each period, and hence is strictly positive. Moreover, the depositors are strictly better off under deposit insurance than the original unstable contract by (30).

Finally, using exactly the same arguments as in the proof of Theorem 3.3, part (b.2), high β ensures that the bank wants to pay $d' + \phi$ at state h, self verifies at state ℓ, and (A3) ensures that the bank has incentive to verify all firms that sent message 1. $\square$

References

[1] Diamond, D. (1984). "Financial Intermediation and Delegated Monitoring," *Review of Economic Studies*, 51, 393-414.

[2] Diamond, D. and P. Dybvig, (1983). "Bank Runs, Deposit Insurance, and Liquidity," *Journal of Political Economy,* 91, 401-419.

[3] Freixas, X. and J.-C. Rochet, (1997). *Microeconomics of Banking.* MIT Press.

[4] Gorton, G. and A. Winton, (2003). "Financial Intermediation," in *Handbook of the Economics of Finance*, edited by G. Constantinides, M. Harris and R. Stulz, Amsterdam: North Holland.

[5] Gale, D. and M. Hellwig, (1985). "Incentive-Compatible Debt Contracts: The One Period Problem," *Review of Economic Studies,* 52, 647-663.

[6] Krasa, S. and A. Villamil, (1992). "Monitoring the Monitor: An Incentive Structure for a Financial Intermediary," *Journal of Economic Theory*, 57, 197-221.

[7] Krasa, S. and A. Villamil, (2000). "Optimal Contracts When Enforcement is a Decision Variable," *Econometrica*, 68, 119-134.

[8] Mookherjee, D. and I. Png, (1989). "Optimal Auditing, Insurance and Redistribution," *Quarterly Journal of Economics,* 104, 399-415.

[9] Philippon, T., (2015). "Has the US Finance Industry Become Less Efficient? On the Theory and Measurement of Financial Intermediation," *American Economic Review,* 105, 1408-1438.

[10] Sun, Y. (2006). "The Exact Law of Large Numbers via Fubini Extension and Characterization of Insurable Risks," *Journal of Economic Theory*, 126(1), 31–69.

[11] Thomas, J. and T. Worrall (1990). "Income Fluctuation and Asymmetric Information: An Example of a Repeated Principal-Agent Problem," *Journal of Economic Theory*, 51, 367-90.

[12] Townsend, R., (1979). "Optimal Contracts and Competitive Markets with Costly State Verification," *Journal of Economic Theory,* 21, 265-293.

[13] Winton, A. (1995). "Costly State Verification and Multiple Investors: The Role of Seniority." *Review of Financial Studies,* 8, 91-123.

[14] Williamson, S., (1986). "Costly Monitoring, Financial Intermediation, and Equilibrium Credit Rationing," *Journal of Monetary Economics*, 18, 159-179.

Supplemental Appendix: for online publication only

A.1 Optimal verification regions

Theorem 2.1 proved the optimality of debt contracts, but didn't characterize their specific form. We now extend this result by finding the optimal verification regions (the W_m) for the debt contract. Let (a, b) be a debt contract, where we denote by b the level of debt for each lender. For each $w \in [0, Mb)$, the *ex ante* expected payoff to a lender for having m lenders coming to verify is

$$\Upsilon(w, m) \equiv \frac{m}{M} \left[u \left(\frac{w - (M - m)b}{m} \right) - \gamma \right] + \frac{M - m}{M} u(b). \tag{31}$$

Let $k := \min \{\arg\max_{m \geq 1} \Upsilon(Mb, m)\}$, i.e., let k be the number of lenders we want verifying at Mb (if there are multiple maximizers at Mb, we take the smallest number of verifiers since that will be the relevant one leading up to Mb).

Let $x_k = Mb$ and for each $m = M, M - 1, ..., k + 1$, define x_m as follows:

$$x_m = \inf\{w \geq (M - m + 1) b : \Upsilon(w, m - 1) \geq \Upsilon(w, m)\}.$$

Lemma A.3. *For each* $m = k, ..., M$, x_m *is the unique* w *such that either* $\Upsilon(w, m) = \Upsilon(w, m - 1)$ *or* $x_m = (M - m + 1) b$. *Moreover,* $x_m > x_n$ *for all* $m < n$.

Proof. Fix some m. Suppose that $\Upsilon((M - m + 1)b, m) > \Upsilon((M - m + 1)b, m - 1)$. Then first note that

$$\frac{\partial}{\partial w} [\Upsilon(w, m) - \Upsilon(w, m - 1)] = \frac{1}{M} \left[u' \left(\frac{w - (M - m)b}{m} \right) - u' \left(\frac{w - (M - m + 1)b}{m - 1} \right) \right] < 0$$

for all $w < Mb$; this follows since u is concave and

$$\frac{w - (M - m)b}{m} - \frac{w - (M - m + 1)b}{m - 1} = \frac{Mb - w}{m(m - 1)}.$$

Consider the case $w = Mb$; we have

$$\Upsilon(Mb, m) - \Upsilon(Mb, m - 1) = \frac{1}{M} \gamma < 0,$$

41

and hence there is a unique x_m such that $\Upsilon(x_m, m) = \Upsilon(x_m, m-1)$. Thus we have that for each m we have either $\Upsilon(x_m, m) = \Upsilon(x_m, m-1)$ or $x_m = (M - m + 1)\, b$.

Next we will show that $x_m > x_{m-1}$ and the result follows by induction. First consider the case where $x_m > (M - m + 1)\, b$. Thus $\Upsilon(x_m, m) = \Upsilon(x_m, m-1)$ and $\Upsilon(x_{m-1}, m-1) \le \Upsilon(x_{m-1}, m-2)$ implies that:

$$
\begin{aligned}
0 &= mu\left(\frac{x_m - (M-m)b}{m}\right) - (m-1)u\left(\frac{x_m - (M-m+1)b}{m-1}\right) - u(b) - \gamma \\
&\ge (m-1)u\left(\frac{x_{m-1} - (M-m+1)b}{m-1}\right) - (m-2)u\left(\frac{x_{m-1} - (M-m+2)b}{m-2}\right) - u(b) - \gamma.
\end{aligned}
$$

Suppose, by contradiction, that $x_{m-1} \ge x_m$. Then,

$$
\begin{aligned}
&mu\left(\frac{x_{m-1} - (M-m)b}{m}\right) - (m-1)u\left(\frac{x_{m-1} - (M-m+1)b}{m-1}\right) \\
&\ge mu\left(\frac{x_m - (M-m)b}{m}\right) - (m-1)u\left(\frac{x_m - (M-m+1)b}{m-1}\right) \\
&\ge (m-1)u\left(\frac{x_{m-1} - (M-m+1)b}{m-1}\right) - (m-2)u\left(\frac{x_{m-1} - (M-m+2)b}{m-2}\right),
\end{aligned}
$$

which implies that

$$
\begin{aligned}
&\frac{m}{2m-2}u\left(\frac{x_{m-1} - (M-m)b}{m}\right) + \frac{(m-2)}{2m-2}u\left(\frac{x_{m-1} - (M-m+2)b}{m-2}\right) \\
&\ge u\left(\frac{x_{m-1} - (M-m+1)b}{m-1}\right).
\end{aligned}
$$

But notice that

$$
\frac{x_{m-1} - (M-m+1)b}{m-1} = \frac{m}{2m-2}\left(\frac{x_{m-1} - (M-m)b}{m}\right) + \frac{(m-2)}{2m-2}\left(\frac{x_{m-1} - (M-m+2)b}{m-2}\right),
$$

and this leads to a contradiction to the concavity of u.

Now consider the case where $x_m = (M - m + 1)\, b$. Because of the definition of x_{m-1}, we have that $x_{m-1} \ge (M - m + 2)\, b > x_m$. $\qquad\square$

A.2   Optimality of debt contract under aggregate uncertainty

The bank contract with the depositors is given by a number d and a function $e(y, m, s)$ that specifies payment to each verifying depositor, where d is the debt level, y is the

available funds to each verifying depositor, m is the fraction who verify, and s is the state. Our earlier results regarding the optimality of debt contracts with firms, however, do not directly extend to the bank contract, since the bank's funds are (at least partially) paid to the depositors. In general, a bank contract with the firm may be denoted $(a_s, \bar{b}_s)_{s=h,\ell}$, where $a_s(w)$ denotes the firm's repayment to the bank when verified return is w and $\bar{b}_s$ is the repayment without verification, both of which may depend on the state s.

Given the contract, the strategies can be defined in an analogous manner to those in Definition 3.1, but with one modification: now both s_f and (s_B^1, s_B^2) may depend on the state s. Moreover, as before, we focus on symmetric equilibria in which the interactions between the bank and the firms in a PBE with the property that the bank verifies all firms that send message 1. The following lemma shows that it is optimal to have debt contracts between the bank and the firms.

Lemma A.4. *Let $\{(a_s, \bar{b}_s)_{s=h,\ell}, (d, e)\}$ be an implementable bank contract. Under (A3), there is another bank contract in which the contract with the firm is a debt contract with verification occurring iff $w < B_s$ and B_s the debt level at state s and which Pareto dominates the original contract.*

Proof. Suppose that the contracts with the firms are given by $(a_s, \bar{b}_s)_{s=h,\ell}$, and suppose that $W_{0,s}$ is the set where no verification occurs from the bank, $s = h, \ell$. Let

$$\eta_s = \frac{1}{M} \left\{ \int_{W_{1,s}} a_s(w) \mathrm{d}F_s(w) - F_s(W_{1,s})c_E + F_s(W_{0,s})\bar{b}_s \right\},$$

and let

$$\zeta_s = \frac{1}{M} \left\{ \int_{W_{1,s}} a_s(w) \mathrm{d}F_s(w) + F_s(W_{0,s})\bar{b}_s \right\}.$$

Here η_s is the revenue (net of monitoring cost) from the firms by the bank, and ζ_s is the available funds at the bank at state s, both in per depositor's term.

Consider the alternative contracts $(B_s)_{s \in S}$ and e' as follows. B_s is such that

$$\int_0^{B_s} w \mathrm{d}F_s(w) + [1 - F_s(B_s)]B_s = M\zeta_s,$$

43

and hence $B_s \leq \bar{b}_s$. Let $W'_{1,s} = [0, B_s) \subset W_{1,s}$, and let

$$\eta'_s = \frac{1}{M} \left\{ \int_0^{B_s} w \mathrm{d}F_s(w) - F_s(B_s)c_E + [1 - F_s(B_s)]B_s \right\} \geq \eta_s.$$

The above inequality is strict unless the contract with the firm is a debt contract. Moreover, the firm's expected payment is exactly the same conditional on each state under the alternative contract as in the original contract.

Now we consider the bank contract with the depositors. Given the bank contract with the depositors, (d, e), let d_s be the average amount paid to the depositors at state s. Strict incentive compatibility for the bank to monitor the firms requires $d_s < \eta_s$ for both $s = h, \ell$. We then design the new contract with the depositors as follows. Keep d as in the original contract and keep the fraction of verifying depositors the same as before for each states. If at state s no depositors verify the bank, then it must be the case that $\zeta_s \geq Mr$ and hence $B_s \geq Mr$. Then, (A3) implies that the bank has incentive to verify the firms at state s. Instead, if the bank is verified at state s and if the bank's available fund and the fraction of verifiers are consistent with bank revenue of ζ_s, give the same payments as the original contract. Otherwise, require the bank to pay off all its available funds. This ensures the bank has the same incentive to monitor the firms.

Finally, since the depositors receive the same amount of payments when verifying in both states and the probability of verifying remains the same, their incentive to verify remains. □

A.3 Feasibility for deposit insurance scheme

In Section 4.2 we consider a deposit insurance scheme that only requires budget-balancedness from the *ex ante* perspective. This would require an external lender who can commit to a contract with the regulator, who will receive the premium at good states and will pay for the insurance at bad states. Here we relax this assumption and consider two possible alternative assumptions. In the first we consider the possibility that the external lender cannot commit, and in the second we consider the case where there are many banks but there is no aggregate uncertainty when we aggregate across the banks.

External funding without lender commitment

Here we consider the situation where the external lender who provides funding for the deposit insurance cannot commit to his future actions. The regulator, however, can commit and offer a contract as follows. We assume that there is a single lender to simplify the analysis but, to be more realistic, we could have a continuum of identical lenders and each is offered the same contract. We assume that the lender is risk-neutral and has discount factor β.

If the bank is in state h, then the regulator pays ϕ (in per depositor term). If the bank is in state ℓ, then the lenders pays τ. Lack of commitment implies that the lender can walk away from the contract at any point of time. Obviously, the lender would not do so in state h. In state ℓ, however, the lender may choose to leave unless the future benefits from staying in the contract is better than leaving the contract. We assume that the regulator only pays the lender who has stayed in the contract always.

Thus, the incentive compatibility constraint for the lender to remain in the contract is given by

$$-\tau + \frac{\beta}{1-\beta}\left[\pi_h \phi - \pi_\ell \tau\right] \geq 0. \tag{32}$$

Here we show that even with this additional constraint, Theorem 4.1 still holds. The only modification needed in the argument is the construction of the deposit insurance scheme, and, instead of having $\pi_h \phi = \pi_\ell \tau$, we have:

$$\pi_h \phi = (\pi_\ell + r)\tau,$$

where $r = \frac{1-\beta}{\beta}$. This implies that the bank in state h pays $d'' < d'$ (the amount constructed in the proof of Theorem 4.1, but the difference converges to zero as β goes to one. Since we obtained strict Pareto improvement in that theorem, this implies that we can still obtain improvement with β sufficiently high. The incentives to monitor are not altered.

Oligopoly banks with idiosyncratic shocks

Suppose that there are continuum of depositors of measure $M \times N$, and there are N industries, each with measure one of firms. We assume that the stochastic returns to each firm are distributed according to $F_s(\cdot)$, where $s \in S = \{h, \ell\}$. While each industry is subject to shocks, the overall economy has no aggregate shock in the sense that the number of industries whose realize state ℓ is a constant, $\bar{n}$. Industry n's shock is ℓ if $\chi(n) \leq \bar{n}$, where χ is a permutation of $\{1, ..., N\}$, and each permutation occurs with the same probability, which is independent across time. Thus, for each individual bank, the probability of a high state is $\pi_h = \frac{N - \bar{n}}{N}$.

In this environment, the equation (9) has a different interpretation. Here π_h also refers to the fraction of banks in state h, and hence (9) requires the deposit insurance scheme to be self-financing: it uses the premium received from banks in state h to finance transfers to banks in state ℓ. We have the following corollary.

Corollary A.1. *Assume (A3). Let $(\mathbf{B}, \mathbf{d})$ be a dynamic bank contract that is not financially stable for each bank. Then, for β sufficiently high, there exists an (ex post) budget-balanced deposit insurance scheme and a bank contract that Pareto dominates it.*

Proof. Note that from depositors' perspective and the firms' perspective the problem is the same as facing a monopolist bank. The incentive to monitor firms is the same for a bank under state s is not changed either. We only need to show that now the budget is balanced for the deposit insurance scheme. Since in equilibrium $N - \bar{n}$ firms are in state h and pays ϕ each while the rest are in state ℓ and receive π each, budget balancedness requires

$$(N - \bar{n})\phi - \bar{n}\pi = 0,$$

which is equivalent to (9) with $\pi_h = \frac{N - \bar{n}}{N}$. □