Federal Reserve Bank of Chicago

Equilibrium Lending Mechanism and Aggregate Activity

By: Cheng Wang and Ruilin Zhou

WP 2000-30

Equilibrium Lending Mechanism and Aggregate Activity[*]

Cheng Wang
Graduate School of Industrial Administration
Carnegie Mellon University
Pittsburgh, PA 15213
chewang@wang.gsia.cmu.edu

and

Ruilin Zhou
Research Department
Federal Reserve Bank of Chicago
P.O. Box 834 Chicago, IL 60690-0834
rzhou@frbchi.org

December, 2000

Abstract

This paper develops a model of the credit market where the equilibrium lending mechanism, as well as the economy's aggregate investment and output, are endogenously determined. It predicts that the optimal contract is one of two kinds: either with intensive monitoring by investors to overcome entrepreneurs' incentive problems, such as most of intermediated financing, or with heavy reliance on entrepreneurs, such as market financing. We show that the observation that bank lending falls relative to corporate bond issuance during recessions can be explained by movements in the economy's real factors, such as a decline in average investment returns, a contraction of credit supply, and paradoxically, maybe even an increase of investment demand (which worsens credit market condition and intensifies incentive problems).

J.E.L Classification: E44, G32.

1. Introduction

What determines a firm's choice of its mechanism of investment financing? How is the choice of a firm's financing mechanism at the micro level related to investment and output at the aggregate level? These questions are at the center of recent discussions with respect to the nature and role of the so-called "credit crunch" which occurred during the most recent 1990-91 U.S. recession. During this recession, the economy's total outstanding loans fell dramatically and more important, the fraction of intermediated loans fell dramatically relative to unintermediated loans including public bonds and commercial paper (Friedman and Kuttner 1993).

There are very few existing theories of the relationship between the mechanism of financing and aggregate economic activity. Kashyap, Stein and Wilcox (1993) find that following a tightening of monetary policy, while there usually is a sharp increase in the amount of commercial paper outstanding, bank loans fall. They argue that monetary contraction tightens the supply of bank credit and hence forces borrowers to switch to commercial paper, and they further view this as evidence on the existence of a loan supply channel of monetary policy transmission. On the other hand, Bernanke, Gertler and Gilchrist (1996) advocate a flight to quality view on the same subject. They postulate both that demand for short-term credit is countercyclical, and that firms differ in their degree of access to credit markets. Thus, during a recession, high-grade firms borrow relatively easily by way of issuing commercial paper while low-grade firms, firms which can only borrow from banks, are constrained.

This paper develops a model of the credit market where the equilibrium lending mechanism, as well as the economy's aggregate investment and output, are endogenously determined. We then use the model to examine how the relationship between the equilibrium financing mechanism and aggregate output varies in response to disturbances to the model's exogenous variables. Suppose the economy receives a negative "real" shock. Specifically, suppose there is a decrease in the potential returns of an average investment project. Then the credit market may respond by switching from intermediated bank loans to unintermediated market lending. Meanwhile, fewer projects will be implemented, the success rates of the implemented projects will be higher (flight to quality), and total investment and output will both fall. Thus we provide a "real" explanation for the observation that economic downturns are often accompanied not only by contractions in total lending, but also by declines in the ratio of bank loans to non-bank lending.

Our theoretical findings are based on a lender (investor)-borrower (entrepreneur) relationship that features adverse selection, moral hazard and costly monitoring. In the model, adverse selection arises in that after the project is funded, the entrepreneur observes a random signal $\theta \in [0, 1]$ which indicates the project's success rate. This signal is private to the entrepreneur unless the investor pays a fixed cost to monitor. The project can then be liquidated or fully undertaken. In the latter case, the entrepreneur must make an unobservable effort to carry out the rest of the

investment process.

We show that the optimal contract has the following characteristics. First, it is always optimal to fully undertake projects with sufficiently high success rates. Let θ^* denote the cut-off level of the realization of the random signal θ below which the project is liquidated and above which the project is fully implemented. Second, either monitoring is never optimal, in which case any project with $\theta \geq \theta^*$ continues to be funded without being monitored; or monitoring is optimal, in which case there exists a second cut-off level of the success rate $\theta_n^* \in (\theta^*, 1]$ such that a project with a θ that falls between θ^* and θ_n^* is monitored. Third, the optimal compensation scheme is a debt contract if the contract prescribes no monitoring, and a combination of debt and equity contract otherwise.

If the optimal policy involves a positive probability of monitoring the entrepreneur, we brand the optimal contract as a form of bank lending; and, if the optimal contact involves no monitoring at all, we classify the optimal contract as market financing. This interpretation of the model is essential for our purpose. In practice, some business enterprises seek financing from financial intermediaries while others borrow directly from the credit market (e.g., commercial paper, corporate bond). A key distinction between the two financing mechanisms is that financial intermediaries often engage in extensive monitoring during the process of financing, whereas typical individual lenders do not monitor, or do so much less. A theoretical explanation for this distinction is that monitoring of private information is more effective when it is delegated to a financial intermediary rather than when done repetitively by individual lenders (Diamond 1984). The idea that banks are delegated monitors is central to the models of financial intermediation based on costly state verification (e.g., Williamson 1986, 1987). Recent studies on the choice of the optimal financing mechanism by Diamond (1991) and Holmstrom and Tirole (1997) have also taken seriously the notion that bank financing is closely related to monitoring. In both papers, financial intermediaries are modeled as monitors who can detect bad projects.

We now explain why a negative productivity shock can cause both the aggregate output and the ratio of intermediated loans to unintermediated loans to fall. In our model, it holds that in the absence of monitoring, the agency costs that must be incurred by the lender are higher if projects with lower success rates are undertaken. Now suppose the economy receives a productivity shock that lowers the return of a successful project. Then fewer projects should be fully funded (that is, θ^* should be higher). But this implies monitoring would be less efficient relative to no-monitoring, which in turn implies the ratio of intermediated loans to unintermediated loans would decrease. Meanwhile, because fewer projects are fully funded, total outstanding loans and aggregate output would be lower.

Our model is also rich enough to permit studies of other interactions between the credit market and the aggregate variables. In particular, in our model, it can be the case that the economy's total output is higher when it has less investment opportunities than when it has more

investment opportunities. This seemingly counterintuitive result can be explained as follows. When the economy is endowed with more investment opportunities, competition for loans will lower the equilibrium expected utility of the borrowers. This, given limited liability, makes the agency problem more severe, thereby causing more liquidation and less output. On the other hand, if the economy is endowed with less investment opportunities, competition for projects will shift the bargaining power from the lender to the borrower, thus raising the equilibrium expected utility of the borrower and lowering agency costs, resulting in less liquidation and more output.

Another lesson we learn from the model is that when the economy is experiencing a decline in bank lending, the economy's total output may rise or fall, depending on the source of the decline. Put differently, "credit crunch" is not necessarily bad news. It depends on what causes the crunch. We show that if the decline in bank lending is caused, say, by an increase in the cost of monitoring or by a decrease in the potential returns of the project, then total output falls as bank lending declines. If the decline in bank lending is caused by a decrease in the economy's endowment of investment projects, then under some conditions the economy's total output could increase while total bank loans fall.

An important feature of our model is that whenever there is a shortage of funds, in equilibrium there is *always* credit rationing of the type discussed by Stiglitz and Weiss (1981) and Williamson (1986, 1987), where among a group of identical borrowers, those who receive loans are strictly better off than those who do not. Credit rationing in our model is motivated sometimes by costly monitoring (as in Williamson 1987) and sometimes by costly over-liquidation. A lower reservation utility of the borrower may imply that his project must be liquidated with an excessively higher probability, which lowers the lender's expected returns on a loan. The notion that credit rationing is a mechanism to avoid excessive liquidation has not been discussed in the literature.

This paper builds on the large literature in contract theory that follows Townsend (1979) in modeling the role of costly monitoring in optimal financial arrangements, including Gale and Hellwig (1985), Williamson (1986, 1987), and Boyd and Smith (1997). At the heart of our model is the interaction between the investor's optimal monitoring policy and optimal financing strategy. In which states of the project should the investor monitor, and what happens subsequently? Could it be optimal that in some states the project is not monitored but fully financed, whereas in other states the project is monitored but subsequently abandoned? These questions, though obviously important for the study of investment financing, have not been addressed explicitly by the existing literature. Holmstrom and Tirole (1998) also model the optimal liquidation decision conditional on the realization of a random signal (a liquidity shock in their environment). But they abstract from the problem of costly monitoring by assuming the random shock is observable (or when it is not, it still does not affect the structure of the optimal contract). Admati and Pfleiderer (1994) have a model which is somewhat similar to ours and they assume monitoring is not costly. Modeling explicitly the process of costly monitoring allows us to study the interaction between

costly information acquisition and the investment financing decision. But as we will show, solving the optimization problem is by no means a trivial task.

Section 2 presents the model. In Section 3, we study the two-agent optimal contract assuming a certain credit market outcome. Section 4 embeds the optimal contract in a perfect competitive credit market, and analyses the market equilibrium. It then considers the implications of the model's comparative statics. Section 5 concludes the paper.

2. The Model

There are three periods, $\tau = 0, 1, 2$. There are two types of agents, investors and entrepreneurs, and there is a continuum of each type such that the measure of the investors is λ, and that of the entrepreneurs is δ. All agents are risk neutral. Investors maximize their expected consumption in period 2, entrepreneurs maximize the expected value of $u(c, e) = c - e$, where c is consumption in period 2 and e is effort exerted in period 1.
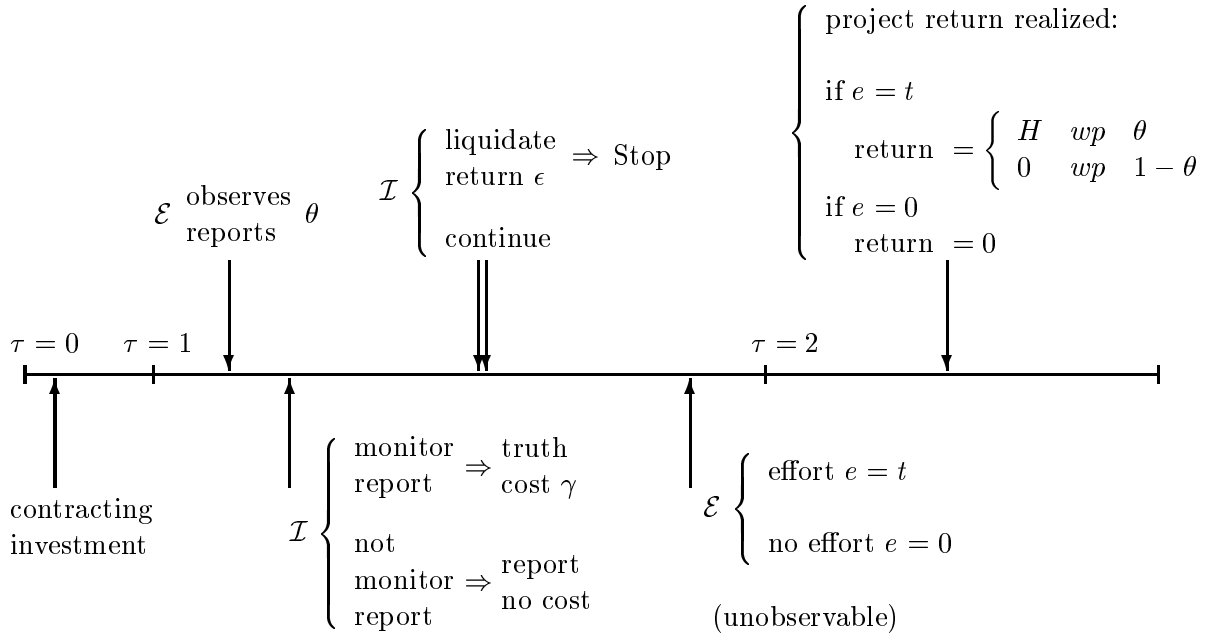
In period 0, each investor has one indivisible unit of investment good, which can either be invested in the credit market which matches worthy projects (entrepreneurs) with investment goods (investors), or earn a certain gross return of one unit of consumption in period 2 through storage. Each investor also has access to $\xi(> 0)$ units of the consumption good in period 2. We will assume that ξ is large enough to fulfill all payments specified by financial contract. Each entrepreneur owns a risky investment project, which requires an investment of one unit of the investment good in period 0, or it simply perishes. No entrepreneur has any initial wealth, and hence he must rely on external financing in order to undertake his project. All investors are ex ante identical, so are all entrepreneurs (hence, their projects).

In period 0, there is a competitive credit market in which investors offer lending contracts to entrepreneurs who exchange investment opportunities for credit and compensation. Given that all agents are risk neutral, and all projects are identical ex ante, without loss of generality, we assume that each entrepreneur obtains funds from at most one investor, and each investor invests in at most one project. That is, any contract is formed exclusively between one entrepreneur and one investor. At the end of period 0, there may be projects unfunded or investment goods unused, depending on the size of each side of the market δ and λ. All contracts traded are identical at a credit market equilibrium, each promises an expected utility equal to u_0^* to the entrepreneur party. This equilibrium expected utility of the entrepreneur will be determined endogenously.

Without loss of generality, we discuss the contracting and investment problem between two generic agents of each type: an investor $\mathcal{I}$ and an entrepreneur $\mathcal{E}$. Assume that $\mathcal{E}$'s project is worth funding and is funded by $\mathcal{I}$. The timing of the events unfold as indicated in Figure 1. At the beginning of period 1 the entrepreneur $\mathcal{E}$ observes a signal θ. Here $\theta \in [0, 1] \equiv \Theta$ is a random variable that represents the potential success rate of the project. We assume that θ is a

4

continuous random variable on Θ with a distribution function $G(\theta)$ and a density $g(\theta)$, and for all $\theta \in \Theta$, $g(\theta) > 0$. The signal θ is directly observable only to the entrepreneur. The investor $\mathcal{I}$ can observe the realization of θ through a costly monitoring process, which requires $\gamma \geq 0$ units of $\mathcal{I}$'s effort. Here, γ may be interpreted as effort required to discover the project's technical feasibility and market profitability. If the investor monitors, then she learns the true realization of θ. Otherwise, she knows only the entrepreneur's report.

Figure 1. The Timing of a Project Development

After the observation of the signal θ by at least one party, a decision must be made as for whether to continue the project or liquidate it. The liquidation value of the project is $\epsilon \geq 0$, measured in units of period 2 consumption. If the investor $\mathcal{I}$ monitors $\mathcal{E}$'s report of θ, then of course the liquidation/continuation decision can be based on the true realization of θ. Otherwise, the liquidation/continuation decision may take into account only $\mathcal{E}$'s report. Suppose the investment is continued, then the entrepreneur $\mathcal{E}$ makes an unobserved effort of either $t > 0$ or zero. In other words, there is moral hazard.

If the project is continued, at $\tau = 2$, the return is realized. If $\mathcal{E}$ makes the required effort t to his continued project, then with probability θ the project succeeds with return $H > 0$, and with probability $1 - \theta$ it fails and yields nothing. If $\mathcal{E}$ does not make the required effort, then the project fails with probability one. We call H the potential return of the project, and θ the project's success rate. At the end of period 2, the contract ends with transferring the specified payment,

5

depending on the commonly observed events occurred during the project development process (monitoring, continuation/liquidation, and project return), from the investor to the entrepreneur. The investor receives the residual return of the project.

We make some further assumptions. First, all payments to the entrepreneur must be non-negative (limited liability). Second, renegotiation is not allowed. In other words, we assume that once the contract is signed, both parties can fully commit, at all stages of the investment process, to the terms of the initial contract. Third, a necessary condition that a project is worthy of investment ex ante is that at $\tau = 1$, with the best possible realization of $\theta \ (= 1)$, the project return H is higher than all the potential costs from then on. That is,

Assumption (1) $t + \epsilon + \gamma \leq H$.

Finally, we make a technical assumption to guarantee the uniqueness of solution to the optimal contracting problem in the sections to follow.

Assumption (2)
$$\frac{-H}{H - (t + \epsilon)} \leq \frac{g'(\theta)}{g(\theta)} \leq \frac{H}{t + \epsilon + \gamma}.$$
Clearly, there is a wide range of distribution functions with support $[0, 1]$, including the uniform distribution, which satisfy the above condition.

In what follows, we first investigate the optimal contract between an investor and an entrepreneur, assuming a certain outcome of credit market competition. We then use the obtained optimal contract to study credit market equilibrium.

3. Two-Agent Optimal Contract

In this section, we determine the form of the optimal contract between a representative investor $\mathcal{I}$ and entrepreneur $\mathcal{E}$ pair. Suppose that the period-0 credit market competition yields that an entrepreneur's compensation from his funded project is at least $u_0 \geq 0$ in terms of expected utility. The optimal contract, then, maximizes the investor's expected utility subject to the constraint that the entrepreneur's expect payoff is no less than u_0. As u_0 varies, the contract moves along the Pareto frontier between the two agents.

3.1. The First-Best Contract

Consider the case where both the realization of θ and entrepreneur $\mathcal{E}$'s effort t are publicly observable. Given that there is no private information and moral hazard problem, entrepreneur $\mathcal{E}$ can be compensated with a fixed payment, denote it x. In addition, a contract must specify a liquidation/continuation policy Φ, a subset of Θ: if $\theta \in \Phi$, then the project is continued;

otherwise, it is liquidated. Given the environment, the optimal contract must implement t as the entrepreneur's effort.

Let Φ' denote the complement of the set Φ. The problem of optimal contracting can then be formulated as follows.

$$(P0) \qquad \max_{x,\Phi} \int_{\Phi} \theta H dG(\theta) + \int_{\Phi'} \epsilon dG(\theta) - x \qquad (1)$$

subject to

$$x - \int_{\Phi} t dG(\theta) \geq u_0. \qquad (2)$$

The objective function (1) represents investor $\mathcal{I}$'s expected payoff.[1] Condition (2) is entrepreneur $\mathcal{E}$'s participation constraint: his expected return is no less than the expected credit market payoff u_0. Clearly, constraint (2) must be binding, since otherwise reducing the value of x can improve $\mathcal{I}$'s expected payoff while holding the participation constraint satisfied. By substituting constraint (2) into the objective function, we can rewrite the optimal contracting problem as

$$\max_{\Phi} \int_{\Phi} \Big(\theta H - t\Big) dG(\theta) + \int_{\Phi'} \epsilon dG(\theta) - u_0. \qquad (3)$$

Obviously, the optimal Φ must be an upper interval of Θ.[2] Let this interval be $[\theta_{\text{fb}}, 1]$, $\theta_{\text{fb}} = argmax_{x \in [0,1]} F(x)$ where

$$F(x) \equiv \int_x^1 \Big(\theta H - t\Big) dG(\theta) + \epsilon G(x).$$

It can be shown that function $F(x)$ is strictly concave under assumption (2),[3] and therefore the maximization problem (3) has a unique solution:

$$\Phi_{\text{fb}} = [\theta_{\text{fb}}, 1], \quad \text{where} \quad \theta_{\text{fb}} = \frac{t + \epsilon}{H}. \qquad (4)$$

Given Φ_{fb}, $\mathcal{E}$'s compensation is determined by $x = u_0 + (1 - G(\theta_{\text{fb}}))t$, and $\mathcal{I}$'s expected payoff (expected net returns on an investment) is given by

$$V_{\text{fb}} = H \int_{\theta_{\text{fb}}}^1 (\theta - \theta_{\text{fb}}) dG(\theta) + \epsilon - u_0 - 1. \qquad (5)$$

---

[1] For convenience, we omit the constant unit cost of date-0 investment in all of the objective functions.

[2] If a project with a lower success rate θ is continued, then a project with a higher θ should also be continued.

[3] We have $F'(x) = -[xH - (t + \epsilon)]g(x)$. Obviously, $F'(0) > 0$, $F'(1) < 0$. Also, $F''(x) = -[xH - (t+\epsilon)]g'(x) - Hg(x)$. So for the function $F(x)$ to be concave, we need

$$-\frac{g'(x)}{g(x)}\Big[xH - (t + \epsilon)\Big] < H.$$

Now if $xH - (t + \epsilon) \geq 0$, then the above inequality certainly holds. If $xH - (t + \epsilon) < 0$, then the concavity condition becomes

$$\frac{g'(x)}{g(x)} < \frac{H}{(t + \epsilon) - xH},$$

which holds, by the second inequality of assumption (2).

Note that given both parties are risk neutral, it is straightforward to show the following holds. The first-best outcome is achievable when there is only moral hazard concerning the entrepreneur's effort, but no information asymmetry with respect to the project's success rate θ. The first-best outcome is also achievable if there is only asymmetric information concerning the success rate θ, but there is no moral hazard.

3.2. Contract with Costly Monitoring and Moral Hazard

Now, we consider our original model where the realization of the project's success rate θ is directly observable only to entrepreneur $\mathcal{E}$. Investor $\mathcal{I}$ can observe the value of θ at a cost $\gamma > 0$. Moreover, there is moral hazard: entrepreneur $\mathcal{E}$'s effort is not observable. In this standard finite-horizon, two-person game with adverse-selection and moral-hazard problems, the revelation principle applies. That is, every equilibrium allocation of any arbitrary mechanism can be implemented as an equilibrium of a revelation mechanism. Therefore, we will focus on incentive compatible mechanisms to characterize the optimal contract.

3.2.1. The Definition of contract

With costly monitoring and moral hazard, there are now three components to a loan contract: (i) a monitoring policy M for verifying the state of the success rate θ, (ii) a liquidation/continuation policy Φ which determines whether or not the project is liquidated after the realization of the state θ, and (iii) a scheme for state contingent compensations to the entrepreneur. Formally, a contract takes the following form:

$$\sigma = \left\{ M; \; \Phi; \; x; \; y(\tilde{\theta}), \tilde{\theta} \notin \Phi; \; R_0(\tilde{\theta}), R(\tilde{\theta}), \tilde{\theta} \in \Phi \right\}.$$

We abstract throughout from stochastic monitoring; thus, the monitoring policy M is a subset of Θ in which verification of the reported state will occur. That is, let $\hat{\theta}$ denote the entrepreneur's report of θ, then monitoring takes places if and only if $\hat{\theta} \in M$.

The liquidation/continuation policy Φ is also a subset of Θ. Unlike in the case of complete information, here Φ as well as compensation schedule must take into account the fact that there is information asymmetry between investor $\mathcal{I}$ and entrepreneur $\mathcal{E}$ concerning the realization of θ. Let $\tilde{\theta}$ denote investor $\mathcal{I}$'s knowledge of the realization of θ on which the liquidation/continuation decision must be conditioned:

$$\tilde{\theta}(\hat{\theta}, \theta) = \begin{cases} \theta, & \text{if } \hat{\theta} \in M, \\ \\ \hat{\theta}, & \text{otherwise.} \end{cases}$$

Then the project is continued if $\tilde{\theta} \in \Phi$, and it is liquidated if $\tilde{\theta} \notin \Phi$.

In the state of liquidation, investor $\mathcal{I}$ seizes the scrap value of the project ϵ, and entrepreneur $\mathcal{E}$ receives a payment $y(\tilde{\theta})$. Conditional on the project being continued, entrepreneur $\mathcal{E}$ is paid $R_0(\tilde{\theta}) \geq 0$ if the project eventually fails; and he is paid $R_0(\tilde{\theta}) + R(\tilde{\theta}) \geq 0$ if the project succeeds with realized return H. Finally, the contract specifies a fixed payment $x \geq 0$ that entrepreneur $\mathcal{E}$ receives in period 2. This payment is not contingent on the state of θ nor the realization of the project's random return.[4]

Given that monitoring is costly, it may not be efficient to always monitor. To simplify matters, we assume that if entrepreneur $\mathcal{E}$ is indifferent between reporting truthfully and lying, he reports truthfully. Therefore, if both the realization θ and the report $\hat{\theta}$ are in the monitoring region M, there is no point to lie. Furthermore, for any realization θ not in the monitoring region M, entrepreneur $\mathcal{E}$ gains nothing by giving a false report in M to induce monitoring. Since in either case, $\mathcal{E}$'s payoff will depend on the truth only. This implies that entrepreneur $\mathcal{E}$ will not submit a false report of θ for monitoring. In other words, we have:

Lemma 1. *If $\hat{\theta} \in M$, then $\hat{\theta}(\theta) = \theta$.*

As mentioned in the introduction, a focus of this paper is the joint determination of the optimal monitoring policy M and liquidation/continuation policy Φ. It is thus useful to define the following subsets of Θ:

$$A \equiv \Phi \cap M, \quad B \equiv \Phi' \cap M, \quad C \equiv \Phi \cap M', \quad D \equiv \Phi' \cap M',$$

where Φ' and M' are the complements of Φ and M, respectively. By Lemma 1, if $\mathcal{E}$'s report of the state $\hat{\theta}$ is in $A[B]$, then monitoring will occur and the project will [will not] continue. On the other hand, if $\hat{\theta} \in C[D]$, then the project is not monitored and it will [will not] continue.

Now consider the set D, the non-monitoring/liquidation region. Suppose that $\theta_1, \theta_2 \in D$ and $y(\theta_1) > y(\theta_2)$. Then, whenever θ_2 is realized, $\mathcal{E}$ could lie and report θ_1 to get the higher payoff $y(\theta_1)$, given that both θ_1 and θ_2 are not monitored. This implies that $y(\theta)$ must be constant on D in order for the contract to be incentive compatible.

Lemma 2. *An incentive compatible contract satisfies $y(\theta) = Y_D$ for all $\theta \in D$.*

Given Lemmas 1 and 2, an incentive compatible contract has to satisfy the following three sets of incentive constraints. First, there should be no incentives for entrepreneur $\mathcal{E}$ to report untruthfully a $\hat{\theta} \in C$ in order to continue the investment process without monitoring (conditions (6) and (8)). Second, there should be no incentives for entrepreneur $\mathcal{E}$ to report untruthfully a $\hat{\theta} \in D$ so that a good project is abandoned to avoid making effort or to receive a better compensation Y_D (conditions (7) and (9)). Third, there should be no incentives for entrepreneur

---

[4]Clearly, x is a mathematically redundant component of the contract, and we have introduced x only for analytical convenience.

$\mathcal{E}$ to shirk whenever the project is continued (condition (10)). Formally,

Truth-telling constraints:

$$\forall \theta \in A \cup C, \ \forall \hat{\theta} \in C \qquad \theta R(\theta) + R_0(\theta) - t \geq \max\left\{\theta R(\hat{\theta}) + R_0(\hat{\theta}) - t, \ R_0(\hat{\theta})\right\} \qquad (6)$$

$$\theta R(\theta) + R_0(\theta) - t \geq Y_D \qquad (7)$$

$$\forall \theta \in B \cup D, \ \forall \hat{\theta} \in C \qquad y(\theta) \geq \max\left\{\theta R(\hat{\theta}) + R_0(\hat{\theta}) - t, \ R_0(\hat{\theta})\right\} \qquad (8)$$

$$\forall \theta \in B \qquad y(\theta) \geq Y_D \qquad (9)$$

Effort constraint:

$$\forall \theta \in A \cup C \qquad \theta R(\theta) + R_0(\theta) - t \geq R_0(\theta) \qquad (10)$$

The entrepreneur's participation constraint is as follows,

$$x + \int_{A \cup C}\left(\theta R(\theta) + R_0(\theta) - t\right) dG(\theta) + \int_{B \cup D} y(\theta) dG(\theta) \geq u_0. \qquad (11)$$

We are now in a position to define optimality. We call a contract optimal if it maximizes investor $\mathcal{I}$'s expected payoff, subject to the incentive constraints, the participation constraint, and the limited-participation constraint for entrepreneur $\mathcal{E}$. That is, an optimal contract solves the following problem,

$$(P1) \qquad \max_{\sigma} \int_{A \cup C} [\theta H - \theta R(\theta) - R_0(\theta)] \, dG(\theta) + \int_{B \cup D} [\epsilon - y(\theta)] dG(\theta) \qquad (12)$$
$$-\mu(A \cup B)\gamma - x$$

$$\text{subject to} \qquad (6)\text{---}(11)$$

$$x \geq 0, \qquad \forall \theta \in B \cup D \quad y(\theta) \geq 0,$$

$$\forall \theta \in A \cup C \quad R_0(\theta) \geq 0, \quad R_0(\theta) + R(\theta) \geq 0 \qquad (13)$$

where μ denotes the probability measure on Θ: for any set $Z \subseteq \Theta$, $\mu(Z) = \int_Z dG(\theta)$.

3.2.2. The optimal contract

We now set out to analyze the properties of the optimal contract. Our first task is to simplify the incentive constraints. The approach we take is to consider a class of optimal contracts, all of which deliver the same expected utilities to both entrepreneur $\mathcal{E}$ and investor $\mathcal{I}$, and then show that each contract in that class is equivalent to a contract whose compensation scheme resembles that of a debt or an equity contract. In the following, any two contracts are said to be equivalent if they satisfy the same set of constraints and promise the same expected payoffs to both the investor and the entrepreneur.

Proposition 1. *For any contract σ that solves $(P1)$, there exists a contract $\hat{\sigma}$ which is equivalent to σ , and $\hat{\sigma}$ has the following properties: for all $\theta \in A \cup C$, $R_0(\theta) = 0$, and for all $\theta \in C$, $R(\theta) = R_C$, where $R_C \geq 0$ is a constant.*

Proposition 1 implies that we can focus on the set of contracts which have a relatively simple compensation structure: conditional on the project being continued, the entrepreneur's compensation is zero if the project fails. Moreover, if the project succeeds, and if there is no monitoring, entrepreneur $\mathcal{E}$'s compensation is independent of his report of θ. The intuition for this result is simple. The debt structure is efficient here partly because it imposes the largest possible punishment for a bad outcome. The constant compensation on C is required by truth-telling constraint. The technical proof of this proposition, however, is somewhat involved because of the tangled truth-telling and effort-making incentive constraints. The proof of Proposition 1 is in the appendix.

Proposition 1 allows us to focus on a set of simpler contracts where the compensation schemes are debt-looking in the continuation regions A and C. Note that by constraint (10), we have $R(\theta) \geq t/\theta > 0$ for any θ in A or C, which implies the non-negativity of $R(\theta)$ on A and R_C. Hence, the optimal contracting problem can be simplified as follows:

$$(P2) \quad \max_{\sigma} \int_A [\theta H - \theta R(\theta)] \, dG(\theta) + \int_C \left(\theta H - R_C\right) dG(\theta) + \int_B (\epsilon - y(\theta)) dG(\theta) \quad (14)$$
$$+ (\epsilon - Y_D)\mu(D) - \mu(A \cup B)\gamma - x$$

subject to

$$x \geq 0; \quad \forall \theta \in B \ y(\theta) \geq 0; \quad Y_D \geq 0, \quad (15)$$
$$x + \int_A (\theta R(\theta) - t) dG(\theta) + \int_C (\theta R_C - t) dG(\theta) + \int_B y(\theta) dG(\theta) + Y_D \mu(D) \geq u_0 \quad (16)$$

and the following set of incentive constraints,

$$\forall \theta \in A \qquad \theta R(\theta) - t \geq \theta R_C - t, \quad (17)$$
$$\theta R(\theta) - t \geq Y_D, \quad (18)$$
$$\forall \theta \in B \qquad y(\theta) \geq \theta R_C - t, \quad (19)$$
$$y(\theta) \geq Y_D, \quad (20)$$
$$\forall \theta \in C \qquad \theta R_C - t \geq Y_D, \quad (21)$$
$$\forall \theta \in D \qquad Y_D \geq \theta R_C - t, \quad (22)$$
$$\forall \theta \in A \qquad \theta R(\theta) \geq t \quad (23)$$
$$\forall \theta \in C \qquad \theta R_C \geq t. \quad (24)$$

The next proposition shows that the optimal liquidation/continuation policy is monotonic. That is, if a project with success rate θ is continued, then any project with a higher rate of success is also continued.

Proposition 2. *For any optimal contract σ that solves problem $(P2)$, there exists a contract $\hat{\sigma}$ satisfying: for all $\theta_1 \in \Phi$ and $\theta_2 \in \Phi'$, $\theta_1 > \theta_2$ and $\hat{\sigma}$ is equivalent to σ.*

Suppose a project with success rate θ_1 is continued, but a project with success rate θ_2 $(\theta_2 > \theta_1)$ is liquidated. Then by switching the positions of θ_1 and θ_2, and by re-arranging the compensation schemes properly, one can achieve a Pareto improvement. The proof of Proposition 2 is in the appendix.

Given Proposition 2, we can then focus, without loss of generality, on contracts with monotonic liquidation/continuation policies; that is, contracts in which the set Φ is an upper interval of Θ. Our next lemma shows that for optimality, this upper interval must not be empty.

Lemma 3. *If σ is an optimal contract, then $\Phi' \neq \emptyset$.*

The proof of Lemma 3 is in the appendix. The next contract specifies the main structure of optimal contract.

Proposition 3. *An optimal contract has the following characteristics:*

(i) *$B = \emptyset$ and $\Phi' = D$.*

(ii) *There are constants θ_m and θ_n, $0 < \theta_m \leq \theta_n \leq 1$ such that*

$$A = [\theta_m, \theta_n), \quad C = [\theta_n, 1], \quad D = [0, \theta_m].$$

(iii) *Moreover, the following compensation scheme is optimal:*

$$\forall \theta \in A = [\theta_m, \theta_n) \ R(\theta) = t/\theta, \quad \forall \theta \in C = [\theta_n, 1] \ R_C = t/\theta_n, \quad Y_D = 0.$$

By Proposition 3, it is never optimal to have the project monitored and then abandoned. Moreover, the optimal monitoring strategy is to monitor those success rate θ which are neither too low, nor too high. Put differently, it is optimal not to monitor when the "news" from entrepreneur $\mathcal{E}$ is sufficiently good or sufficiently bad. The entrepreneur's compensation is zero when the project is abandoned and when the project is continued but fails. When the project is continued and succeeds with return H, entrepreneur $\mathcal{E}$'s compensation is nonlinear in the realization of θ: it is relatively high but decreasing in the success rate θ in the region where monitoring occurs, and it is low (but positive) and constant across the region where monitoring does not occur.

Note that the optimal monitoring strategy is not monotonic over the whole state space Θ, although it is monotonic conditional on the investment being continued. Also, given that the investment is continued, entrepreneur $\mathcal{E}$'s *expected* net compensation is monotonic and piecewise linear in θ: it is zero for all $\theta \in A$ and $(t/\theta_n)\theta > 0$ for all $\theta \in C$.

The intuitions for Proposition 3 are as follows. There is no need to monitor a project that is to be abandoned $(B = \emptyset)$, since it is a waste of resources without any return. Conditional on continuation, it is optimal to monitor lower rather than higher reports of θ because doing so minimizes the cost of monitoring. To see this, suppose there is a θ' in A that is greater than the lowest θ in C. Then continuing a project with report θ' without monitoring (that is, move θ' into

C) can result a net gain for investor $\mathcal{I}$ without violating any other constraints. Entrepreneur $\mathcal{E}$'s expected payoff can be maintained by increasing the fixed payment x by the amount of reduction at θ' from $R(\theta')$ to R_C. The change is also incentive compatible. First, the effort-making constraint is satisfied at θ' since it is satisfied at the lowest θ in C. Second, there are no incentives for the entrepreneur to misreport a realization θ' as some other θ'' in D, since his payoff in the liquidation stage Y_D is lower than $\theta' R_C - t$, his expected payoff if the project continues. Third, there are no incentives for him to misreport a different θ'' in C other than θ', since the payoff is the same with both reports on C. However, this change strictly improves investor $\mathcal{I}$'s expected payoff because it reduces monitoring cost in state θ'.

To explain part (iii), notice that it is optimal to set the entrepreneur's compensation, Y_D in the liquidation region D and $R(\theta)$ and R_C in the continuation regions A and C, just high enough that proper incentives are given for truthful reporting and effort making. Holding the levels of Y_D, $R(\theta)$ and R_C too high is potentially costly: it may cause the entrepreneur's overall expected compensation to exceed his reservation utility, since $x \geq 0$ must hold. Now given (i) and (ii), it is easy to check that the compensation scheme specified in the proposition is the lowest possible that still satisfies all the incentive constraints.

Given Proposition 3, the optimal contract is characterized fully by variables x, θ_m, and θ_n, with $x \geq 0$ and $\theta_m \leq \theta_n$. Hence, the optimal contracting problem ($P2$) can be rewritten as follows:

$$(P3) \quad \max_{x, \theta_m, \theta_n} \int_{\theta_m}^{\theta_n} \Big(\theta H - t - \gamma\Big) dG(\theta) + \int_{\theta_n}^{1} \Big(\theta H - \theta \frac{t}{\theta_n}\Big) dG(\theta) + \epsilon G(\theta_m) - x \qquad (25)$$

subject to

$$x + T(\theta_n) \geq u_0, \qquad (26)$$

$$x \geq 0, \quad \theta_m, \theta_n \in [0,1], \quad \theta_m \leq \theta_n, \qquad (27)$$

where

$$T(\theta_n) \equiv \int_{\theta_n}^{1} \Big(\theta \frac{t}{\theta_n} - t\Big) dG(\theta). \qquad (28)$$

The term $T(\theta_n)$ is the lowest reservation utility of entrepreneur $\mathcal{E}$ in order to induce his full effort for all continued projects in non-monitoring region $[\theta_n, 1]$. We have

$$T'(\theta_n) = -\frac{t}{(\theta_n)^2} \int_{\theta_n}^{1} \theta dG(\theta) < 0.$$

Thus, in the absence of monitoring, more must be promised to entrepreneur $\mathcal{E}$ if more projects are continued.

Obviously, the problem ($P3$) has a solution, denote it $\{x^*, \theta_m^*, \theta_n^*\}$. To solve problem ($P3$), we consider the cases where the entrepreneur's participation constraint (26) is binding and slack separately.

First, suppose that constraint (26) is binding at the optimum. Then, substituting $x = u_0 - T(\theta_\mathsf{n})$ into the objective function (25) and maximize it without constraint (27), we have

$$\theta_\mathsf{m}^* = \theta_\mathsf{n}^* = \theta_\mathsf{fb} = \frac{1}{H}(t + \epsilon) \tag{29}$$

and

$$x^* = u_0 - T(\theta_\mathsf{fb}). \tag{30}$$

That is, no monitoring is necessary, and the optimal cutoff level for project continuation coincides with the first-best continuation level θ_fb. Of course, for $\{x^*, \theta_\mathsf{m}^*, \theta_\mathsf{n}^*\}$, given by (29) and (30), to be the solution, it has to satisfy $x^* > 0$, given that $\theta_\mathsf{fb} \in [0, 1]$ by assumption (1). That is,

$$T(\theta_\mathsf{fb}) \leq u_0. \tag{31}$$

When condition (31) holds, $\{x^*, \theta_\mathsf{m}^*, \theta_\mathsf{n}^*\}$ given by (29) and (30) is the solution to $(P3)$. In fact, condition (31) is the necessary and sufficient condition for this to be the solution and for the first-best continuation/liquidation to be achievable.

Next, suppose that condition (31) does not hold. Then it must hold that $x^* = 0$, for otherwise, given $T'(\theta_\mathsf{n}) < 0$, it is possible to reduce the values of x and θ_n simultaneously and to increase the investor's expected payoff without violating the participation constraint (26). With $x^* = 0$, constraint (26) becomes $T(\theta_\mathsf{n}) \geq u_0$. Let $\bar{\theta}_\mathsf{n}$ be the level of θ at which the constraint binds, that is,

$$T(\bar{\theta}_\mathsf{n}) = u_0. \tag{32}$$

Since $u_0 = T(\bar{\theta}_\mathsf{n}) < T(\theta_\mathsf{fb})$, and T is decreasing in θ_n, we have $\theta_\mathsf{fb} < \bar{\theta}_\mathsf{n} \leq 1$. Also because $T(\theta_\mathsf{n})$ is decreasing, the participation constraint $T(\theta_\mathsf{n}) \geq u_0$ is equivalent to $\theta_\mathsf{n} \leq \bar{\theta}_\mathsf{n}$. Now we can rewrite the optimal contracting problem $(P3)$ as follows:

$$(P4) \quad \max_{\theta_\mathsf{m}, \theta_\mathsf{n}} O(\theta_\mathsf{m}, \theta_\mathsf{n}) \equiv \int_{\theta_\mathsf{m}}^1 \Big(\theta H - t\Big) dG(\theta) + \epsilon G(\theta_\mathsf{m}) - T(\theta_\mathsf{n}) - \int_{\theta_\mathsf{m}}^{\theta_\mathsf{n}} \gamma dG(\theta) \tag{33}$$

$$\text{subject to} \quad \theta_\mathsf{m}, \theta_\mathsf{n} \in [0, 1], \quad \theta_\mathsf{n} \geq \theta_\mathsf{m}, \quad \theta_\mathsf{n} \leq \bar{\theta}_\mathsf{n}. \tag{34}$$

For future reference, let $(\hat{\theta}_\mathsf{m}, \hat{\theta}_\mathsf{n})$ be the solution for the unconstraint problem (33). Then,

$$\hat{\theta}_\mathsf{m} = \frac{t + \epsilon + \gamma}{H} \tag{35}$$

$$\frac{t}{(\hat{\theta}_\mathsf{n})^2} \int_{\hat{\theta}_\mathsf{n}}^1 \theta dG(\theta) - \gamma g(\hat{\theta}_\mathsf{n}) = 0. \tag{36}$$

Furthermore, let $\tilde{\theta}_\mathsf{n}$ be the solution to the following equation,

$$\tilde{\theta}_\mathsf{n} = \frac{1}{H}(t + \epsilon) + \frac{t}{(\tilde{\theta}_\mathsf{n})^2 \, g(\tilde{\theta}_\mathsf{n}) H} \int_{\tilde{\theta}_\mathsf{n}}^1 \theta dG(\theta). \tag{37}$$

It is easy to show that $\theta_{fb} < \hat{\theta}_m \leq 1$, $0 \leq \hat{\theta}_n < 1$, and $\theta_{fb} < \tilde{\theta}_n < 1$.

Problem $(P4)$ is well defined: the objective function $O(\theta_m, \theta_n)$ is strictly concave, and the constraint set defined by (34) is convex. So it has a unique solution. The detailed solution is given in the appendix. The following proposition summarizes the solution for problem $(P3)$.

Proposition 4. *The optimal contract takes one of the following three forms:*

(i) *If* $u_0 \geq T(\theta_{fb})$, *then*
$$\sigma_n^{fb} : \quad \theta_n^* = \theta_{fb}, \; M^* = \emptyset, \; \Phi^* = [\theta_{fb}, 1], \; and \; x^* = u_0 - T(\theta_{fb}).$$

(ii) *If* $u_0 < T(\theta_{fb})$, *and* $\hat{\theta}_m < \min\{\hat{\theta}_n, \bar{\theta}_n\}$, *then* $\theta_m^* = \hat{\theta}_m$, $\theta_n^* = \min\{\hat{\theta}_n, \bar{\theta}_n\}$,
$$\sigma_m^{sb} : \quad \theta_{fb} < \theta_m^* < \theta_n^*, \; M^* = [\theta_m^*, \theta_n^*), \; \Phi^* = [\theta_m^*, 1], \; and \; x^* = 0,$$

(iii) *If* $u_0 < T(\theta_{fb})$ *and* $\hat{\theta}_m \geq \min\{\hat{\theta}_n, \bar{\theta}_n\}$, *then* $\theta_m^* = \theta_n^* = \min\{\tilde{\theta}_n, \bar{\theta}_n\}$,
$$\sigma_n^{sb} : \quad \theta_{fb} < \theta_n^*, \; M^* = \emptyset, \; \Phi^* = [\theta_n^*, 1], \; and \; x^* = 0.$$

Proposition 4 states the precise conditions under which the optimal contract takes what particular form. One critical element is u_0, the reservation utility of entrepreneur $\mathcal{E}$. If u_0 is high enough, the optimal contract σ_n^{fb} achieves the first-best outcome without monitoring. When u_0 is below a certain threshold; $u_0 < T(\theta_{fb})$, only second-best outcome can be obtained. In such a case, the optimal contract may be one with some monitoring (reports in region $M^* = [\theta_m^*, \theta_n^*)$, contract σ_m^{sb}), or without monitoring at all (contract σ_n^{sb}). In either cases, the cut-off level for continuation (θ_m^* under σ_m^{sb}, θ_n^* under σ_n^{sb}) is the higher than that of the first-best solution. That is,

Corollary 4.1. *Whenever the first-best is not attainable, there is always over-liquidation at the optimum.*

Proposition 4 also shows that unless the first-best is attainable, the non-state-contingent component of the compensation x must be zero. Thus, when the first-best is not attainable, the optimal contract takes the form of debt (σ_n^{sb}) or combined debt and equity (σ_m^{sb}) contract. Moreover, at the optimum, the entrepreneur earns positive compensation only in states where the project is carried out without being monitored; that is, only when the project is sufficiently good and when the project is ultimately successful. Given (θ_m^*, θ_n^*), the optimal payoff schedule for entrepreneur $\mathcal{E}$ is given by Proposition 3. That is,

$$\forall \theta \in [0, \theta_m^*) \; y^*(\theta) = 0, \qquad \forall \theta \in [\theta_m^*, \theta_n^*) \; R^*(\theta) = t/\theta, \qquad \forall \theta \in [\theta_n^*, 1] \; R^*(\theta) = t/\theta_n^*. \quad (38)$$

Let $V(u_0)$ denote investor $\mathcal{I}$'s net expected payoff from the project as a function of the reservation utility u_0 of entrepreneur $\mathcal{E}$. Then,

$$V(u_0) = H \int_{\theta_m^*}^1 \left(\theta - \theta_{fb}\right) dG(\theta) - \int_{\theta_m^*}^{\theta_n^*} \gamma dG(\theta) + \epsilon - (x + T(\theta_n^*)) - 1. \quad (39)$$

15

Not surprisingly, when the first-best is attainable, $V(u_0)$ coincides with V_{fb} as defined in (5). We call V the investor's value function. For the optimal contract to be valid, the project has to yield positive expected net return for the investor, that is, $V(u_0) \geq 0$.

By proposition 4, given the parameters of the model, H, γ, ϵ, and t, there are two possible cases: (1) $\hat{\theta}_{\mathsf{m}} < \hat{\theta}_{\mathsf{n}}$ at which for some values of the entrepreneur's reservation utility u_0, the optimal contract involves monitoring, and (2) $\hat{\theta}_{\mathsf{m}} \geq \hat{\theta}_{\mathsf{n}}$ at which there is never monitoring at optimum regardless of u_0. Define

$$\Lambda = \hat{\theta}_{\mathsf{n}} - \hat{\theta}_{\mathsf{m}}. \tag{40}$$

Consider case (1) first ($\Lambda > 0$). By proposition 4, the value function (39) can be divided into five segments, as depicted in Figure 2. When, $u_0 \geq \hat{u}_0$, where

$$\hat{u}_0 = T(\theta_{\mathsf{fb}}) \tag{41}$$

the optimal contract is the non-monitoring first-best $\sigma_{\mathsf{n}}^{\mathsf{fb}}$. The threshold between contract $\sigma_{\mathsf{m}}^{\mathsf{sb}}$ or $\sigma_{\mathsf{n}}^{\mathsf{sb}}$ being optimal, call it $\bar{u}_0$, is determined by $\bar{\theta}_{\mathsf{n}} = \hat{\theta}_{\mathsf{m}}$, that is,

$$\bar{u}_0 = T(\hat{\theta}_{\mathsf{m}}). \tag{42}$$

When the optimal contract is $\sigma_{\mathsf{m}}^{\mathsf{sb}}$, i.e. when $u_0 \in [0, \bar{u}_0)$, the cutoff level above which continued projects are not monitored, θ_{n}^*, can be either $\hat{\theta}_{\mathsf{n}}$ or $\bar{\theta}_{\mathsf{n}}$. The value function V is different depending on which value θ_{n}^* takes. In particular, when $\theta_{\mathsf{n}}^* = \bar{\theta}_{\mathsf{n}}$, entrepreneur $\mathcal{E}$ is paid exactly u_0, and V is strictly decreasing as u_0 rises. But when $\theta_{\mathsf{n}}^* = \hat{\theta}_{\mathsf{n}}$, entrepreneur $\mathcal{E}$ is paid $T(\hat{\theta}_{\mathsf{n}})$, which is independent of u_0, and as a result V is constant in u_0. Let u_0^1 be the boundary between the two regions, which is determined by $\hat{\theta}_{\mathsf{n}} = \bar{\theta}_{\mathsf{n}}$, that is,

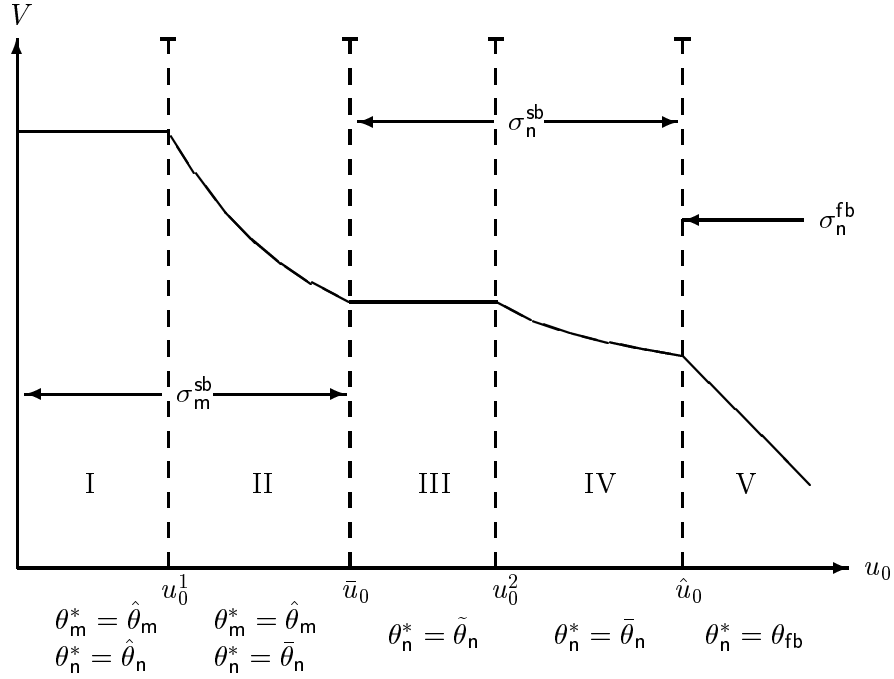$$u_0^1 = T(\hat{\theta}_{\mathsf{n}}). \tag{43}$$

Similarly, when the optimal contract is $\sigma_{\mathsf{n}}^{\mathsf{sb}}$, i.e. when $u_0 \in [\bar{u}_0, \hat{u}_0)$, the cutoff level below which projects are liquidated, θ_{n}^*, can be either $\tilde{\theta}_{\mathsf{n}}$ or $\bar{\theta}_{\mathsf{n}}$. When $\theta_{\mathsf{n}}^* = \bar{\theta}_{\mathsf{n}}$, more projects is liquidated as u_0 decreases, entrepreneur $\mathcal{E}$ is paid exactly u_0, and function V is strictly decreasing in u_0. But when $\theta_{\mathsf{n}}^* = \tilde{\theta}_{\mathsf{n}}$, the liquidation cutoff level is fixed at $\tilde{\theta}_{\mathsf{n}}$, entrepreneur $\mathcal{E}$ is paid $T(\tilde{\theta}_{\mathsf{n}})$, which is independent of u_0, and as a result V is constant in u_0. Let u_0^2 be the boundary between these two regions, which is determined by $\tilde{\theta}_{\mathsf{n}} = \bar{\theta}_{\mathsf{n}}$, that is,

$$u_0^2 = T(\tilde{\theta}_{\mathsf{n}}). \tag{44}$$

It is easy to verify that $0 \leq u_0^1 \leq \bar{u}_0 \leq u_0^2 \leq \hat{u}_0$. In case (2) ($\Lambda \leq 0$), the value function V corresponds to only a portion of the function under case (1): segments III, IV and V. To summaries, the value function V may consist of either five or three segments.

- If $\Lambda > 0$, $V(u_0)$ may be divided into five segments, as depicted in Figure 2:

  I: $\forall\, u_0 \in [0,\, u_0^1)$, $\sigma_{\mathsf{m}}^{\mathsf{sb}}$ is optimal, $(\theta_{\mathsf{m}}^*,\, \theta_{\mathsf{n}}^*) = (\hat{\theta}_{\mathsf{m}},\, \hat{\theta}_{\mathsf{n}})$, V is constant.

  II: $\forall\, u_0 \in [u_0^1,\, \bar{u}_0)$, $\sigma_{\mathsf{m}}^{\mathsf{sb}}$ is optimal, $(\theta_{\mathsf{m}}^*,\, \theta_{\mathsf{n}}^*) = (\hat{\theta}_{\mathsf{m}},\, \bar{\theta}_{\mathsf{n}})$, V is strictly decreasing.

  III: $\forall\, u_0 \in [\bar{u}_0,\, u_0^2)$, $\sigma_{\mathsf{n}}^{\mathsf{sb}}$ is optimal, $\theta_{\mathsf{m}}^* = \theta_{\mathsf{n}}^* = \tilde{\theta}_{\mathsf{n}}$, V is constant.

  IV: $\forall\, u_0 \in [u_0^2,\, \hat{u}_0)$, $\sigma_{\mathsf{n}}^{\mathsf{sb}}$ is optimal, $\theta_{\mathsf{m}}^* = \theta_{\mathsf{n}}^* = \bar{\theta}_{\mathsf{n}}$, V is strictly decreasing.

  V: $\forall\, u_0 \geq \hat{u}_0$, $\sigma_{\mathsf{n}}^{\mathsf{fb}}$ is optimal, $\theta_{\mathsf{m}}^* = \theta_{\mathsf{n}}^* = \theta_{\mathsf{fb}}$, V is linear and strictly decreasing.

- If $\Lambda \leq 0$, then $u_0^1 = \bar{u}_0 = 0$, and V is divided into III, IV and V three segments as above.

Figure 2. The Investor's Value Function

3.2.3. Analyzing the optimal contract

Suppose that in addition to assumption (1) and (2), $V(u_0) \geq 0$, that is, investment in the project is beneficial for both parties. Two questions are of particular interest. First, under what conditions does $T(\theta_{\mathsf{fb}}) \leq u_0$ hold and hence the first-best is attainable? Second, suppose that the first-best is not attainable, then under what conditions is monitoring optimal? Both questions can be studied with the criterion given in Proposition 4.

Consider first the condition $T(\theta_{\mathsf{fb}}) \leq u_0$, which is necessary and sufficient for attaining the first-best outcome. From the analysis above, $T(\theta_{\mathsf{fb}})$ is the minimum amount of expected compensation

17

needed, in absence of monitoring, to induce truth-telling and effort-making if all projects with potential success rate above θ_{fb} are to be continued. For this reason, we call $T(\theta_{\text{fb}})$ the incentive cost to first-best financing. Condition (31) requires that this incentive cost to be lower than entrepreneur $\mathcal{E}$'s reservation utility u_0. The following proposition summarizes the conditions under which the first-best is attainable. For convenience, define

$$\Psi(H,\ t,\ \epsilon,\ u_0) \equiv \frac{u_0}{t}\theta_{\text{fb}} - \int_{\theta_{\text{fb}}}^{1}(\theta - \theta_{\text{fb}})dG(\theta) \geq 0. \tag{45}$$

Then, condition $T(\theta_{\text{fb}}) \leq u_0$ is equivalent to $\Psi \geq 0$. Note that Ψ is independent of the monitoring cost γ.

Proposition 5. *Suppose that assumptions (1) and (2) hold, and $V(u_0) \geq 0$. Then holding other parameters constant,*

(i) *$\Psi \geq 0$ if and only if $u_0 \geq \hat{u}_0$.*

(ii) *There exists $\hat{H} \geq t + \epsilon$ such that $\Psi \geq 0$ if and only if $H \leq \hat{H}$;*

(iii) *If $\Psi \geq 0$ holds for some $\hat{\epsilon}$, then it holds for all $\epsilon \geq \hat{\epsilon}$.*

Part (i) of the proposition is obvious. Parts (ii) and (iii) are given by the facts that function Ψ is increasing in θ_{fb} and that θ_{fb} is decreasing in H but increasing in ϵ. The proof is omitted. When the first-best is not attainable, the optimal contract is either one at which some projects are monitored or one at which no project is monitored. The following proposition characterizes the parameter space for each of these two forms of second-best financing mechanism to be optimal.

Proposition 6. *Suppose that assumptions (1) and (2) hold, $V(u_0) \geq 0$, and that $\Psi < 0$. Let $\hat{H}$ and $\hat{u}_0$ be defined as in Proposition 5. Holding other parameters constant,*

(i) *For all $u_0 \in [0, \bar{u}_0)$, $\sigma_{\text{m}}^{\text{sb}}$ is optimal, and for all $u_0 \in [\bar{u}_0, \hat{u}_0)$, $\sigma_{\text{n}}^{\text{sb}}$ is optimal.*

(ii) *There exists $\bar{H} > \hat{H}$ suc that for all $H \in (\hat{H}, \bar{H})$, $\sigma_{\text{n}}^{\text{sb}}$ is optimal, and for all $H > \bar{H}$, $\sigma_{\text{m}}^{\text{sb}}$ is optimal.*

(iii) *There exists $\bar{\gamma} \in [0, H - t - \epsilon)$ such that for all $\gamma \leq \bar{\gamma}$, $\sigma_{\text{m}}^{\text{sb}}$ is optimal; and for all $\gamma \geq \bar{\gamma}$, $\sigma_{\text{n}}^{\text{sb}}$ is optimal.*

(iv) *Suppose that $\sigma_{\text{m}}^{\text{sb}}$ is optimal for some $\bar{\epsilon} > 0$, then $\sigma_{\text{m}}^{\text{sb}}$ is optimal for all $\epsilon \leq \bar{\epsilon}$.*

The proof for Proposition 6 is given in the appendix. In both lemmas, we omit the discussion of the effect of moral hazard cost t on the form of optimal contract because of its complexity.[5]

---

[5]We know only when the effort cost t is very large or very small, the first-best financing is achievable. This is because in both cases, the incentive cost $T(\theta_{\text{fb}})$ is small: when t is very small, the direct effect of a small effort cost implies that the incentive cost is small; when t is large, the indirect effect of inducing more liquidation (θ_{fb} is high) implies that the incentive cost required for the small amount of continued project is also low. However, in general, there is no monotone relationship between t and the incentive cost. As t increases, the compensation required to overcome the moral hazard problem for the continued project increases, but the optimal level of total amount of the project continued decreases. These two opposite effects are what create the potential non-monotonicity of the relationship.

We now explain the intuition behind Propositions 5 and 6.

There are four cost factors at work. First, monitoring cost γ is paid for project whose success rate is in the monitoring range M^*. Second, entrepreneur $\mathcal{E}$ has to be paid incentive cost $T(\theta_{\mathsf{n}})$, which is a result of giving the entrepreneur enough incentive to report truthfully and to make the required effort t for all continued but not-monitored projects as if they are the ones with the lowest signal θ in the region. Third, as Corollary 4.1 indicated, there will be over-liquidation when the first-best is not attainable. And fourth, entrepreneur $\mathcal{E}$'s expected net payoff from the project has to meet his reservation utility u_0. The form of optimal contract is usually a result of balancing two of the relevant factors out of these four. We discuss the influence of each of the parameters of the model on the choice of optimal financing through their effects on these factors, assuming assumptions (1) and (2) are satisfied, and $V(u_0) > 0$.[6]

$\mathcal{E}$'s reservation utility u_0:
$$\underbrace{\quad\overset{\sigma^{\mathsf{sb}}_{\mathsf{m}}}{\rule{2.5cm}{0pt}}}_{\bar{u}_0}\underbrace{\overset{\sigma^{\mathsf{sb}}_{\mathsf{n}}}{\rule{2.5cm}{0pt}}}_{\hat{u}_0}\overset{\sigma^{\mathsf{fb}}_{\mathsf{n}}}{\rule{2cm}{0pt}}$$

All four factors may be at play here, although not concurrently. Figure 3 depicts the relationship between the incentive cost function $T(\theta)$ and the reservation utility u_0, and their effects on the optimal contract, when monitoring contract is optimal for some u_0 (corresponds to Figure 2, the five segmented value function V). When the credit market awards a high expected utility $u_0 (> \hat{u}_0)$ to entrepreneur $\mathcal{E}$, the first-best outcome is achievable. In such a case, the incentive cost $T(\theta_{\mathsf{fb}})$ is part of the u_0 payment. No monitoring is necessary, since entrepreneur $\mathcal{E}$'s stake in the project is high enough to avoid any of his incentive problems.
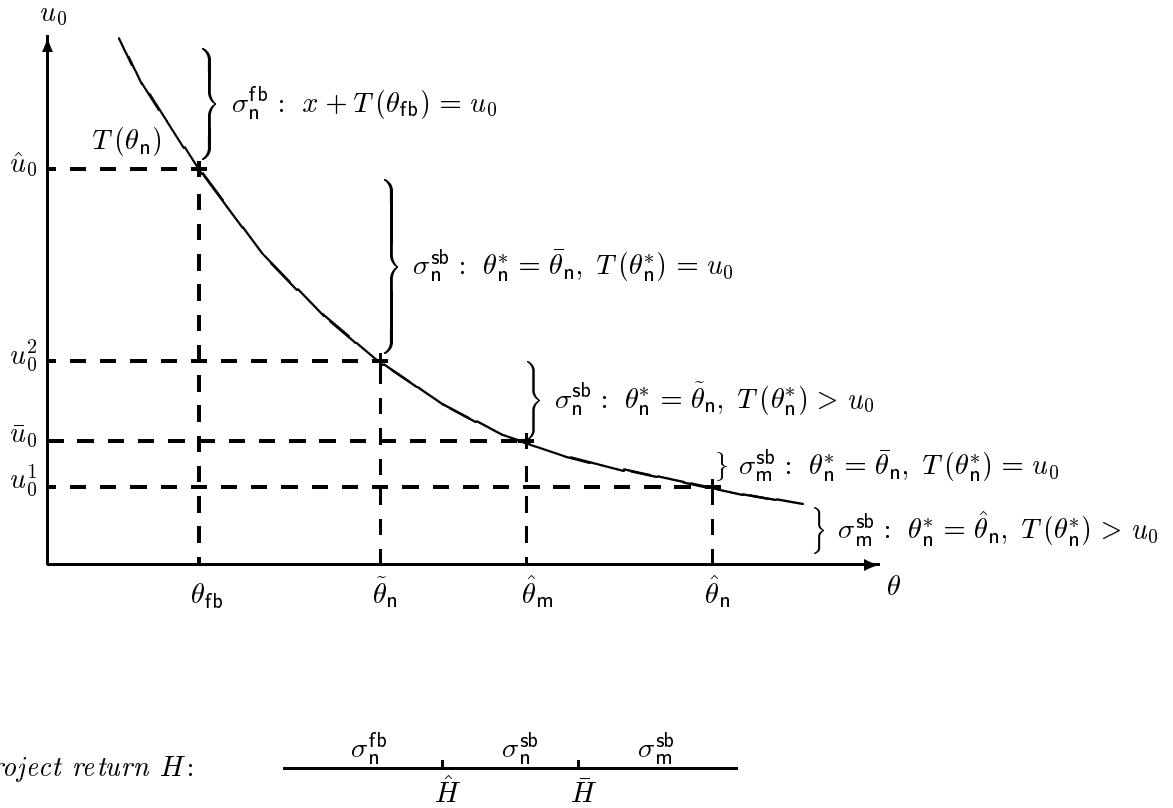
When u_0 is in the intermediate range ($u_0 \in (\bar{u}_0, \hat{u}_0)$, note that it is possible $\bar{u}_0 = 0$), it is not worthwhile for investor $\mathcal{I}$ to pay the incentive cost to first-best financing $T(\theta_{\mathsf{fb}})$, which is higher than u_0. It is also too costly to pay the monitoring cost and liquidating more projects than necessary. Continuity argument suggests that raising the liquidation threshold θ^*_{n}, and hence reducing the incentive cost $T(\theta^*_{\mathsf{n}})$ to u_0 preserve the best interests of both parties. This leads to the optimal contract to be the one with non-monitoring, $\sigma^{\mathsf{sb}}_{\mathsf{n}}$. However, when $u_0 \in (\bar{u}_0, u^2_0]$, setting the liquidation threshold θ^*_{n} by equating $T(\theta_{\mathsf{n}})$ to u_0 ($\theta^*_{\mathsf{n}} = \bar{\theta}_{\mathsf{n}}$) may resulting too much liquidation. In such a case, entrepreneur $\mathcal{E}$ is paid more than his market share of the project u_0 since the gain to continue the profitable project (or the cost of liquidating it) outweighs the incentive cost.

When u_0 is low ($u_0 < \bar{u}_0$, in the case of $\bar{u}_0 > 0$, which may be true when γ is low), it is optimal to monitoring some projects, and the chosen contract is $\sigma^{\mathsf{sb}}_{\mathsf{m}}$. While the trade-off between monitoring cost and liquidation cost determines the boundary between liquidation and continuation θ^*_{m}, the boundary between monitoring and non-monitoring among the continued projects is more complicated. Its determination partitioned this set of u_0 into two subsets. If $u_0 \in [\bar{u}_0, u^1_0)$, the non-monitoring region is chosen so that the incentive cost to induce effort $T(\theta^*_{\mathsf{n}})$

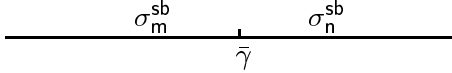[6]Note that the cut-off points, such as $\bar{u}_0$, $\hat{H}$, etc. may be zero.

is exactly u_0, and the rest of continuation region are monitored. If $u_0 \leq u_0^1$, the trade-off between paying the monitoring cost γ and the incentive cost due to non-monitoring tip the balance in favor of the incentive cost: it is beneficial for the investor to pay a higher incentive cost than u_0 in exchange for monitoring less projects. Therefore, even though the credit market may have assigned zero or a very small share of project surplus to the entrepreneur, his expect payoff will not go down that low since it is optimal for the investor to not monitor him and hence pay him the incentive cost than paying the high monitoring cost.

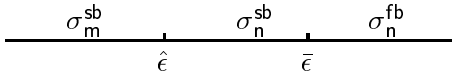Figure 3. Reservation Utility u_0, Incentive Cost $T(\theta)$, and Optimal Contract

With very low project return H ($< \hat{H}$), a project's potential success rate has to be very high in order for it to be continued, that is, the first-best liquidation level θ_{fb} is very high. With few high-success-rate projects being continued, the incentive cost of not monitoring these projects are also small. In particular, it is lower than the entrepreneur's reservation utility u_0. Therefore, the non-monitoring contract achieves the first-best financing. With very high return H ($> \bar{H}$), on the other hand, even projects with low potential success rate is worth to be continued. Because many projects are continued, paying the incentive cost of non-monitoring according to the lowest-success-rate continued project for the entire set of continued projects is very expensive. Hence, monitor the lower success rate projects and not monitoring the higher ones (contract $\sigma_{\mathsf{m}}^{\mathsf{sb}}$)

is optimal. There is a lower bound for H, $\bar{H}$, below which paying incentive cost of non-monitoring is cheaper than monitoring. In such a case, the non-monitoring contract σ_n^{sb} is optimal.

Monitoring cost γ:

$$\begin{array}{c|c} \sigma_m^{sb} & \sigma_n^{sb} \\ \hline & \bar{\gamma} \end{array}$$

This parameter is only relevant when the first-best financing is not attainable. The key trade-off here is paying the monitoring cost or the incentive cost due to non-monitoring. It is intuitive when the monitoring cost γ is low, the optimal contract should be the one with monitoring σ_m^{sb}. And when it is high, not monitoring but paying the incentive cost to induce truth-telling and effort-making is less costly, and hence, σ_n^{sb} is optimal.

Liquidation value ϵ:

$$\begin{array}{c|c|c} \sigma_m^{sb} & \sigma_n^{sb} & \sigma_n^{fb} \\ \hline \hat{\epsilon} & \bar{\epsilon} & \end{array}$$

The liquidation value ϵ works exactly the opposite way as project return H, through its effect on the level of project continuation. With very high ϵ ($> \bar{\epsilon}$), the opportunity cost of continue a project is high. Hence, few projects are continued. But with very low ϵ ($< \hat{\epsilon}$), the close to nothing scrape value of a project will not have much effect on its continuation decision, and hence relatively, more projects are continued. The effect of ϵ on the amount of projects continued is monotonic.

This concludes our discussion of the two-agent optimal contract.

4. Equilibrium

In this section we first describe what a credit market equilibrium is, and then compare allocations across equilibria under different parameter values of the model. We focus on the following questions: What determines the equilibrium total number of projects which are fully implemented? (What determines the economy's equilibrium output?) How are the economy's total investment and output related to the equilibrium lending mechanism?

We have assumed up to now that the parameters of the model, H, t, ϵ, γ, and the density function $g(\theta)$, satisfy assumptions (1) and (2). We need to make a third assumption: at least for some credit-market solution u_0, it is worthwhile for investors to invest in projects rather than the storage technology. That is,

Assumption (3) There exists $u_0 \geq 0$ such that $V(u_0) \geq 0$.

Since investors' value function $V(u_0)$ is a weakly decreasing function of u_0 by Lemma 4, assumption (3) implies the following lemma.

Lemma 6. *Suppose that assumptions (1)—(3) hold. Then for the optimal contract σ^*, there exists $\underline{u}_0 > 0$ satisfying $V(\underline{u}_0) = 0$, such that for all $u_0 \leq \underline{u}_0$, $V(u_0) \geq 0$, and for all $u_0 > \underline{u}_0$, $V(u_0) < 0$.*

We need to assume that the parameters of the model satisfy assumptions (1)—(3) in order to have any investment made.

The equilibrium notion we use is competitive: the short side of the market extract all the surplus from trades. There are two possibilities. The first is the case when the economy's total supply of loanable funds exceeds the total demand for funds, that is, $\delta < \lambda$. In this case, competition for projects among lenders will work to maximize the expected payoff of entrepreneurs. Specifically, it will drive the expected payoff of each entrepreneur u_0 down to $\underline{u}_0$ at which the expected net payoffs for investors are zero. Thus in equilibrium entrepreneurs will extract all of the surplus associated with the invested projects. Depending on where $\underline{u}_0$ is located, the optimal contract can be the monitoring $\sigma_{\mathsf{m}}^{\mathsf{sb}}$ ($\underline{u}_0 \in [0, \bar{u}_0]$), the non-monitoring second-best $\sigma_{\mathsf{n}}^{\mathsf{sb}}$ ($\underline{u}_0 \in [\bar{u}_0, \hat{u}_0]$), or the non-monitoring first-best $\sigma_{\mathsf{n}}^{\mathsf{fb}}$ ($\underline{u}_0 \geq \hat{u}_0$).

The second case occurs when the economy's total supply of funds is less than the total demand for funds, that is, $\delta > \lambda$. In this case, not all projects will be funded. Competition for funds will drive u_0 to zero, although the expected payoff of each entrepreneur may be above zero (due to incentive problems) at a level where the investor's expected returns are maximized. The optimal contract is the the monitoring $\sigma_{\mathsf{m}}^{\mathsf{sb}}$ if regions I and II in Figure 2 are not empty ($\bar{u}_0 > 0$). Otherwise, it is the non-monitoring second-best $\sigma_{\mathsf{n}}^{\mathsf{sb}}$.

In the special case where $\delta = \lambda$, the two parties can divide the surplus from invested projects in any arbitrary way. For simplicity, we assume that in such a case, entrepreneurs get all the surplus, since this is the most likely case to achieve the first-best. Formally,

Proposition 7. *Suppose that the parameters of the model satisfies assumptions (1)—(3). Then, there are two possible equilibria.*

(i) *If $\delta \leq \lambda$, then in equilibrium all projects are funded, and an equilibrium is a pair (u_0^*, σ^*), where $u_0^* = \underline{u}_0$ (that is, $V(u_0^*) = 0$), and σ^* is the optimal contract which gives expected utility u_0^* to the entrepreneur.*

(ii) *If $\delta > \lambda$, then in equilibrium the measure of the projects funded is λ, and an equilibrium is a pair (u_0^*, σ^*), where $u^* = 0$ and σ^* is the optimal contract which presumes zero expected payoff to the entrepreneur, but its actual prescription is $T(\theta_{\mathsf{n}}^*)$.*

Obviously we have imposed a very simple market structure in the definition of the credit-market equilibrium. This simple competition mechanism may be interpreted as a special case of a bargaining process which in principle can take a very general form.[7]

---

[7]For example, this bargaining process may dictate that the equilibrium fraction of the surplus associated with

4.1. Credit Rationing

Notice that whenever $\delta > \lambda$, in equilibrium there is *always* credit rationing of the type discussed by Stiglitz and Weiss (1981) and Williamson (1987), where among a group of identical borrowers some receive loans and some don't, and those who do are strictly better off than those don't.

Why is credit rationed? Sometimes it is because of costly monitoring, as in Williamson (1986, 1987).[8] Specifically, when $\sigma_{\mathsf{m}}^{\mathsf{sb}}$ is optimal, lowering the entrepreneur's reservation utility u_0 implies a higher $\bar{\theta}_n$, which in turn implies that the expected monitoring cost is higher. Sometimes credit is rationed because of costly liquidation. In particular, when $\sigma_{\mathsf{n}}^{\mathsf{sb}}$ is optimal, as u_0 decreases, more projects must be liquidated in order to make the contract incentive compatible. The notion that credit rationing is a mechanism to avoid excessive liquidation has not been discussed in the literature. Stiglitz and Weiss (1981) model credit rationing as a mechanism to reduce costly ex post default on loans.

4.2. Comparative Statics

In the section of two-agent optimal contract, we discuss the effect of the parameters of the model on the optimal contract, for a given credit market outcome u_0. In this section we study the general-equilibrium effects of shocks to project return H, monitoring cost γ, the investment fund supply (the measure of investors λ), and the demand for funds (the measure of entrepreneurs δ) on the equilibrium financing mechanism and output. The results obviously depend on the initial condition, in particular, which party gets bigger share of the investment return, that is, $u_0^* = 0$ or $u_0^* = \underline{u}_0$. We first summarize some properties the value function V which are needed to analyze the response of $\underline{u}_0$ to changes in the parameters of the model.

The value function V are implicit function of the exogenous parameters H, γ, ϵ,, and explicit function of u_0.

Lemma 4. *Holding other parameters of the model constant, the sign of the derivative of V with respect to parameter x, $\partial V/\partial x$, for x being H, γ, ϵ, as well as u_0 on the five (or three)*

---

a project that goes to the investor is a function of δ and λ in the form of say $\eta(\frac{\delta}{\lambda})$, with $\eta \in [0, 1]$ and $\eta' < 0$. The special case we use in this paper simply sets

$$\eta(\frac{\delta}{\lambda}) = \left\{ \begin{array}{ll} 1 & \text{if} \quad \frac{\delta}{\lambda} > 1 \\ 0 & \text{if} \quad \frac{\delta}{\lambda} < 1. \end{array} \right. \tag{46}$$

Adopting the more general form of the competition mechanism does not affect our results qualitatively.

[8]Williamson considers a standard costly state verification model.

segments of V are as follows.

	I	II	III	IV	V	*overall*
H	+	+	+	+	+	*increasing*
γ	–	–	0	0	0	*decreasing*
ϵ	+	+	0	+	+	*increasing*
u_0	0	–	0	–	–	*decreasing*

This lemma can be verified directly against the solution of the optimal contract given in Proposition 4. As parameters H, γ and ϵ changes, in addition to the level of V, the validity of condition $\Lambda > 0$ (the case where for some u_0, the optimal contract involves monitoring), as well as the boundaries $\hat{u}_0$, $\bar{u}_0$, u_0^1, u_0^2 that divide the value function V into five or three segments may change correspondingly. Using the definition of Λ, $\hat{u}_0$, $\bar{u}_0$, u_0^1, u_0^2 given in equations (40)—(44) as well as that of $\bar{\theta}_n$, $\hat{\theta}_n$ and $\tilde{\theta}_n$ given in equations (32), (36) and (37), the following lemma summarizes these effects.

Lemma 5. *Holding other parameters of the model constant, the sign of the derivative of y with respect to parameter x, $\partial y / \partial x$, for y being Λ, u_0^1, $\bar{u}_0$, u_0^2, $\hat{u}_0$, and x being H, γ, ϵ are as follows.*

	Λ	u_0^1	$\bar{u}_0$	u_0^2	$\hat{u}_0$
H	+	0	+	+	+
γ	–	+	–	0	0
ϵ	–	0	–	–	–

The derivations of both lemmas are omitted here.

4.2.1. *Disturbances to project return H*

Consider first the effect of a change in the level of H. Fix the model's other parameters at levels such that assumptions (1)—(3) are satisfied. Suppose the economy is experiencing a negative technology shock which lowers the level of H from H^o to H^n. Assume that with H^n, assumptions (1)—(3) remain valid.

Suppose first that regardless the initial environment, after the shock, $\Lambda^n \leq 0$ (at which only non-monitoring contract is optimal). Then, equilibrium lending mechanism either shifts from bank lending to market financing or continue to be market financing. With lower H, it is less profitable to continue some projects that would have been fully funded under H^o. Lower expected return of each continued project and increased liquidation implies that the total output of the economy will fall, more than the drop in H.

Next, suppose that in the initial environment, $\Lambda^o > 0$ (so it is possible to have intermediated lending to be optimal), and after the shock, by Lemma 5, Λ decreases but $\Lambda^n > 0$. Suppose that

the economy has an over-supply of projects, that is, $u_0^* = 0$. Then the financing mechanism was and continues to be intermediated lending ($\sigma_{\mathsf{m}}^{\mathsf{sb}}$). If the economy has an over-supply of funds, that is, $u^* = \underline{u}_0$, the effect of the shock is more complicated. On the one hand, by Lemma 4, the investors' value function V is an increasing function of H and a decreasing function of u_0. Hence, as H decreases from H^o to H^n, $\underline{u}_0$ is also reduced, say from $\underline{u}_0^o$ to $\underline{u}_0^n$. On the other hand, by Lemma 5, three of the four boundaries between segments of V, $\bar{u}_0$, u_0^2 and $\hat{u}_0$ also decrease. The end result is likely that the lending mechanism does not change in response to the shock. However, it is possible, although rare, that the initial funding mechanism is market lending ($\sigma_{\mathsf{n}}^{\mathsf{sb}}$ or $\sigma_{\mathsf{n}}^{\mathsf{fb}}$), and the new market clearing u_0^* ($= \underline{u}_0^n$) drops enough to a level that is less than the new boundary $\bar{u}_0^n$ (see Figure 2). In this case, the lending activities moves to borrowing from banks. Whether or not the lending mechanism is affected, the amount of projects liquidated increases, and hence total output falls.

Thus, our model predicts that, as H decreases, the economy will see more projects being liquidated. There is a flight for quality in the sense that projects which are fully executed have higher probabilities to succeed. Furthermore, if firms borrow from banks to finance investment, the drop in H is likely to trigger a switch to bond financing (although not necessarily). This is consistent with what happened during the 1990-91 recession. Note that as H falls, the economy's total output falls more than proportionally. This is because, as H decreases, not only each firm produces less, but also there are fewer firms producing. In other words, an earnings shock is amplified through the credit market.

Conversely, as H rises, which is often associated with economic boom, more projects will be fully implemented, total output of the economy rising, and the optimal lending mechanism is likely to be intermediated financing.

4.2.2. *Disturbances to monitoring cost γ*

Consider now the effect of a change in the values of γ, while other parameters of the model hold constant. Suppose the economy is experiencing an improvement in monitoring technology such that monitoring cost γ drops from γ^o to γ^n. Assume that both before and after the change, assumptions (1)—(3) are satisfied.

First, suppose that in the initial environment, $\Lambda^o \leq 0$, that is, the initial optimal contract is the non-monitoring $\sigma_{\mathsf{n}}^{\mathsf{sb}}$ or $\sigma_{\mathsf{n}}^{\mathsf{fb}}$. The drop of monitoring cost raises the value of Λ to $\Lambda^n > 0$ (by Lemma 5). If the economy has an over-supply of projects, that is, $u_0^* = 0$, then the equilibrium financing mechanism shifts to bank lending ($\sigma_{\mathsf{m}}^{\mathsf{sb}}$). If the economy has an over-supply of funds, that is, $u^* = \underline{u}_0$, bank-lending is optimal if $\underline{u}_0$ is now in the $\sigma_{\mathsf{m}}^{\mathsf{sb}}$ region, otherwise, bond financing remains optimal. Next, suppose that in the initial environment, $\Lambda^o > 0$. Then by Lemma 5, $\Lambda^n > 0$. In this case, regardless who has the upper hand on the credit market, the equilibrium financing form does not change. (When $u^* = \underline{u}_0$, a drop in γ induces a rise in $\underline{u}_0$ since V is

decreasing in γ in regions I and II, but it also leads to a right shift of the boundary between $\sigma_{\mathsf{m}}^{\mathsf{sb}}$ and $\sigma_{\mathsf{n}}^{\mathsf{sb}}$, $\bar{u}_0$. Given that the rest of function V does not change, the new $\underline{u}_0$ remains in region II.) In all these scenarios, lending mechanism either remains unchanged or moves into intermediated financing. Either way, the drop in γ will leads to less liquidated projects, hence, higher total output.

If there is an increase in the monitoring cost γ, the effect will be complete opposite, with increased liquidation, lower output, and more likely market lending.

4.2.3. *Shocks to demand for funds δ or supply of funds λ*

Suppose the economy resides initially at an equilibrium where there is an over-supply of funds, i.e., $\delta < \lambda$. So all projects are funded, and the entrepreneurs have the upper hand in the credit market, earning the maximum payoff possible $\underline{u}_0$. The equilibrium lending mechanism can be either borrowing from banks or issuing corporate bond.

Imagine now the economy receives a "real" shock which increases the number of investment opportunities from δ^o to δ^n, and $\delta^n > \lambda$. That is, there is an increase in the demand for funds while the supply of funds remain unchanged. This reversal of power on the credit market implies that now the entrepreneurs' share of the investment return is reduced to zero, i.e, $u_0 = 0$. The new optimal contract can be either $\sigma_{\mathsf{m}}^{\mathsf{sb}}$ or $\sigma_{\mathsf{n}}^{\mathsf{sb}}$. The only possible change of lending mechanism occurs when $\Lambda > 0$, the initial optimal contract prescribes non-monitoring ($\sigma_{\mathsf{n}}^{\mathsf{sb}}$ or $\sigma_{\mathsf{n}}^{\mathsf{fb}}$), and the new one requires monitoring ($\sigma_{\mathsf{m}}^{\mathsf{sb}}$). In such a case, lending activities shift from bond market to bank loans. Regardless of the lending mechanism, the induced drop in entrepreneurs' reservation utility is likely to trigger an increase in the amount of projects liquidated, unless $\underline{u}_0$ is in the flat portion of the value function V (region I or III). However, the increase of investment opportunities implies that the amount of projects funded rises from δ^o to λ, to fully utilize the available funds. The combined effect on total output is unclear, depending on whether the downward push of the increased liquidation can be overturned by the upward lift of the increased investment. It is possible that the effect of more liquidation dominates that of more funding, hence total output falls despite the increased investment. This seemingly counterfactual result is rooted in the logic that the tightening of credit market for the entrepreneurs may produce the adverse effect of intensifying incentive problems.

Next, consider the implications of a "monetary" shock that changes the supply of loanable funds, starting from the same initial credit market equilibrium (excess of funds, $u_0^* = \underline{u}_0$). Specifically, suppose that a sequence of monetary policy innovations manage to bring λ to below δ.[9] The reduction in available funds shifts the market power to the side of the investors, and leads to a similar reduction of the entrepreneurs' entitlement of investment return u_0^* from $\underline{u}_0$ to zero.

---

[9]We do not provide explanations for why a tight monetary policy induces a contraction in the total supply of credit.

The response to this decline in u_0^* is the same as those induced by the increased demand for funds discussed above, with increased liquidation and a potential shift of lending mechanism from issuing bond to borrowing from banks. The only difference here is that when available loans shrink, total investment also shrink. Therefore, the economy will see a definite fall in total output.

In our model, a fall in the supply of credit can create two effects: an interest rate effect and a credit effect. The interest rate effect occurs right when λ crosses δ, where there is a discrete downward jump in both total investment and total output caused by a sudden increase in the investor's expected return on a loan (rate of interest). As the expected return on loans increases, the utility of the entrepreneurs fall, agency costs increase, more projects are liquidated after the observation of the random signal, and aggregate output is lower.

After λ has crossed δ, the credit effect takes over. As λ keeps falling, total investment and total output fall continuously while the rate of interest remains flat. The same credit effect is discussed in Stiglitz and Weiss (1981). But a somewhat interesting point this paper offers is: a decrease in the total supply of loans may cause aggregate output to fall more than proportionally. A fall in loan supply causes less projects to be funded initially (a pure credit effect), and among those receive initial funding, more are to face liquidation subsequently (an agency effect).

5. Concluding Remarks

We have constructed and studied a model of the credit market in which both the economy's total output and the equilibrium source of financing are endogenously determined. In contrast to the literature, we focus on two important elements of external financing. One is that the equilibrium contract is optimal with respect to the environment rather than exogenously imposed. The other is the effect of credit market condition on the equilibrium lending mechanism. Among other things, we show that the observation that bank lending falls relative to corporate bond issuance during recessions can be explained by movements in the economy's real factors, including the availability of investment opportunities and the potential returns of an average investment project.

A major simplifying assumption of the model is that the economy's total demand for and supply of funds are exogenously fixed. This can be relaxed. For instance, one could imagine that the availability of funds is an increasing function of the expected return on a loan to the investor, or one could also assume that a higher expected return on a project to the entrepreneur brings a supply of more projects. But as long as these relationships are not sufficiently elastic, it is clear that the comparative statics properties of the model will remain valid.

Appendix

Proof of Proposition 1.

Step 1. We show that there exists a contract $\tilde{\sigma} = \{M; \Phi; \tilde{x}; y(\theta), \theta \in B, Y_D; R(\theta), R_0(\theta), \theta \in A; \tilde{R}(\theta), \tilde{R}_0(\theta), \theta \in C\}$ such that $\tilde{\sigma}$ is equivalent to σ, and that $\tilde{R}(\theta)$ and $\tilde{R}_0(\theta)$ are constants on C. Note that $\tilde{\sigma}$ is identical to σ except for $\tilde{R}(\theta)$, $\tilde{R}_0(\theta)$, and $\tilde{x}$. Without loss of generality, assume that C has a minimum point, and let $\theta_1 \equiv \min_{\theta \in C} \theta$. Let $\tilde{R}(\theta) = R(\theta_1)$, $\tilde{R}_0(\theta) = R_0(\theta_1)$ for all $\theta \in C$, and allow $\tilde{x}$ to be determined later.

(i) We first show $\tilde{\sigma}$ is incentive compatible. We need to show only that the revision on C satisfies conditions (6)—(8), and (10). Note that conditions (6) and (8) are obviously satisfied on C, given that $\theta_1 \in C$, and $\tilde{R}(\theta) = R(\theta_1)$, $\tilde{R}_0(\theta) = R_0(\theta_1)$, for all $\theta \in C$. Since condition (7) holds for $\theta = \theta_1$, we have for any $\theta \in C$, $\theta \geq \theta_1$,

$$\theta \tilde{R}(\theta) + \tilde{R}_0(\theta) \geq \theta_1 R(\theta_1) + R_0(\theta_1) \geq Y_D.$$

Thus $\tilde{\sigma}$ also satisfies (7) for all $\theta \in C$. Similarly, condition (10) holds for $\theta = \theta_1$, which implies $\theta_1 R(\theta_1) \geq t$. Then, for any $\theta \in C$, we have $\theta \tilde{R}(\theta) \geq \theta_1 R(\theta_1) \geq t$. That is, (10) is satisfied with any $\theta \in C$.

(ii) With $\tilde{\sigma}$ instead of σ, the entrepreneur's expected utility is different only in C. Let $\tilde{x}$ be defined as follows:

$$\tilde{x} = x + \int_C \left(\theta R(\theta) + R_0(\theta) \right) dG(\theta) - \int_C \left(\theta R(\theta_1) + R_0(\theta_1) \right) dG(\theta).$$

We need to show $\tilde{x} \geq 0$. But by (6), $\forall \theta \in C$,

$$\theta R(\theta) + R_0(\theta) \geq \theta R(\theta_1) + R_0(\theta_1) = \theta \tilde{R}(\theta) + \tilde{R}_0(\theta).$$

That is, with $\tilde{\sigma}$, for all $\theta \in C$, entrepreneur $\mathcal{E}$'s expected payoff is less than or equal to that of the original contract. We therefore have: $\tilde{x} \geq x \geq 0$.

By (ii), investor $\mathcal{I}$'s payoff is the same with contract $\tilde{\sigma}$ as with contract σ. So, we have shown that $\tilde{\sigma}$ is equivalent to the original contract σ.

Step 2. We further demonstrate that the contract $\tilde{\sigma}$ is equivalent to a third contract $\hat{\sigma} = \{M; \Phi; \hat{x}; y(\theta), \theta \in B, \hat{Y}_D; \hat{R}(\theta), \hat{R}_0(\theta), \theta \in A \cup C\}$, which is otherwise identical to $\tilde{\sigma}$ except

$$\forall \theta \in A \qquad \hat{R}_0(\theta) = 0, \qquad \hat{R}(\theta) = R(\theta) + \frac{1}{\theta} R_0(\theta) \tag{47}$$

$$\forall \theta \in C \qquad \hat{R}_0(\theta) = 0, \qquad \hat{R}(\theta) = R_C \equiv \tilde{R}(\theta) = R(\theta_1) \tag{48}$$

$$\hat{Y}_D = Y_D - R_0(\theta_1), \qquad \hat{x} = \tilde{x} + \mu(C) R_0(\theta_1) + \mu(D)(Y_D - R_0(\theta_1)). \tag{49}$$

(i) We show that this new contract $\hat{\sigma}$ promises the same expected utilities as does $\tilde{\sigma}$ to both entrepreneur $\mathcal{E}$ and investor $\mathcal{I}$. The entrepreneur's expected payoff on A under $\hat{\sigma}$ is the same pointwise as under σ since for each $\theta \in A$,

$$\theta R(\theta) + R_0(\theta) = \theta(R(\theta) + \frac{1}{\theta}R_0(\theta)) + 0 = \theta\hat{R}(\theta) + \hat{R}_0(\theta).$$

By (48), under $\hat{\sigma}$, if the project with θ in C succeeds, entrepreneur $\mathcal{E}$ receives the expected payoff R_C that he would receive under $\tilde{\sigma}$. His total expected payment on C when the project fails, $\mu(C)R_0(\theta_1)$, and part of the payment on D, $\mu(D)(Y_D - R_0(\theta_1))$ (which is positive by (8)), are moved from C and D, respectively, into the constant payment $\hat{x}$ (an increase from $\tilde{x}$). Therefore, the two contracts give the same expected payoffs to both agents.

(ii) We show that the new contract $\hat{\sigma}$ is incentive compatible. First, since the changes on A do not affect the entrepreneur's expected payoff pointwise, the left-hand side of the relevant constraints (6) and (7) are the same as those under $\tilde{\sigma}$. Seconde, note that for any $\theta \in \Theta$, and any $\theta' \in C$, $\tilde{R}_0(\theta') = R_C \geq 0 = \hat{R}_0(\theta')$, and $\theta\tilde{R}(\theta') + \tilde{R}_0(\theta') - t \geq \theta R_C - t = \theta\hat{R}(\theta') + \hat{R}_0(\theta') - t$. Furthermore, $Y_D \geq \hat{Y}_D$. That is, the right-hand sides of conditions (6)—(9) under $\hat{\sigma}$ are all smaller than that under $\tilde{\sigma}$. Therefore, for any $\theta \in A \cup B$, conditions (6)—(9) are satisfied under $\hat{\sigma}$. Next, given that the definition of $\hat{Y}_D$ by (49), and that conditions (6)—(8) are satisfied for any $\theta \in C \cup D$ under $\tilde{\sigma}$, they are also satisfied under contract $\hat{\sigma}$. Last, since constraint (10) is satisfied under $\tilde{\sigma}$, for any $\theta \in A$, $\theta R(\theta) \geq t$, and $\theta_1 R_C \geq t$. By (47), $\hat{R}(\theta) \geq R(\theta)$, so for any $\theta \in A$, $\theta\hat{R}(\theta) \geq \theta R(\theta) \geq t$. For any $\theta \in C$, $\theta \geq \theta_1$, so $\theta\hat{R}(\theta) = \theta R_C \geq \theta_1 R_C \geq t$. Therefore, condition (10) holds under $\hat{\sigma}$.

We have shown that incentive constraints (6)—(10) hold for $\hat{\sigma}$, and that both agents receive the same expected payoff under contract $\hat{\sigma}$ as under $\tilde{\sigma}$. Therefore, the two contracts are equivalent.
∎

Proof of Proposition 2.

We first introduce some notation. Define

$$X >^* Y \quad \Longleftrightarrow \quad \forall\, x \in X, \forall\, y \in Y,\ x > y.$$

Let $\mathcal{P} = \mu \times \mu$ be the product measure on $\Theta \times \Theta$, and let $X \tilde{>} Y$ denote "$X >^* Y$ almost surely",

$$X \tilde{>} Y \quad \Longleftrightarrow \quad \mathcal{P}\Big\{(x,y) \,|\, x \in X,\, y \in Y \text{ and } y > x\Big\} = 0.$$

The proposition states that for any given optimal contract σ, there is contract which is equivalent to σ and which satisfies $\hat{\Phi} >^* \hat{\Phi}'$.

We first show that σ satisfies $\Phi \tilde{>} \Phi'$, which is equivalent to showing $A \tilde{>} B$, $A \tilde{>} D$, $C \tilde{>} B$, and $C \tilde{>} D$. Before proceeding, assume each of the four sets A, B, C and D has positive measure. (If one of the sets has measure 0, the corresponding assertion holds automatically.)

29

(i) We show $A \tilde{>} B$. Suppose not, then there exist $\Theta_A \subseteq A$ and $\Theta_B \subseteq B$ such that $\Theta_B >^* \Theta_A$. Without loss of generality, suppose that Θ_A and Θ_B satisfy $\mu(\Theta_A) = \mu(\Theta_B) \neq 0,$[10] and that $R(\theta)$ has a minimum on Θ_A. Now consider an alternative contract $\tilde{\sigma}$ which is identical to σ except

(a) $\tilde{B} = B \cup \Theta_A \setminus \Theta_B$, and $\forall \theta_A \in \Theta_A$,

$$\tilde{y}(\theta_A) = \frac{1}{\mu(\Theta_B)} \int_{\Theta_B} y(\theta) dG(\theta).$$

(b) $\tilde{A} = A \cup \Theta_B \setminus \Theta_A$, and $\forall \theta_B \in \Theta_B$,

$$\tilde{R}(\theta_B) = R_{min}(\Theta_A) \equiv \min_{\theta \in \Theta_A} R(\theta),$$

(c) If $\int_{\Theta_B} \theta_B R_{min}(\Theta_A) dG(\theta_B) < \int_{\Theta_A} \theta_A R(\theta_A) dG(\theta_A)$, then

$$\tilde{x} = x + \int_{\Theta_A} \theta_A R(\theta_A) dG(\theta_A) - \int_{\Theta_B} \theta_B R_{min}(\Theta_A) dG(\theta_B)$$

otherwise, $\tilde{x} = x$.

We need only verify that the incentive constraints (17)—(20) and (23) hold for $\tilde{\sigma}$. Since $R(\theta_A) \geq R_C$ for all $\theta_A \in \Theta_A$, it holds that $\tilde{R}(\theta_B) = R_{min}(\Theta_A) \geq R_C$, for all $\theta_B \in \Theta_B$, or $\tilde{R}(\theta_B)$ satisfies (17). Now define $\breve{\theta}_A \equiv \mathrm{argmin}_{\theta \in \Theta_A} R(\theta_A)$. By (18) and (23), $\breve{\theta}_A R(\breve{\theta}_A) \geq Y_D t \geq t$. Thus for all $\theta_B \in \Theta_B$,
$$\theta_B \tilde{R}(\theta_B) \geq \breve{\theta}_A \tilde{R}(\theta_B) = \breve{\theta}_A R(\breve{\theta}_A) \geq Y_D t \geq t.$$

That is, $\tilde{R}(\theta_B)$ satisfies (18) and (23). Since σ satisfies constraints (19) and (20), we have for all $\theta \in \Theta_B$, $y(\theta) \geq \max\{\theta R_C - t, Y_D\}$. Therefore, for all $\theta_A \in \Theta_A$,

$$\tilde{y}(\theta_A) = \frac{1}{\mu(\Theta_B)} \int_{\Theta_B} y(\theta) dG(\theta) \geq \max\left\{ \frac{R_C}{\mu(\Theta_B)} \int_{\Theta_B} \theta dG(\theta) - t, \ Y_D \right\} \geq \max\left\{ R_C \theta_A - t, \ Y_D \right\},$$

or $\tilde{y}(\theta_A)$ satisfies (19) and (20). Thus we have shown that $\tilde{\sigma}$ is incentive compatible.

Next, we show that $\tilde{\sigma}$ Pareto dominates σ. By construction, $\int_{\Theta_A} \tilde{y}(\theta_A) dG(\theta_A) = \int_{\Theta_B} y(\theta_B) dG(\theta_B)$. Suppose $\int_{\Theta_B} \theta_B R_{min}(\Theta_A) dG(\theta_B) \geq \int_{\Theta_A} \theta_A R(\theta_A) dG(\theta_A)$. Then moving from σ to $\tilde{\sigma}$ the entrepreneur's expected payoff is changed by

$$\int_{\Theta_B} \theta_B R_{min}(\Theta_A) dG(\theta_B) - \int_{\Theta_A} \theta_A R(\theta_A) dG(\theta_A) > 0,$$

and the investor's expected payoff is changed by

$$\int_{\Theta_B} \theta_B (H - R_{min}(\Theta_A)) dG(\theta_B) - \int_{\Theta_A} \theta_A (H - R(\theta_A)) dG(\theta_A) > 0,$$

---

[10]Given θ is continuously distributed, the sets Θ_A and Θ_B can be cut arbitrarily small to satisfy this property.

since $\Theta_B >^* \Theta_A$. Thus both parties are better off under $\tilde{\sigma}$ than under σ.

Suppose $\int_{\Theta_B} \theta_B R_{min}(\Theta_A) dG(\theta_B) < \int_{\Theta_A} \theta_A R(\theta_A) dG(\theta_A)$. Then under $\tilde{\sigma}$ the entrepreneur's expected payoff decreases on Θ_B compared to what she receives under σ on Θ_A, but the decrease is made up exactly by the increase of x to $\tilde{x}$, so her total expected payoff remains the same. Now the investor's expected payment to the entrepreneur is the same, but the investor's expected payoff is increased by

$$H\left(\int_{\Theta_B} \theta_B dG(\theta_B) - \int_{\Theta_A} \theta_A dG(\theta_A)\right) > 0.$$

This is because projects with higher success rates are continued. Again, $\tilde{\sigma}$ Pareto dominates σ.

(ii) We show $A \tilde{>} D$. Suppose not, then there exist $\Theta_A \subseteq A$ and $\Theta_D \subseteq D$ such that $\Theta_D >^* \Theta_A$. Without loss of generality, suppose Θ_A and Θ_D satisfy $\mu(\Theta_A) = \mu(\Theta_D) \neq 0$, and $R(\theta)$ has a minimum on Θ_A. Now consider an alternative contract $\tilde{\sigma}$ which is identical to σ except

(a) $\tilde{D} = D \cup \Theta_A \setminus \Theta_D$, and $\forall \theta_A \in \Theta_A$,

$$\tilde{y}(\theta_A) = Y_D.$$

(b) $\tilde{A} = A \cup \Theta_D \setminus \Theta_A$, and $\forall \theta_D \in \Theta_D$,

$$\tilde{R}(\theta_D) = R_{min}(\Theta_A) \equiv \min_{\theta \in \Theta_A} R(\theta), \ \forall \theta_D \in \Theta_D.$$

(c) If $\int_{\Theta_D} \theta_D R_{min}(\Theta_D) dG(\theta_D) < \int_{\Theta_A} \theta_A R(\theta_A) dG(\theta_A)$,

$$\tilde{x} = x + \int_{\Theta_A} \theta_A R(\theta_A) dG(\theta_A) - \int_{\Theta_D} \theta_D R_{min}(\Theta_A) dG(\theta_D),$$

otherwise, $\tilde{x} = x$.

Now since $\Theta_D >^* \Theta_A$ and (22) holds for σ, we have $Y_D \geq \theta_D R_C - t \geq \theta_A R_C - t$ holds for all θ_A and θ_D, and hence constraint (22) is satisfied by contract $\tilde{\sigma}$.

As in the proof for $A \tilde{>} B$, we can show that $\tilde{R}(\theta_D)$ satisfies constraints (17), (18), and (23), and thus $\tilde{\sigma}$ is incentive compatible. As in the proof for $A \tilde{>} B$, we can show that $\tilde{\sigma}$ Pareto dominates σ, a contradiction.

(iii) We show $C \tilde{>} B$. Suppose not. Without loss of generality, assume that there exists $\Theta_B \subseteq B$ and $\Theta_C \subseteq C$ such that $\Theta_B >^* \Theta_C$, and $\mu(\Theta_B) = \mu(\Theta_C) \neq 0$. Now consider an alternative contract $\tilde{\sigma}$ which is identical to σ except

(a) $\tilde{B} = B \cup \Theta_B \setminus \Theta_C$, and $\forall \theta_C \in \Theta_C$,

$$\tilde{y}(\theta_C) = \frac{1}{\mu(\Theta_B)} \int_{\Theta_B} y(\theta) dG(\theta).$$

(b) $\tilde{C} = C \cup \Theta_C \setminus \Theta_B$, and $\forall \theta_B \in \Theta_B$,

$$\tilde{R}(\theta_B) = R_C.$$

Since every $\theta_B \in \Theta_B$ satisfies (19) and (20), $y(\theta_B) \geq \max\{\theta_B R_C - t, Y_D\}$. Then, for any $\theta_C \in \Theta_C$,

$$\tilde{y}(\theta_C) = \frac{1}{\mu(\Theta_B)} \int_{\Theta_B} y(\theta) dG(\theta) \geq \max\left\{\frac{R_C}{\mu(\Theta_B)} \int_{\Theta_B} \theta dG(\theta) - t, \ Y_D\right\} \geq \max\left\{R_C \theta_C - t, \ Y_D\right\}.$$

That is, $\tilde{y}(\theta_C)$ satisfies (19) and (20). Also, take an arbitrary $\theta_C \in \Theta_C$, $\theta_C R_C - t \geq Y_D$ and $\theta_C R_C \geq t$. Since for any $\theta_B \in \Theta_B$, $\theta_B > \theta_C$, we have $\theta_B R_C - t \geq Y_D$ and $\theta_B R_C \geq t$, or, constraints (21) and (24) are satisfied on Θ_B. So the modified contract satisfies all the relevant incentive constraints.

By construction, $\int_{\Theta_C} \tilde{y}(\theta_C) dG(\theta_C) = \int_{\Theta_B} y(\theta_B) dG(\theta_B)$. But since $\Theta_B >^* \Theta_C$, the entrepreneur's expect payoff is increased by

$$\int_{\Theta_B} \theta_B R_C dG(\theta_B) - \int_{\Theta_C} \theta_C R_C dG(\theta_C) > 0,$$

and the investor's expected payoff is increased by

$$\int_{\Theta_B} \theta_B (H - R_C) dG(\theta_B) - \int_{\Theta_C} \theta_C (H - R_C) dG(\theta_C) > 0.$$

That is, both agents' expected payoffs are strictly higher under $\tilde{\sigma}$ than under σ.

(iv) Last, we show $C \tilde{>} D$. Constraints (21) and (22) directly imply that $C >^* D$, which further implies $C \tilde{>} D$.

To summarize, we have shown that $\Phi \tilde{>} \Phi'$. Given that contract σ satisfies $\Phi \tilde{>} \Phi'$, it is trivial to show that there is an equivalent contract $\hat{\sigma}$ that satisfies $\hat{\Phi} >^* \hat{\Phi}'$. Since $\Phi >^* \Phi'$ can only be violated on a measure zero set, we can rearrange monitoring and continuation/liquidation policies on this measure-zero set to eliminate the violations without affecting the payoffs. ∎

Proof of Lemma 3.

Consider an optimal contract σ. By Proposition 2, we assume $\Phi = [\theta_1, 1]$ and $\Phi' = [0, \theta_1)$. Suppose $\Phi' = \emptyset$. Then consider contract $\hat{\sigma}$ which is otherwise identical to σ except

(a) $\hat{D} = [0, t/H)$, $\hat{A} = A \cap [t/H, 1]$, $\hat{C} = C \cap [t/H, 1]$.

(b) $\hat{Y}_D = 0$.

(c) If $R_C > H$, then $\hat{R}_C = H$; if $R_C \leq H$ (including the case $C = \emptyset$), then $\hat{R}_C = R_C$.

(d) $\hat{x} = x + \int_{A \cap \hat{D}} (\theta R(\theta) - t) dG(\theta) + \int_{C \cap \hat{D}} (\theta R_C - t) dG(\theta) + \int_{\hat{C}} \theta(R_C - \hat{R}_C) dG(\theta).$

Notice that since $t > 0$, $\mu(\hat{D}) \neq 0$.

By construction, for all $\theta \in \hat{D}$, $\theta\hat{R}_C - t \leq \frac{t}{H}\hat{R}_C - t \leq 0 = \hat{Y}_D$, and hence constraint (22) is satisfied on $\hat{D}$. The contract $\hat{\sigma}$ satisfies constraint (17) since $R_C \geq \hat{R}_C$. $\hat{\sigma}$ also satisfies (18) since (23) holds under σ. If $\hat{R}_C = R_C$, then clearly constraints (21) and (24) are both satisfied. If $H = \hat{R}_C < R_C$, then for all $\theta \in \hat{C}$, $\theta \geq t/H$, $\theta\hat{R}_C = \theta H \geq t$, hence constraints (21) and (24) are also satisfied. Therefore, $\hat{\sigma}$ is incentive compatible. Finally, the expected payment to the entrepreneur under contract $\hat{\sigma}$ is the same as under contract σ, but the investor's expected payoff is increased by $\epsilon\mu(\hat{D}) - \int_{\hat{D}}(\theta H - t)dG(\theta) > 0$. This contradicts the fact that the σ is optimal. ∎

Proof of Proposition 3.

We first show that given the optimal continuation/liquidation policy Φ, the optimal monitoring region A is a lower interval of Φ and the non-monitoring region C is the complement upper interval of Φ. Suppose this is not true. That is, suppose there is an optimal contract σ such that a subset Θ_A of A is embedded in C, that is, for all $\theta_A \in \Theta_A$, $\theta_A > \inf_C \theta$. Without loss of generality assume $\mu(\Theta_A) \neq 0$.

Consider contract $\hat{\sigma}$ which is otherwise identical to σ except

(a) $\hat{A} = A \setminus \Theta_A$, $\hat{C} = C \cup \Theta_A$, and $\forall \theta_A \in \Theta_A, \hat{R}(\theta_A) = R_C$.

(b) $\hat{x} = x + \int_{\Theta_A} \theta(R(\theta) - R_C)dG(\theta)$.

To show that $\hat{\sigma}$ is incentive compatible, we need only check that constraints (21) and (24) are satisfied for all $\theta_A \in \Theta_A$. Since for all $\theta \in C$, $\theta R_C - t \geq Y_D$, we have $(\inf_C \theta)R_C - t \geq Y_D$, which in turn implies for all $\theta_A \in \Theta_A$, $\theta_A R_C - t \geq Y_D$ given that $\theta_A > \inf_C \theta$. That is, (21) is satisfied. Constraint (24) is implied by (21) since $Y_D \geq 0$.

By construction of $\hat{x}$, the entrepreneur's expected payoff remains the same under $\hat{\sigma}$. But the investor gains by the savings of the monitoring cost $\gamma\mu(\Theta_A) > 0$. This contradicts the fact that the contract σ is optimal.

Next, we show that $B = \emptyset$. Let σ be an optimal contract which has $B \neq \emptyset$. By Proposition 2, $\Phi >^* \Phi'$, and from the above proof, $C >^* A$. Hence we can let $A = [\theta_m, \theta_n)$ and $C = [\theta_n, 1]$, where $\theta_m \leq \theta_n$. Consequently, $\Phi = A \cup C = [\theta_m, 1]$ and $\Phi' = B \cup D = [0, \theta_m)$. Notice Lemma 3 implies $\theta_m > 0$.

Consider an alternative contract $\hat{\sigma}$ which is otherwise identical to σ except

(a) $\hat{R}_C = t/\theta_n$.

(b) $\hat{D} = D \cup B$, and $\hat{Y}_D = 0$.

(c) $\hat{x} = x + \int_C \theta(R_C - \hat{R}_C)dG(\theta) + \int_B y(\theta)dG(\theta) + \int_D Y_D dG(\theta)$.

Using the equations $R_C \geq t/\theta_n = \hat{R}_C$ and $\hat{Y}_D = 0$, it is easy to check that the contract $\hat{\sigma}$ satisfies all the incentive constraints including (17), (18), (21), (22), (24), as well as the non-negative constraints (13). Moreover, the construction of $\hat{x}$ implies that the entrepreneur's expected compensation under $\hat{\sigma}$ is the same as under σ. However, under $\hat{\sigma}$ the investor's expected payoff is increased by the savings of the monitoring cost $\gamma\mu(A \cup B) > 0$. This contradicts the assumption that σ is optimal.

Finally, we show (iii) holds. Suppose σ is optimal and it has a compensation scheme that differs from what is given by the proposition. We need only show that σ is equivalent to a contract $\hat{\sigma}$ whose compensation scheme takes the form that is given by the proposition. Let the compensation scheme of $\hat{\sigma}$ be given by $\hat{R}(\theta) = t/\theta$ for all $\theta \in A$, $\hat{R}_C = t/\theta_n$, $\hat{Y}_D = 0$, and $\hat{x} = x + \int_A \theta(R(\theta) - t/\theta)dG(\theta) + \int_C \theta(R_C - \hat{R}_C)dG(\theta) + \mu(D)Y_D$. It is easy to check that $\hat{\sigma}$ satisfies incentive constraints (17)–(24). Since the compensation schedule of the contract σ also satisfies these constraints, in particular, for all $\theta \in A$, $R(\theta) \geq t/\theta = \hat{R}(\theta)$, for all $\theta \in C$, $R_C \geq t/\theta_n = \hat{R}_C$, and $Y_D \geq 0 = \hat{Y}_D$, we have $\hat{x} \geq x$. Clearly, the compensation scheme of $\hat{\sigma}$ conforms with the proposition, and $\hat{\sigma}$ is equivalent to σ. $\blacksquare$

Solution to Problem $(P4)$

We first show that by assumptions (1) and (2), the objective function $O(\theta_m, \theta_n)$ is strictly concave in both θ_m and θ_n.

The function $O(\theta_m, \theta_n)$ is strictly concave if its Hessian matrix is negative definite. By the definition of function $O(\theta_m, \theta_n)$ in equation (33), $\partial O(\theta_m, \theta_n)^2/\partial\theta_n\partial\theta_m = 0$. So, we only need to show that the second derivatives with respect to θ_m and θ_n are strictly negative.

$$\frac{\partial O(\theta_m, \theta_n)^2}{\partial\theta_m^2} = -[\theta_m H - (t + \epsilon + \gamma)]g'(\theta_m) - Hg(\theta_m).$$

If $\theta_m H - (t + \epsilon + \gamma) \geq 0$, then $\partial O(\theta_m, \theta_n)^2/\partial\theta_m^2 < 0$ since by the first inequality of assumption (2),

$$\frac{-H}{H - (t + \epsilon + \gamma)} < \frac{-t}{H - (t + \epsilon)} \leq \frac{g'(\theta_m)}{g(\theta_m)}.$$

If $\theta_m H - (t + \epsilon + \gamma) < 0$, then $\partial O(\theta_m, \theta_n)^2/\partial\theta_m^2 < 0$ is equivalent to

$$\frac{g'(\theta_m)}{g(\theta_m)} < \frac{H}{(t + \epsilon + \gamma) - \theta_m H}$$

which is implied by the second inequality of assumption (2). With respect to θ_n,

$$\frac{\partial O(\theta_m, \theta_n)^2}{\partial\theta_n^2} = -\frac{t}{\theta_n}g(\theta_n) - \frac{2t}{\theta_n^3}\int_{\theta_n}^1 \theta dG(\theta) - \gamma g'(\theta_n) < -\frac{t}{\theta_n}g(\theta_n) - \gamma g'(\theta_n) \equiv J(\theta_n).$$

34

By assumptions (1) and (2),

$$-\frac{t}{\gamma\theta_n} < -\frac{t}{\gamma} \le \frac{-t}{H-(I_1+t+\epsilon)} \le \frac{g'(\theta_m)}{g(\theta_m)}$$

hence, $J(\theta_n) < 0$, or equivalently, $\partial O(\theta_m,\theta_n)^2/\partial\theta_n^2 < 0$. So, $O(\theta_m,\theta_n)$ is strictly concave in both θ_m and θ_n.

Given that the objective function $O(\theta_m,\theta_n)$ is strictly concave, and that the constraint set defined by (34) is convex, problem $(P4)$ has a unique solution. We can solve $(P4)$ with *Lagrange*'s method. Let ρ_1 be the multiplier for constraint $\theta_m \le \theta_n$, and ρ_2 be the multiplier for constraint $\theta_n \le \bar{\theta}_n$. Then the Lagrange is given by

$$L(\theta_m,\theta_n,\rho_1,\rho_2) = \int_{\theta_m}^{1}\Big(\theta H - t\Big)dG(\theta) + \epsilon G(\theta_m) - T(\theta_n) - \int_{\theta_m}^{\theta_n}\gamma dG(\theta) + \rho_1(\theta_n-\theta_m) + \rho_2(\bar{\theta}_n-\theta_n).$$

The first-order conditions are

$$\frac{\partial L}{\partial\theta_m} = \Big(-H\theta_m + t + \epsilon + \gamma\Big)g(\theta_m) - \rho_1 = 0 \tag{50}$$

$$\frac{\partial L}{\partial\theta_n} = \frac{t}{(\theta_n)^2}\int_{\theta_n}^{1}\theta dG(\theta) - \gamma g(\theta_n) + \rho_1 - \rho_2 = 0 \tag{51}$$

$$\frac{\partial L}{\partial\rho_1} = \theta_n - \theta_m \ge 0, \quad \rho_1 \ge 0 \quad \text{with complementary slackness} \tag{52}$$

$$\frac{\partial L}{\partial\rho_2} = \bar{\theta}_n - \theta_n \ge 0, \quad \rho_2 \ge 0 \quad \text{with complementary slackness.} \tag{53}$$

Depending on which of the constraints binds, there are four possible solutions for $\{\theta_m^*,\theta_n^*\}$.

(a) $\rho_1 = 0$, $\rho_2 = 0$. Then neither constraint binds. By (50) and (51), $\theta_m^* = \hat{\theta}_m$ and $\theta_n^* = \hat{\theta}_n$.

(b) $\rho_1 = 0$, $\rho_2 > 0$. Then by (50), $\theta_m^* = \hat{\theta}_m$, and θ_n^* is given by the binding constraint: $\theta_n^* = \bar{\theta}_n$.

(c) $\rho_1 > 0$, $\rho_2 = 0$. Then $\theta_m = \theta_n$ and $\theta_n < \bar{\theta}_n$. Substituting $\rho_2 = 0$ and θ_m by θ_n, we get equation (37) from (50) and (51). Given that $\tilde{\theta}_n$ is its solution, $\theta_n^* = \theta_m^* = \tilde{\theta}_n$.

(d) $\rho_1 > 0$, $\rho_2 > 0$. Then both constraints binds: $\theta_m^* = \theta_n^* = \bar{\theta}_n$.

To summarize, the solution to $(P4)$ can be one of the two classes, depending on whether $\theta_m^* = \theta_n^*$:

(1) When $\hat{\theta}_m < \min\{\hat{\theta}_n, \bar{\theta}_n\}$, which includes cases (a) and (b), $\theta_m^* < \theta_n^*$. Then the monitoring region M^* is not empty, $M^* = [\theta_m^*, \theta_n^*)$, and the project-continuation region is $\Phi^* = [\theta_m^*, 1]$, where $\theta_m^* = \hat{\theta}_m$, $\theta_n^* = \min\{\hat{\theta}_n, \bar{\theta}_n\}$, and $\theta_m^* > \theta_{fb}$.

(2) When $\hat{\theta}_m \ge \min\{\hat{\theta}_n, \bar{\theta}_n\}$, which includes cases (c) and (d), $\theta_m^* = \theta_n^*$. Hence, the monitoring region is empty, $M^* = \emptyset$, and the project-continuation region is given by $\Phi^* = [\theta_n^*, 1]$, where $\theta_n^* = \min\{\tilde{\theta}_n, \bar{\theta}_n\}$, and $\theta_n^* > \theta_{fb}$.

Proof of Proposition 6.

(i) This is given by the discussion of value function V.

(ii) Both $\hat{\theta}_n$ and $\bar{\theta}_n$ do not depend on H. When $H = \hat{H}$, $\bar{\theta}_n = \theta_{fb} \leq (t + \epsilon + \gamma)/H = \hat{\theta}_m$, since $u_0 = T(\theta_{fb}) = T(\bar{\theta}_n)$. Then, regardless of $\hat{\theta}_n$, $\hat{\theta}_m > \bar{\theta}_n \geq \min\{\hat{\theta}_n, \bar{\theta}_n\}$. When $H \to \infty$, $\hat{\theta}_m \to 0 < \min\{\hat{\theta}_n, \bar{\theta}_n\}$. Since $\hat{\theta}_m$ is a continuous function of H, and $\min\{\hat{\theta}_n, \bar{\theta}_n\}$ does not depend on H, there exists an $\bar{H} > \hat{H}$, such that for all $H \in (\hat{H}, \bar{H}]$, $\hat{\theta}_m \geq \min\{\hat{\theta}_n, \bar{\theta}_n\}$, which by Proposition 4, implies that the optimal contract is the one without monitoring σ_n^{sb}, and for all $H > \bar{H}$, $\hat{\theta}_m < \min\{\hat{\theta}_n, \bar{\theta}_n\}$, which, by the same proposition, implies that the optimal contract is the one with monitoring σ_m^{sb}.

(iii) The following facts are relevant to the proof of this statement.

(a) Given condition (31) does not hold, $\theta_{fb} < \bar{\theta}_n$, and both θ_{fb} and $\bar{\theta}_n$ are not functions of γ.

(b) When $\gamma \to 0$, $\hat{\theta}_m \to \theta_{fb} < \bar{\theta}_n$ and $\hat{\theta}_n = 1$, hence, $\hat{\theta}_m < \min\{\hat{\theta}_n, \bar{\theta}_n\} = \bar{\theta}_n$. When $\gamma = H - t - \epsilon$ (where $H - t - \epsilon$ is the maximum γ that is allowed by assumption (1)), $\hat{\theta}_m = 1 > \min\{\hat{\theta}_n, \bar{\theta}_n\}$.

(c) It is obvious that $\hat{\theta}_m$ is an increasing function in γ. Also, $\hat{\theta}_n$ as a solution to equation (36) is a decreasing function of γ, since totally differentiate (36) with respect to θ_n and γ at $\hat{\theta}_n$, we have

$$d\hat{\theta}_n/d\gamma = g(\hat{\theta}_n) \left/ \frac{\partial O(\theta_m, \theta_n)^2}{\partial \theta_n^2} \right|_{\theta_n = \hat{\theta}_n} < 0$$

given that function O is strictly concave. Since $\bar{\theta}_n$ does not depend on γ, $\min\{\hat{\theta}_n, \bar{\theta}_n\}$ is also a decreasing function of γ.

Both $\hat{\theta}_m$ and $\min\{\hat{\theta}_n, \bar{\theta}_n\}$ are continuous functions of γ. By (c), $\hat{\theta}_m$ is increasing in γ and $\min\{\hat{\theta}_n, \bar{\theta}_n\}$ is decreasing in γ. By (b), as $\gamma \to 0$, $\hat{\theta}_m < \min\{\hat{\theta}_n, \bar{\theta}_n\}$, but at $\gamma = H - t - \epsilon$, $\hat{\theta}_m > \min\{\hat{\theta}_n, \bar{\theta}_n\}$. Therefore, there exists a $\bar{\gamma} \in (0, H - t - \epsilon)$ such that $\hat{\theta}_m = \min\{\hat{\theta}_n, \bar{\theta}_n\}$, for all $\gamma < \bar{\gamma}$, $\hat{\theta}_m < \min\{\hat{\theta}_n, \bar{\theta}_n\}$, and for $\gamma \in [\bar{\gamma}, H - t - \epsilon]$, $\hat{\theta}_m \geq \min\{\hat{\theta}_n, \bar{\theta}_n\}$. Hence by Proposition 4, the optimal contract is the one with monitoring σ_m^{sb} for $\gamma < \bar{\gamma}$, and it is the one without monitoring σ_n^{sb} for $\gamma \in [\bar{\gamma}, H - t - \epsilon]$.

(iv) Both $\hat{\theta}_n$ and $\bar{\theta}_n$ do not depend on ϵ. By assumption, there exists an $\bar{\epsilon} > 0$ satisfying assumptions (1) and (2) such that the optimal contract is the one with monitoring σ_m^{sb}. Hence, by Proposition 4, $\hat{\theta}_m = (t + \bar{\epsilon} + \gamma)/H < \min\{\hat{\theta}_n, \bar{\theta}_n\}$. Therefore, for any $\epsilon \leq \bar{\epsilon}$,

$$\hat{\theta}_m = \frac{t + \epsilon + \gamma}{H} \leq \frac{t + \bar{\epsilon} + \gamma}{H} < \min\{\hat{\theta}_n, \bar{\theta}_n\}.$$

That is, the condition for case (ii) of Proposition 4 is satisfied. It then follows the optimal contract is the one with monitoring σ_m^{sb}. ∎

References

Admati, A. and P. Pfleiderer (1994). "Robust Financial Contracting and the Role of Venture Capitalists," *Journal of Finance* Vol. XLIX, No. 2, 371-402.

Bernanke, B., M. Gertler, and S. Gilchrist (1996). "The Financial Accelerator and the Flight to Quality," *The Review of Economics and Statistics* Vol. LXXVIII, No. 1, 1-15.

Boyd, J.H. and B.D. Smith (1997). "Capital Market Imperfections, International Credit Market, and Nonconvergence," *Journal of Economic Theory* 73, 335–364.

Diamond, Douglas W. (1984). "Financial Intermediation and Delegated Monitoring," *Review of Economic Studies* 51, 393-414.

Diamond, Douglas W. (1991). "Monitoring and Reputation: The Choice between Bank Loans and Directly Placed Debt," *Journal of Political Economy* 99, 689-721.

Friedman, B., and K. Kuttner (1993). "Economic Activity and the Short-Term Credit Markets: An Analysis of Prices and Quantities," *Brookings Papers on Economic Activity*, 193-266.

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

Holmstrom, B. and J. Tirole (1997). "Financial Intermediation, Loanable Funds, and the Real Sector," *Quarterly Journal of Economics* Vol. CXII, 663-92.

Holmstrom, B. and J. Tirole (1998). "Private and Public Supply of Liquidity," *Journal of Political Economy* 106, 1-40.

Kashyap, K., J.C. Stein and D.W. Wilcox (1993). "Monetary Policy and Credit Conditions: Evidence from the Composition of External Finance," *American Economic Review* 83, 78-98.

Stiglitz, J. and A. Weiss (1981). "Credit Rationing in Markets with Imperfect Information," *American Economic Review* Vol. 71 No.3, 393-409.

Townsend, R. (1979). "Optimal Contracts and Competitive Markets with Costly State Verification," *Journal of Economic Theory* 22, 265-293.

Williamson, S. (1986). "Costly Monitoring, Financial Intermediation, and Equilibrium Credit Rationing," *Journal of Monetary Economics* 18, 159-79.

Williamson, S. (1987). "Costly Monitoring, Loan Contracts, and Equilibrium Credit Rationing," *Quarterly Journal of Economics* 102, 135-146.