The Allocation of Liability, Delegated Monitoring, and Modes of Financing

Tianxi Wang*

University of Essex

Abstract

Based on the allocation of liability, the paper examines all the modes of organizing finance and delegated monitoring in an economy with costly state verification. Besides *the mode of financial intermediation* (FI), a mode of direct finance, *Conglomeration*, accommodates delegated monitoring and obtains the benefit of diversification also, which, therefore, do not necessarily make FI viable, as the extant literature claims. The paper empirically predicts that the prevalence of bank financing could increase with the costs of banks' services or the rates of banks' loans. The paper is the first to examine the market organization of a financial service through the allocation of liability.

Key words: Allocation of Liability Modes of Financing Delegated Monitoring Financial Intermediation Conglomeration

*Email: wangt@essex.ac.uk. Correspondence: Department of Economics, University of Essex, Wivenhoe Park, Colchester, Essex, CO4 3SQ, UK. I am indebted to John Moore for his enormous time, patience, and advice. I am grateful to Gordon Kemp for his help in language. I thank for helpful comments Madhav Aney, Jan Bena, Erlend Berg, Sanjay Bernaji, Sudipto Bhattacharya, Christian Hellwig, Giovanni Ko, Xuewen Liu, Alan Morrison, Motty Perry, and the seminar participants at FMG and STICERD of the London School of Economics and at the ES European conference in the summer of 2007.

1 Introduction

Why is financial intermediation viable, given it apparently adds one level of agency problems? The reason, the literature starting with Diamond (1984) asserts,[1] is that this mode of financing accommodates monitoring service which addresses the friction of costly state verification and adds value under sufficient diversification. This literature of banking, however, fail to realize that the provision of the service does not have to be bundled with financial intermediation. That is, monitoring can be accommodated by modes of *direct finance* under which the expert who has the human capital for monitoring provides the service only, while funds flow directly from investors to end users (called entrepreneurs). The paper, to make up the gap, considers in a unified model all the possible modes of financing, including *the Mode of Financial Intermediation* (FI).

The first difficulty of comparing various modes of financing is to find a way of differentiating them. Apparently, funds change hands once more under FI than under a mode of direct finance. But this simple fact can hardly make any economic difference, as Holmstrom and Tirole (1997) confirm.[2] The paper is the first to suggest that the modes of financing should be differentiated according to the allocation of liability to investors. If it is taken by the service provider (i.e. the expert) alone, the mode is FI, where she is liable to repay investors, while entrepreneurs are liable to repay her. If entrepreneurs are liable to repay investors directly, the mode is of direct finance.

The model of the paper consists of three kinds of agents: Entrepreneurs supply productive projects; investors supply capital to finance the projects and resort to costly auditing to verify a project's outcome, success or failure;[3] and an expert supplies monitoring service through which

[1] See also Williamson (1986), Krasa and Villamil (1992), Winton (1995), Diamond (1996), and Hellwig (2000), among others; and for a good survey, see Gorton and Winton (2002).

[2] They recognize two ways of organizing the service of ex ante monitoring, depending only on the way in which funds flow, but they find these two ways give equivalent economic outcomes.

[3] Townsend (1979) for the first time studies the contractual problem with this type of friction. Mookerjee and Png (1989) extend it considering the case with stochastic auditing strategies.

she knows of the projects' outcomes. The paper first considers the case of two entrepreneurs, which allows for minimum diversification, and then extends to the case of a large number of entrepreneurs.

For the two-entrepreneur case, consider first the *Mode of Independent Finance* (IF), under which the two entrepreneurs are independently financed by investors, namely, an entrepreneur is liable to repay the investor of his project only, not those of the other. Under this mode, in the state when only one entrepreneur succeeds, the failed one declares default and is audited.

Consider then FI, under which the expert sets up a bank that finances the entrepreneurs and is financed by investors. In the above state, the repayment from the successful entrepreneur is shown to suffice for the bank to clear its liability to the investors, whereby it is not audited, and it does not audit the failed entrepreneur, since the bank (i.e. the expert) already knows of his actual failure through monitoring; therefore, costly auditing is saved.

The same kind of cost savings materializes under a mode of direct finance, named *Conglomeration,* which also accommodates monitoring. Under this mode, the liability to investors is taken by a conglomerate in which each project becomes a division, its entrepreneur the division manager, and the expert the chief financial officer who collects funds from the divisions to repay the liability of the whole conglomerate. As was the case under FI, funds from one successful division suffice to clear the whole liability and hence auditing occurs only when both divisions (i.e. projects) fail.

There are other modes of financing, all of them examined by the paper, but the real race is between FI and Conglomeration. The two modes have the same information structure;[4] they suffer the same kinds of incentive problems. They differ *only* in the allocation of liability. Under FI, entrepreneurs are liable to repay the expert who is in turn liable to repay the investors, whereas under Conglomeration, entrepreneurs are liable to repay the investors directly, while employing the expert for monitoring service. The liability is taken upon the *collateral* asset which the investment of the borrowed capital forms and from which the liability repayment funds

[4]Namely, each entrepreneur knows only the outcome of his own project, the expert both, the investors neither.

originates. The collateral for the investors is the pool of the projects under Conglomeration and is the bank asset (namely the bank's claim upon the projects) under FI. This difference matters, which I illustrate with a metaphor.[5]

Suppose each of the expert and the entrepreneurs has a box, and knows what is in it, to know which investors incur serious costs to open it, while the expert sees into an entrepreneur-box at minor costs. Under FI, investors give their capital to the expert and the collateral for them is the expert-box, while under Conglomeration, they give their capital to entrepreneurs and the collateral for them is the two entrepreneur-boxes put together. In any case, the contract stipulates that the investors are repaid before the expert and that if they are not fully repaid, they open the collateral box(es) and take whatever inside. What boxes compos the collateral, which is determined by the allocation of liability, makes differences as follows.

On the one hand, when default is declared, the investors have to open *one* expert-box under FI, but *two* entrepreneur-boxes under Conglomeration. If the reduction in the number of to-be-opened boxes saves auditing costs, FI has a *Number Advantage*.

On the other hand, the investors appropriate the collateral box(es) as the punishment of mis-reporting default. The higher the collateral's value, the harsher the punishment, and hence the smaller the probability in which auditing is needed.[6] The funds in the expert-box flow from the entrepreneur-boxes. Hence, the collateral for investors under FI is worth less than that under Conglomeration in any state. Thus, Conglomeration has a *Collateral Advantage*.

Therefore, the race is between FI's Number Advantage and Conglomeration's Collateral Advantage. FI arises in equilibrium if and only if the former outweighs the latter, namely, there are enough cost savings generated by banking technologies that help set up and rationalize a unified book of all the assets. The advantage of FI is thus not organizational, as suggested by Diamond (1984), but technological, as suggested by Diamond and Rajan (2001), where bankers have advantages in redeploying assets.

[5] I thank John Moore for suggesting the metaphor.

[6] The paper, as Mookerjee and Png (1989), allows stochastic auditing.

FI and Conglomeration are further compared for the case of a large number of entrepreneurs. Under both modes, the average incentive costs approach 0 with the number of entrepreneurs going to infinity. That is, both modes obtain the benefit of full diversification. Moreover, as was in the case of two entrepreneurs, FI dominates Conglomeration only if the former's number advantage outweighs the latter's collateral advantage.

For both cases, the higher the monitoring cost, the weaker the collateral advantage of Conglomeration over FI, and hence the greater the chance FI dominates Conglomeration:[7] Increase in the monitoring cost does not affect the value of Conglomeration's collateral for investors (the projects), but pushes up the value of FI's (the bank asset), since the increase forces the entrepreneurs to pay more to the bank (namely the expert) in order to satisfy her participation constraint. In fact, for the case of a large number of entrepreneurs, the paper shows that if the costs of auditing the bank under FI increase linearly with the number of entrepreneurs, then the chance of FI dominating Conglomeration goes to 0 when the monitoring cost goes to 0 and the number of entrepreneurs goes to infinity.

The comparative static result with respect to the monitoring cost bears empirical predictions at odds with what would be expected from cost or price considerations. If the monitoring cost is directly measurable, the comparative static predicts that the higher the cost of banks' monitoring service, the more prevalent the bank financing. If the monitoring cost is proxied by the average rate of bank loans, then the prediction is that the prevalence of bank financing, for example, across industries, increases with the average rate of bank loans.

The exercise of examining the market organization of a financial service through liability allocation bears a general implication upon the regulation of banks. In the recent crisis, thousands of billions of dollars were poured into bailing out banks, because, at least partly, these banks are believed indispensable in providing certain beneficial, or even essential, services for the financing of real sectors. However, for any given such service, banking is only one organizational mode of providing it and it could be provided under some modes of direct finance, as the paper shows for

[7] To be sure, if monitoring is too costly, it is not worth being provided under any mode.

ex post monitoring. By an exercise parallel to the paper, based on the allocation of liability, all the modes of organizing the service and finance can be examined in *a level playing field*. Only through this kind of exercises can the belief of banks' indispensability be justified.

Relation to Literature

As aforementioned, the paper extends the literature on the viability of FI, by considering all the other modes of financing besides FI and IF, the two modes examined by the literature; see, among others, Diamond (1984, 1996), Williamson (1986), Krasa and Villamil (1992), Winton (1995), and Hellwig (2000); and for a good survey, see Gorton and Winton (2002). Moreover, the present paper draws two distinctive conclusions. First, whereas that literature maintains that the benefit of diversification is the key to the viability of FI, the paper finds that this benefit, as also obtained by Conglomeration, is irrelevant, and the key is instead the number advantage of FI. Second, the paper finds that the monitoring cost favors FI, whereas the opposite is predicted in that literature, where only FI accommodates monitoring and incurs the cost.

This paper adds to the literature that compares several modes of financing in a unified model; see Gertner, Scharfstein and Stein (1994), Boot and Thakor (1997), and Bond (2004), among others.[8] None of these papers takes the perspective of liability allocation: Gertner, Scharfstein and Stein (1994) differentiate the various modes based on the allocation of ownership of the project, Boot and Thakor (1997) based on whether the agents within a mode compete or cooperate, and Bond (2004) based on the distribution of seniority. Nor do they address the organization of a particular service in terms of how to combine human capital and physical capital: In the first paper of the three, the two types of capital are both owned by the investor, thus presumedly bundled together; and the latter two papers do not have any particular service.

The paper is related to the literature on how an informed agent profits from his private information on a financial asset: Either he directly sells his information to clients, or indirectly, he invests for them with his information to earn a commission; see Admati and Pfleiderer (1990),

[8] By mode of financing, I mean a way of organizing financial markets, rather than a type of securities (debt or equity) to be issued, which is what the term means in some papers (e.g. Constantinides and Grundy (1989)).

Brennan and Chordia (1993), and Biais and Germain (2002) among others. Compared to this literature, the papers presents a richer set of organizational modes: While FI is arguably parallel to the indirect way of selling information, Conglomeration is by no means parallel to the direct sale and demands more of organization than the latter, and there are other (three) modes.

The rest of the paper is organized as follows. Section 2 sets up the model. Section 3 examines for the two-entrepreneur case all the possible modes of financing one by one, and then finds the equilibrium mode. Section 4 extends it to the case of a large number of entrepreneurs. Section 5 concludes. Technical proofs are relegated to Appendix B, while Appendix A discusses some assumptions of the model.

2 The Model

The model is isomorphic to the two-entrepreneur case of Diamond (1984). I first describe the model and then return at the last subsection to its relationship to Diamond (1984).

2.1 Agents and Production

The economy has two dates: Investment and contracting occurs at T_0 and return and repayment at T_1. There are two entrepreneurs, many experts, and a lot more potential investors.

Each of the two entrepreneurs, E_1 and E_2, has an independent and identical project. A project needs a unit capital of investment, and returns R with probability q, and nothing with probability $1-q$. Both entrepreneurs are penniless at T_0. Each potential investor has a small amount of capital, but in aggregation capital is abundant. Experts have neither physical capital nor projects, but have the human capital for monitoring (see below for its meaning). An entrepreneur can only accommodate one expert.

All the agents are risk neutral and protected by limited liability, and the risk free rate is 0. Entrepreneurs have all the bargaining power, that is, equilibrium is driven by maximizing their expected profit.

2.2 Information and Collusion

The friction of financing a project is that only the entrepreneur costlessly observes its outcome, success or failure. For others to find out the outcome, two information technologies are feasible, with different costs and strength.

The weak technology is monitoring. If an expert has been monitoring a project from T_0, then she knows of its outcome at T_1, but her knowledge is soft and not sufficient to convince others of what she knows. The cost of monitoring a project is m.[9] The strong technology is auditing, which produces before the court of law hard evidence of the outcome after it has been realized. The cost of auditing a project is C. In accordance with the information strength, $m < C$. Only experts know how to monitor, but the investors can resort to auditing, provided they afford C collectively.[10]

An expert can observe the outcome of a project at a minor cost. If she never colludes with the entrepreneur, the investors can simply rely on her word of mouth to know the outcome and the unobservability of the outcome is not a problem any more. To exclude such a trivial solution, I allow for all possible collusion between the expert and the entrepreneurs; for that purpose, I assume that any side transfer between some or all of these non-investors is outright observable to none but the parties involved. This collusion problem is implicitly assumed by the literature of banking following Diamond (1984), but is explicitly examined in the present paper.

The problem of collusion plainly precludes Maskin-Moore-Repullo mechanisms from functioning for an entrepreneur-expert pair: Monitoring adds no value if one expert monitors only a single entrepreneur because the two would act as one party and the information structure is not change. However, it may add value if one expert monitors both entrepreneurs, because then it qualitatively changes the information structure by engendering a party that knows the

[9] As experts are penniless, the monitoring cost is non-pecuniary, for example the opportunity cost of the time devoted to monitoring.

[10] Notice that here the auditing cost does not vary with the number of the agents to whom the truth is disclosed, which, however, is assumed by Bond (2004). This differenc explains why seniority plays no role in the present paper but a key role in Bond (2004).

outcomes of both projects, namely the monitor. The presence of such a party, I will show, is indispensable to materializing the benefit of diversification. Together with the restraint of only one expert allowed for one entrepreneur, it follows that if monitoring service is provided at all, it is provided by one expert only, who is to monitor both entrepreneurs. This expert is labelled as X.

Denote by $C_2 \geq C$ the cost of auditing X's account (or pocket).

For simplicity, I assume that the action of monitoring is contractible; adding the according moral hazard problem gains nothing but the complexity of exposition. Thus, if X is employed, she will monitor and then at T_1, the information structure is as follows. An entrepreneur knows the outcome of his own project, but not the other's; X knows both; and investors know neither, but can spend C to check a project's outcome or C_2 to check what is in X's pocket.

Additional assumptions are made below.

Assumption 1: At T_0 the investors commit to a *stochastic* auditing strategy.

The commitment of the investors (or the lender) is commonly assumed by the literature on costly state verification to get rid of voluntary default.[11] For this paper, it facilitates the mechanism design approach and merely plays a technical role. In its absence, so the investors decide the probability of auditing ex post, the trade-off between FI and Conglomeration is still the number advantage versus the collateral advantage.

Stochastic auditing is disallowed by Diamond (1984, 1996), but studied by Mookerjee and Png (1989). As Hart (1995) points out, it generally allows of Pareto improvement. In the present paper, the assumption of stochastic auditing not only technically facilitates the mechanism design approach but also uncovers the indispensability of monitoring to materializing the benefit of diversification; see the remark following Lemma 2. Further discussion of the assumption is to be seen in Appendix A.

[11] If the lender cannot commit ex ante and instead decides ex post whether to audit the borrower, then the equilibrium features a positive probability of voluntary default: If the borrower never voluntarily defaults, the lender will never audit; but then the borrower always voluntarily defaults.

Hereinafter, the investors will be dealt with as one party.

Assumption 2: $S \equiv qR - (1-q)C \geq 1$, $(1-q)C \geq \frac{qR}{2}$, and $0 < m < C\frac{(1-q)q^2R}{(q^2R+2(1-q)S)S}$.

S is, in some sense, the social revenue of a project, which returns revenue R in case of success and incurs cost C in case of failure; hence $S \geq 1$ roughly means that a project has a positive net present value. The assumption $(1-q)C \geq \frac{qR}{2}$ ensures that the "expected" auditing cost ($(1-q)C$) is large enough to leverage up various modes of financing; what happens if instead $(1-q)C < \frac{qR}{2}$ is examined in detail in Appendix A. The last part of the assumption stipulates that monitoring service is not too costly; otherwise it will certainly not be provided under any mode.

Assumption 3: Securities (i.e. contracts) issued to investors bear repayments that weakly increase with the economic fundamental.

That is, the investors are promised to be repaid more when more projects succeed. This feature would be a property of the equilibrium securities, if I assumed that the investors can collude with X by arguments in a line of Innes (1990), or if I restricted, as the literature of banking following Diamond (1984) did, the contract between an entrepreneur and the expert is binary, namely, independent of the outcome of the other entrepreneur's project. With the assumption removed, nothing below but Lemma 3, which concerns the optimal contractual arrangement under FI, would change, and the change arises only for $q < 0.63$. The assumption is further discussed in Appendix A.

2.3 The Allocation of Liability and Modes of Financing

In this economy, the mode of financing depends on the allocation of liability. The relationship of liability comes with the transfer of capital: If a party A takes capital from another party B at T_0, then A bears the responsibility (namely *liability*) to repay B at T_1. The asset formed by A's investment of the capital so taken, in this model, is all that A possesses, and is thus the only source of funds available to repay B; this asset is thus the *collateral* for B. In this economy, the friction is that the state of the collateral is not outright verifiable to the court of law; it is verifiable only through costly auditing. To incentivize A to honor the liability, any contract A issues to B must therefore lay down the following stipulations, that is, the liability of A to B must bear the following implications.

I, *Contingent on A's report of the performance of the collateral, B has the rights to audit (any part of) the collateral to verify whether A reports the truth.*

II: *If the auditing uncovers a fraud in A's report, B has the rights to appropriates part or the whole of the collateral.* Ex ante, the harsher the punishment, the lower the incentive A has to lie, namely, the more commitment A makes to repay the liability, which benefits A. Therefore, *the optimal contracts always entail the maximum punishment, that is, the appropriation of the entire collateral.*[12]

III: *A is allowed to dispose of (for example consume) anything out of the collateral if and only if B declares A's liability to him has been cleared to the max.*

Based on the allocation of liability, the following modes of financing can be envisaged.

1. *Independent Finance* (IF): Each entrepreneur independently takes the liability to his investors. That is, two projects are financed separately; and each entrepreneur is liable to repay his investors only, not to the investors of the other project, and the collateral for his investors is his project, who will get nothing if it fails.

2. *Joint Liability without Monitoring*: The two entrepreneurs, without using monitoring service, take the liability jointly. That is, the two project are financed jointly and pooled together as the collateral for all the investors, who get repaid (possibly in part) if any project succeeds.

3. *The Mode of Financial Intermediation* (FI): All the entrepreneurs are liable to repay X, and X alone is liable to repay investors. That is, X becomes a bank that takes capital

[12] The appropriation might be triggered by the auditing of only part of the collateral. In this case, appropriating the whole collateral may require to audit all the other parts of the collateral beforehand, especially when the lender is the investors.

from investors and uses it to finance all the entrepreneurs at T_0, and at T_1 is repaid by the entrepreneurs out of their projects' revenues and repay investors out of these repayments. Put differently, an entrepreneur's collateral for the bank is his project and the bank's collateral for the investors is the bank asset, namely, the bank's claims upon the projects' revenues.

4. *Conglomeration*: The two entrepreneurs jointly take the liability to investors and meanwhile employ X for monitoring service. More concretely, the liability to investors is taken by a conglomerate in which each project becomes a division run by the entrepreneur and monitored by X who forms the headquarters. The collateral for the investors is the asset of the conglomerate, namely the pool of the projects, as under mode 2.

5. *The Mix Mode*: One entrepreneur and X jointly take the liability to investors. That is, the liable entity both runs this entrepreneur's project and finances the other project as the intermediary. The collateral for the investors then consists of the directly financed project and the intermediary asset within the other project

Remark: Implication III of liability above rules out those hybrid modes under which the asset to be audited by investors is not the same as the asset to be appropriated by them; for example, the mode under which investors audit only the account of X, while they appropriate the projects in case of uncovering a fraud. Under this mode, the entrepreneurs would always declare default, fill nothing into the X's account, and promise to share the projects' revenues with her when the auditing of her account is finished. According to implication III, the revenue is kept untouched during the auditing, thus ready to be shared afterwards. Therefore, the promise is credible, and the collusion works.

By contrast, under FI, where investors audit the account of X (now the bank) only, this collusion will not work, since the promise of sharing the revenues is not credible: The entrepreneurs will dispose of all the revenues while the bank (namely the expert) is being audited, leaving nothing for her. They can do that under FI, because, according to III, they are free to dispose of what remains in their pockets as soon as the bank declares that they have failed and done the maximum to clear their liabilities to it.

It will be shown that the five modes listed above exhaust all the possible modes of financing. Note that Diamond (1984), and the literature that follows, only considers IF and FI.

2.4 Timing and Mechanisms

The timing of events is as follows.

Date $\mathbf{T}_0$: The entrepreneurs choose the mode of financing, in a cooperative way.[13] Then they design the contractual arrangement (called *mechanism*) between them and investors, and X if she is involved, under the chosen mode. I only consider the mechanisms symmetric between the two entrepreneurs.

The securities are then issued to potential investors. Those who have bought the securities, as a whole, commit to a probably stochastic auditing strategy.

Date $\mathbf{T}_1$: The outcome of each project is realized and observed by the entrepreneur and X (if she is involved). X may arrange collusion with the entrepreneurs. Then, the entity liable to repay the investors reports the performance of its asset, namely the collateral, and is ready to repay them accordingly. Contingent on the report, they audit the collateral with the probability committed at T_0. If auditing is not actually exercised or finds no frauds in the report, they are repaid based on the report; if auditing is exercised and uncovers any fraud, they appropriate the entire collateral.

Only investors exercise auditing in equilibrium path; If X is involved, an entrepreneur never dares to evade his liability to her – if he lied about his success, which X knew through monitoring, X would expose his lie *with certainty* to the court of law, making him lose all. Thus, under any mode, a mechanism consists of the auditing strategy of the investors and the repayment to them,

[13] At this time, they have no conflit of interests and act as one and the same designer. I abstract away the game probably played between them at this time.

and the repayment to X (if being involved), all contingent on the reported performance (or *state*) of the collateral for the investors.

Specifically, under IF, an entrepreneur's mechanism is represented by (d, l): Either his project succeeds and then he repays d to his investors; or it fails and then is audited with probability l. Unlike IF, modes 2, 3, 4 and 5 all feature joint liability, that is, the revenue of the collateral for investors depends on the outcomes of both projects, rather than that of a single one. The collateral thus has three states, $s = 0, 1$, and 2, defined by the number of successful projects, occurring with probabilities $(1-q)^2$, $2q(1-q)$, and q^2 respectively. A mechanism is thus represented by $\{D_s, l_s\}_{s=0,1,2}$ under mode 2, and $\{d_s, D_s; l_s\}_{s=0,1,2}$ under modes 3, 4 and 5 where X is involved: D_s denotes the payment to the investors in state s, d_s the gross payment to X, and l_s the probability of auditing the collateral.

2.5 The Relationship to Diamond (1984)

The model is isomorphic to the two-entrepreneur case of Diamond (1984), as follows.[14]

A. Delegated monitoring is directly assumed by this paper, whereas justified with avoiding cost replication by Diamond (1984). I make the variation to get around the concern that if multiple monitors are feasible, some mechanism can be designed to cheaply elicit the information between them, with auditing not needed at all.[15] In any case, the variation affects nothing of this paper's analysis: Even if delegated monitoring is justified in Diamond's way, it is still true that the service can be accommodated by several modes (modes 3, 4 and 5) rather than solely by FI, as presumed by Diamond (1984).

B. The auditing costs of this paper correspond to the non-pecuniary penalties of Diamond

[14] While the extension of the model to the case of a large number of entrepreneurs, which will be addressed in Section 4, is isomorphic to the complete version of Diamond (1984), the two-entrepreneur case can be thoroughly examined with the mechanism design approach and is thus put as the main model.

[15] For example, the investors of an entrepreneur can randomly choose ten persons as monitors whose identities are secret to each other, let them report their findings separately to the investors, and harshly punish any who reports different things from the majority. This mechanism is likely to work.

(1984). Both are the deadweight loss incurred when (actual) default happens, and in both papers, to reduce this loss with the help of monitoring drives the modes of financing other than IF. One difference is that the auditing costs are *ex post* borne by the investors in this paper, whereas the penalties, though imposed by the investors, are borne by the borrower (an entrepreneur or the bank) in Diamond (1984). This difference affects nothing of equilibrium, however, since in this paper the auditing costs are borne *ex ante* by the entrepreneurs through the IR constraint for the investors.

C. This paper, as most studies on the costly state verification, assumes commitment by the investors, with which Diamond (1984) dispenses due to specific features of his model. The cost of verification, in whatever form, poses a deadweight loss. Therefore, ex post, negotiating it away is in the interest of all the parties involved, if there is a way for them to share the value saved. However, no such a way exists in Diamond (1984), where the verification cost is non-pecuniary and borne by the borrower, which altogether mean that the investors can gain nothing from the saving of the penalties.

So far, I finish setting up the model. In the following section, I first prove the exhaustiveness of modes 1 through 5 above, then examine the five modes one by one, then find the equilibrium mode by comparing between them, and finally present some empirical predictions of the model.

3 The Modes of Financing

3.1 All the Possible Modes

The lemma below lays down the exhaustiveness of the list above and holds true independent of Assumptions 1, 2 and 3. Immediately after the proof, an intuition is expounded regarding how these modes perform differently.

Lemma 1 *IF, Joint Liability without Monitoring, FI, Conglomeration, and the Mixed Mode exhaust all the possible modes of financing.*

Proof. Consider what the collateral for investors could be. The funds to repay them, in the end, originate from the revenue of the real sector projects, either directly flowing out of them, or indirectly from the asset of which the return originates from them, or in some way mixed between the former two channels.

For the first case, either the two projects become two separate collateral assets, which gives rise to IF, under which monitoring is useless; or the two projects are pooled into one collateral asset, which gives rise to Conglomeration or Joint Liability without Monitoring, depending on whether monitoring is provided or not.

The second case, where the asset within the projects forms the collateral asset, leads to a mode that features financial intermediation. Only an expert (X) can sensibly play the role of the intermediary because it demands an information advantage over the entrepreneurs. Therefore, the case gives rise to FI.

In the mix case, the collateral asset consists of one project plus the intermediary asset within the other. Again, only X can sensibly be the intermediary. Thus, the liability is taken jointly by her and the entrepreneur of the directly financed project, which gives rise to the Mixed Mode.

■

As we saw, only investors exercise auditing in equilibrium path. To incentivize the entity that takes the liability to them to report the truth, they must commit to auditing with high enough probabilities. But auditing entails serious costs, and its exercise should be minimized. The allocation of liability affects the exercise of auditing through Implications I and II of liability.

Implication I determines the number of assets to be audited, and as we saw in the introduction, leads to the number advantage of FI.

Implication II, we saw in the introduction, leads to an inverse relationship between the value of the collateral and the probability in which auditing is exercised: The higher the value of the collateral, the harsher the punishment for telling a lie, the less the incentive of lying, and hence the smaller the probability of auditing needed. Implication II, as we saw in the introduction, drives the collateral advantage of Conglomeration.

I move on to examine the five modes in order. Under each mode, I apply a mechanism design approach to find the optimal mechanism. By the revelation principle, I only consider direct mechanisms with truth-telling equilibria. A feasible mechanism thus has to satisfy the incentive compatibility (IC) constraints for the entity liability to repay investors to report the truth, and the individual rationality (IR) constraints for the investors and X (if involved) to participate.

3.2 IF and Joint Liability without Monitoring

Under IF, for each entrepreneur, a mechanism is (d, l): If he reports that his project succeeds, he repay d to his investors; if he reports failure, his project is audited with probability l.

The IC constraint for the entrepreneur to honestly report his success is:

$$d \leq lR \tag{1}$$

The left hand side (LHS), d, is his outlay if he reports honestly his success, whereas the right hand side (RHS) is what he expects to lose if he lies and claims to have failed: Then with probability l, the project is audited, the fraud uncovered, and the whole revenue, R, appropriated by the investors; with probability $1 - l$, the project is not audited and he escapes the liability.

The IR constraint for the investors is:

$$1 + (1 - q)lC \leq qd \tag{2}$$

The LHS is their costs of financing the project: They contributes 1 unit of capital for its investment at T_0, and at T_1 when it fails, they expect to incur costs C to audit it with probability l. The RHS is what they expect to get in gross.

The entrepreneur chooses (d, l) to minimize d subject to the IC and IR constraints above. Both constraints are binding in the optimization. Therefore, the optimal mechanism, denoted by $\{d^I, l^I\}$, is $\{\frac{R}{S}, \frac{1}{S})$, where $S \equiv qR - (1 - q)C$. Note that $S \geq 1$ by Assumption 3; otherwise $l^I > 1$ and IF is not feasible.

Move on to mode 2, Joint Liability without Monitoring, under which the entrepreneurs take the joint liability, but do not observe each other's outcome. The mode is proven equivalent to IF in the following paragraph.

Under mode 2, a mechanism of the mode is $\{D_s, l_s\}_{s=0,1,2}$: In (reported) state s, the investors are repaid with D_s and audit a (reportedly) failed project with probability l_s. $D_0 = 0$ by limited liability, and for $s > 0$, each successful entrepreneur equally contributes $\frac{D_s}{s}$. A successful entrepreneur, if telling truthfully his success, expects to outlay $d^J = (1-q)D_1 + q\frac{D_2}{2}$. If lying, his project is audited with probability $l^J = (1-q)l_0 + ql_1$. Then, the IC constraint is $d^J \leq l^J R$, equivalent to (1), the IC constraint under IF. The IR is $2 + C((1-q)^2 \cdot 2l_0 + 2q(1-q) \cdot l_1) \leq 2q(1-q)D_1 + q^2 D_2$. Its RHS equals $2qd^J$ and RHS equals $2 + 2l^J C$. Therefore, the IR constraint is equivalent to $1 + l^J C \leq qd^J$, equivalent to (2), the IR constraint under IF. With both constraints under mode 2 equivalent to the counterparts under IF, mode 2 is equivalent to IF.

To summarize,

Lemma 2 *Under IF a successful entrepreneur outlays $d^I = \frac{R}{S}$ and a failed project is audited with probability $\frac{1}{S}$. Mode 2 is equivalent to IF, that is, without monitoring, the joint liability alone does not materialize the benefit of diversification.*

Remark: Even if investors are risk averse, the equivalence of mode 2 to IF still holds: Although, at first glance, mode 2 gives the advantage of delivering a diversified portfolio to investors, the investors can obtain the same portfolio under IF by diversifying across the two entrepreneurs.

The assumption of stochastic auditing is crucial to the indispensability of monitoring. In its absence, here is an counter example. Under mode 2, let $D_2 = 2D_1 = 2D; l_0 = 1, l_1 = l_2 = 0$. The IR constraint is $qD \geq 1 + (1-q)^2 C$ and the IC is $(1-q)R \geq D$. If $q(1-q)R \geq 1 + (1-q)^2 C \Leftrightarrow (1-q)S \geq 1$, there exists a range of (q, R, C) within which both constraints are satisfied. Then Mode 2 dominates Mode 1: Under the former, auditing occurs only when both projects fail, whereas under the latter, a project is audited whenever it fails.

At first glance, the joint liability alone seems to be able to save auditing costs in state 1: If the liability of the failed entrepreneur is repaid by the successful one, the former is saved from being audited. This potential cost saving is, however, thwarted by the following incentive compatibility problem. Exactly because the joint liability makes a failed project less likely audited, it gives a successful entrepreneur a higher incentive to hide his success and let the other one shoulder the whole liability. Monitoring helps alleviate this problem, because it brings about a party (namely X) who knows both outcomes and whose silence an entrepreneur, if hoping to hide his success, has to buy, which is costly.

There are various modes that accommodate the monitoring service. The first is FI, examined in the following subsection.

3.3 FI

Under FI, each entrepreneur is liable to repay X upon his project and X is liable to repay investors upon the bank asset. In equilibrium path, X does not need to audit the projects – the entrepreneurs never dare to evade their liabilities to the bank, as we saw – and the investors audit the bank at costs $C_2 \geq C$.

Definition 1 *FI has a "Number Advantage" if* $C_2 < 2C$.

The total costs of auditing the two projects are $2C$. $C_2 < 2C$ means that banking technically saves auditing costs. This cost saving is called "Number Advantage", as it owes to the fact that the bank sets up *one unified book* of all the assets and rationalizes it. The number advantage can be measured by $\frac{2C}{C_2}$.[16]

In this subsection, the marks of some equations are suffixed with "b" to indicate that they are peculiar to FI, where monitoring is provided by the "b"ank.

[16] Note that no number advantage is assumed by Diamond (1996), where the costs are the destruction of the low output, L, and thus $C = L$ and $C_2 = 2L$.

A mechanism is $\{d_s, D_s; l_s\}_{s=0,1,2}$: In (reported) state s, each successful entrepreneur repays d_s to X (the bank) and she then passes D_s to the investors; and the investors audit the bank with probability l_s as they have committed at T_0.

By limited liability for X and the entrepreneurs,

$$D_0 = d_0 = 0,\ D_1 \leq d_1 \leq R \text{ and } D_2 \leq 2d_2 \leq 2R \tag{LL}$$

Consider the IC constraints for the bank to truthfully report the performance of its asset; only the possibly binding constraints are presented in this subsection, the non-binding ones put in Appendix B. As we saw, the main problem in connection with using monitoring service is that X could collude with the two entrepreneurs. There are two kinds of collusion. One is grand collusion between all the three non-investors, and the other is partial collusion between X and one entrepreneur. The difference between the two is that the former targets the investors while the latter targets the entrepreneur outside the collusion. The two kinds of collusion give rise to two groups of IC constraints.

Consider first the grand collusion proof constraints, which command that in any state, the three non-investors as a whole are better off with truth-telling than with mis-reporting, or equivalently, the amount outlaid to the investors with truth-telling is no greater than that with mis-reporting.

The IC constraint for the bank not to mis-report state 2 as state 1 is:

$$D_2 \leq l_1 \cdot 2d_2 + (1 - l_1)D_1 \tag{G21b}$$

In state 2, the bank repays D_2 to the investors if telling the truth, whereas the the RHS is what the bank expects to lose if instead it lies that the state is 1: Then, with probability l_1, the bank is audited and the whole asset, worth $2d_2$ in state 2, is appropriated; with probability $1 - l_1$, the report is accepted as the truth and accordingly the bank outlays D_1 to the investors.

The IC constraint for the bank not to mis-report state 2 as state 0, (G20b), is not binding.

The IC constraint for the bank not to mis-report state 1 as state 0 is:

$$D_1 \leq l_0 d_1 \tag{G10b}$$

In state 1, truth telling leads the bank to outlay D_1, whereas mis-reporting the state as state 0 leads the whole bank asset, worth d_1 in the state, appropriated with probability l_0.

The IC constraint for the bank not to mis-report state 1 as state 2 is never binding, since $D_2 \geq D_1$ by Assumption 3.

Consider then the constraints that prevent partial collusion, by which X may exploit a successful entrepreneur.[17] In state 2, she may arrange collusion in which she collects $t < d_2$ from E_1 to buy his silence when she declares that only E_2 has succeeded. E_1 is happy to gain $d_2 - t > 0$. With the collusion, X collects d_1 from E_2, t from E_1, and outlays D_1 to the investors, thus obtaining $d_1 + t - D_1$, if this fraud is not uncovered by auditing, which happens with probability $1 - l_1$; otherwise, she obtains nothing. Without the collusion, she obtains $2d_2 - D_2$ in net. Therefore, *X has no incentive to arrange any partial collusion if and only if* $2d_2 - D_2 \geq (1 - l_1)(d_1 + t - D_1)$ *for any* $t < d_2$*, or equivalently,*

$$2d_2 - D_2 \geq (1 - l_1)(d_1 + d_2 - D_1) \tag{P2b}$$

The partial collusion proof constraint in state 1, (P1b), is not binding.[18]

The investors, facing the bank contract $\Psi \equiv \{d_s, D_s\}_{s=0,1,2}$, commit to an auditing strategy, denoted by $\{l_s^B(\Psi)\}_{s=0,1,2}$, that minimizes the auditing costs subject to the constraints proof against grand collusion (namely (G21b), (G10b), and (G20b)), but not to those proof against partial collusion: They do not care how the bank deals with the entrepreneurs. For example, so long as $D_2 = D_1$, they will not audit the bank in state 1 ($l_1 = 0$), but if $d_2 - D_2 < d_1 - D_1$ (equivalent to $d_2 < d_1$), (P2b) will be violated and some successful entrepreneur will be exploited in state 2. We saw $l_2^B(\Psi) = 0$ and by (G21b), $l_1 \geq \frac{D_2 - D_1}{2d_2 - D_1}$, so the minimum l_1, $l_2^B(\Psi)$, equals $\frac{D_2 - D_1}{2d_2 - D_1}$, making (G21b) binding.

[17] A failed entrepreneur has nothing to be exploited.

[18] But note that the reason is different from that for (G12b) to be nonbinding. Reporting state 1 as state 2 certainly damages the group of the three non-investors as a whole, but may nevertheless benefit the subgroup of X and the failed entrepreneur.

I move on to the IR constraints for the investors and for X respectively. The IR constraint for the investors is:

$$2 + [(1-q)^2 \cdot l_0 + 2q(1-q)l_1]C_2 \leq q^2 D_2 + 2q(1-q)D_1 \qquad \text{(IR-Ib)}$$

The LHS are the costs for them: the investment costs (2) plus the expected auditing costs, and the RHS gives what they are repaid in gross.

The IR for X is:

$$2m \leq q^2(2d_2 - D_2) + 2q(1-q)(d_1 - D_1) \qquad \text{(IR-X)}$$

She incurs $2m$ in monitoring the two projects and gets the difference between the asset paid in and the liability paid out in states 1 and 2.

The entrepreneurs' problem is then as follows.

Problem 1 $\min_{\{d_s;D_s,l_s\}_{s=0,1,2}} 2q(1-q)d_1 + q^2 \cdot 2d_2$, *subject to*

(1): $l_s = l_s^B(\Psi)$ for $s = 0, 1, 2$, where $\Psi \equiv \{d_s, D_s\}_{s=0,1,2}$.

(2): (LL), (P2b), (P1b), (IR-Ib), (IR-X), and $D_1 \leq D_2$ (Assumption 3).

The solution depends on whether (IR-X) is binding or not. We start with the case of it unbinding; for the case let V^B denote the net rent to X. Let $F \equiv \frac{2+(1-q)^2 C_2}{2q-q^2}$.

Lemma 3 *If (IR-X) is not binding, the optimal mechanism under FI is: $D_0 = d_0 = 0$, $D_1 = D_2 = d_1 = d_2 = F$, $l_0 = 1$ and $l_1 = l_2 = 0$; and X expects to get gross rent $q^2 F$. Therefore, (IR-X) is not binding if and only if $m < \frac{1}{2}q^2 F$, and in that case, $V^B = q^2 F - 2m > 0$.*

Proof. See Appendix B. ■

Critical to the *shape* of the mechanism is $l_1 = 0$: It implies $D_1 = D_2 \equiv D$ by binding (G21b) (namely the liability contract is debt); Furthermore, $l_1 = 0$ and $D_1 = D_2$, by the binding (P2b), imply that $d_1 = d_2 \equiv d$ (namely the contract between an entrepreneur and X is binary).

$l_1 = 0$ is driven by the trade-off between auditing costs and the rent to X. She obtains the rent because she is the only one who knows of the state of the economy before auditing. However, auditing eradicates this information advantage of her, and thus squeezes her rent, by disclosing the state to the public. The trade-off tilts to set $l_1 = 0$, to minimize the exercise of auditing, because the auditing cost C is assumed large enough by Assumption 2.[19]

The *level* of D is found by binding (G10b): $D = l_0 d$. To find the *level* of d, substitute into the binding (IR-Ib) $d_1 = d_2 \equiv d$ and $D_1 = D_2 = l_0 d$, and rearrange:

$$l_0 = \frac{2}{(2q - q^2)d - (1-q)^2 C_2} \tag{3}$$

When (IR-X) is not binding, l_0 can move freely. As d decreases with l_0 by (3), it is minimized at $l_0 = 1$, with the minimum equal to $\frac{2+(1-q)^2 C_2}{2q-q^2} \equiv F$. So we have lemma 3 above.

By (3), l_0 decreases with d, which measures the value of the bank asset (the collateral under FI). This link is due to Implication II of collateral, which leads to an inverse relationship between the value of the collateral and the probability of auditing needed.

Move on to the case in which (IR-X) is binding. Substitute $d_1 = d_2 \equiv d$ and $D_1 = D_2 = l_0 d$ into binding (IR-X):

$$[q^2(2 - l_0) + 2q(1-q)(1-l_0)]d = 2m \tag{4}$$

Let $d^B(m)$ and $l^B(m)$ defined as the solution for d and l_0 of the simultaneous equations of (3) and (4), at a given m.

Lemma 4 *If $m \geq \frac{1}{2}q^2 F$, (IR-X) is binding, and the optimal mechanism is: $D_0 = d_0 = 0$, $D_1 = D_2 = l^B(m)d^B(m)$, $d_1 = d_2 = d^B(m)$, $l_0 = d^B(m)$, and $l_1 = l_2 = 0$, where $(d^B(m), l^B(m))$ is the solution for (d, l_0) of (3) and (4)*

---

[19] For example, suppose $m = 0$ and consider the alternative mechanism: $\frac{D_2}{2} = D_1 = d_1 = d_2 = \frac{2+(1-q^2)C_2}{2q}; l_0 = l_1 = 1$. The mechanism is feasible. It invokes auditing in state 1 as well as in state 0, but gives Ms X no ex post rent in any state. This mechanism is inferior to the mechanism specified in the lemma: $\frac{2+(1-q^2)C_2}{2q} > \frac{2+(1-q)^2 C_2}{2q-q^2} \Leftrightarrow C_2(1-q)(3-q) > 2$, which holds true because $C_2 \geq C$ and $C(1-q) \geq 1$, derived from Assumption 2: $2(1-q)C - (1-q)C \geq qR - (1-q)C \geq 1$.

Proof. If $m \geq \frac{1}{2}q^2 \cdot F$, by Lemma 3, (IR-X) is binding. So X get no net rent. With the concern about her rent removed, the trade-off between the rent and the auditing costs drives $l_1 = 0$, to minimize the latter. It follows, as was in the case of (IR-X) unbinding, that $d_1 = d_2 \equiv d$ and $D_1 = D_2 \equiv l_0 d$, which altogether implies (3) through binding (IR-Ib) and (4) through binding (IR-X). Therefore, $d = d^B(m)$ and $l_0 = l^B(m)$. ∎

Note that $d^B(m)$ strictly increases with m; intuitively, the higher the monitoring cost, the more is needed to pay X (the bank) in order to satisfy her IR constraint. Then $l^B(m) = \frac{2}{(2q-q^2)d^B(m)-(1-q)^2C_2}$ strictly decreases with m. Moreover, both functions are independent of R.

Put the two cases together, the grand collusion proof constraints determine the shape of the liability contract ($D_1 = D_2 \equiv D$), partial collusion proof constraints the shape of the asset contract ($d_1 = d_2 \equiv d$), and (IR-Ib) determines the *level* of D, (IR-X) the *level* of d.

To get a uniform formula for the mechanisms of both cases, extend to define $(d^B(m), l^B(m)) = (F, 1)$ if $m < \frac{1}{2}q^2 F$. And let $C^B \equiv (1-q)^2 C_2 l^B(m)$ denote the expected auditing costs under FI. Then, the net rent to X is $V^B = \max(0, q^2 \cdot \frac{2+C^B}{2q-q^2} - 2m)$: If $m < \frac{1}{2}q^2 F$, then $l^B(m) = 1$ and $F = \frac{2+C^B}{2q-q^2}$; then, the max term is strictly positive and equals $q^2 F - 2m$; and otherwise, $q^2 \cdot \frac{2+C^B}{2q-q^2} < q^2 F \leq 2m$, and hence the max term equals 0 as V^B. To summarize:

Proposition 1 *Let* $F \equiv \frac{2+(1-q)^2C_2}{2q-q^2}$. *Under FI, the optimal mechanism is:* $D_0 = d_0 = 0$, $d_1 = d_2 = d^B(m)$, $D_1 = D_2 = l^B(m)d^B(m)$, $l_0 = l^B(m)$, *and* $l_1 = l_2 = 0$, *where* $(d^B(m), l^B(m)) = (F, 1)$ *if* $m < \frac{1}{2}q^2 F$, *and equals the solution for* (d, l_0) *of (3) and (4) otherwise. That is, a successful entrepreneur outlays* $d^B(m)$, *and the (book) rate of return of bank loans is* $d^B(m) - 1$; *the bank is audited only in state 0, with probability* $l^B(m)$; *and X obtains net rent* $V^B = \max(0, q^2 \cdot \frac{2+C^B}{2q-q^2} - 2m)$, *where* C^B *denotes the expected auditing costs under FI.* $d^B(\cdot)$ *increases and* $l^B(\cdot)$ *decreases with* m, *and both are independent of* R.

3.4 Conglomeration

Under this mode, the liability is taken by a conglomerate, where each project becomes a division run by the entrepreneur, and monitored by X. Here funds could flow, differently from under FI, directly between investors and the entrepreneurs. However, this difference does not matter at all. To highlight this point, I suppose that under Conglomeration X also intermediates with the flow of funds, as she did under FI; she might thus be regarded as the chief financial officer of the conglomerate sitting in the headquarters. Then, Conglomeration and FI differ only in the allocation of liability, and in all the other respects are the same.

In this section, the marks of some equations are suffixed with "h" to indicate that they are peculiar to the mode of Conglomeration, where monitoring is provided by the "h"eadquarters.

Under Conglomeration, a mechanism is $\{d_s, D_s; l_s\}_{s=0,1,2}$: In (reported) state s, a successful entrepreneur contributes d_s to X and she then passes D_s to the investors on behalf of the conglomerate; the investors audit *each* project with probability l_s for $s = 0$ or 2, and audit the *reportedly failed* project with probability l_1 for $s = 1$.[20] As under FI, the limited liability constraint is still (LL). And as under FI, the two kinds of collusion problems give rise to two groups of IC constraints; again, only the binding ones are presented in the subsection, and the non-binding ones put in Appendix B.

Consider first the grand collusion proof constraints, which command that the conglomerate outlays no more to the investors if reporting the true state than lying about it. The IC for the conglomerate not to mis-report state 2 as state 1 is:

$$D_2 \leq l_1 \cdot 2R + (1 - l_1) \cdot D_1 \tag{G21h}$$

In state 2, reporting the truth leads the conglomerate to outlays D_2, whereas the RHS gives what it expects to outlay if mis-reporting the state as state 1: Then, with probability l_1, the

[20] If the report is $\{s, f\}$ (project 1 succeeds and 2 fails), the truth is either $\{s, f\}$ or $\{s, s\}$. If the truth is $\{f, f\}$ or $\{f, s\}$, there is no incentive to mis-report it as $\{s, f\}$. Thus, only the reportedly failed project needs to be audited in state 1.

reportedly failed project is audited, the fraud uncovered, causing the conglomerate to lose the whole collateral (namely the two projects), worth $2R$ in state 2; otherwise it outlays D_1, according to the reported state.

The IC for the conglomerate not to mis-report state 2 as state 0, (G20h), is not binding.

The IC for the conglomerate not to mis-report state 1 as state 0 is:

$$D_1 \leq l_0 R \tag{G10h}$$

The IC for the conglomerate not to mis-report state 1 as state 2, (G12h), is never binding, since $D_2 \geq D_1$ by Assumption 3, as under FI. Thus $l_2 = 0$.

Move on to consider partial collusion, in which X may form a subcoalition with an entrepreneur (for example E_1) to exploit the other (for example E_2) who is successful. The IC constraint for X not to arrange partial collusion mis-reporting state 2 as state 1 is:

$$d_2 + R - D_2 \geq (1 - l_1)(d_1 + R - D_1) \tag{P2h}$$

In state 2, the X-E_1 subcoalition gets $d_2 + R - D_2$ without collusion: X get d_2 from E_2, E_1 contributes R, and X outlays D_2 to the investors. The RHS gives what the subcoalition expects to get if X declares that only E_2 succeeds (so the economy is in state 1): With probability l_1, the fraud is uncovered and the subcoalition (and E_2 as well) gets nothing; otherwise, X get d_1 from E_2, E_1 contributes R, and D_1 is laid out.

(P2h) can also be derived in a way parallel to that leading to (P2b). In state 2, X obtains $2d_2 - D_2$ if honoring the contracts. But she could arrange partial collusion in which she collects $t < d_2$ from E_1 and declares that E_2 only has succeeded and should thus shoulder the whole liability of the conglomerate. With the collusion, she obtains $(t + d_1 - D_1)(1 - l_1)$, as under FI. The difference is that under FI, E_1 always got $R - t$ with the collusion, even if the fraud was uncovered, since his account was not a part of the collateral, thus beyond appropriation by the investors, whereas under Conglomeration, it is under appropriation, and hence only with the probability of $1 - l_1$ does E_1 obtain $(R - t)$ with the collusion. He obtains $R - d_2$ without it.

Thus, he is willing to partake it only if t satisfies $(1-l_1)(R-t) > R-d_2$. Therefore, if and only if for any such t, $2d_2 - D_2 \geq (1-l_1)(t+d_1-D_1)$, which gives rise to (P2h), X has no incentive to arrange partial collusion in state 2.

The IC constraint for X not to arrange part collusion in state 1, (P1h), is not binding.

As under FI, the investors do not care how the internal contract ($\{d_s\}_{s=0,1,2}$) is arranged, and are only concerned with the grand collusion proof constraints ((G21h), (G10h), and (G20h)). Therefore, facing the security $\Psi \equiv \{D_s\}_{s=0,1,2}$, they commit to the auditing strategy with the minimum $\{l_s\}_{s=0,1,2}$ subject to the three constraints. Let the strategy be $\{l_s^H(\Psi)\}_{s=0,1,2}$. We saw $l_2^H(\Psi) = 0$, and the binding (G21h) leads to $l_1 = \frac{D_2-D_1}{2R-D_1}$.

I move on to the IR constraints for the investors and X. The investors audit two projects in state 0 and one in state 1. The IR constraint for them is

$$2 + [(1-q)^2 \cdot 2l_0 + 2q(1-q)l_1]C \leq q^2 D_2 + 2q(1-q)D_1 \qquad \text{(IR-Ih)}$$

The IR to X is (IR-X), the same as that under FI, since she still obtains the difference between the paid-in from the entrepreneurs and the paid-out to the investors.

The entrepreneurs' problem is then:

Problem 2 $\min_{\{d_s;D_s,l_s\}_{s=0,1,2}} 2q(1-q)d_1 + q^2 \cdot 2d_2$, *subject to*

(1): $l_s = l_s^H(\Psi)$ for $s = 0,1,2$, where $\Psi = \{D_s\}_{s=0,1,2}$.

(2): (LL), (P2h), (P1h), (IR-Ih), (IR-X), and $D_1 \leq D_2$ (Assumption 3).

The solution depends on whether (IR-X) is binding or not. We start with the case of it nonbinding. Let $H \equiv \frac{2R}{q^2R+2(1-q)S}$, $l^H \equiv \frac{2}{q^2R+2(1-q)S} = \frac{H}{R}$, and V^H denote the net rent to X.

Lemma 5 *If (IR-X) is not binding, then the optimal mechanism under Conglomeration is: $D_0 = d_0 = 0$, $D_1 = D_2 = d_1 = d_2 = H$, $l_0 = l^H$ and $l_1 = l_2 = 0$; and X expects to obtain gross rent q^2H. Therefore, it is nonbinding if and only if $m < \frac{1}{2}q^2H$, and in that case, $V^H = q^2H - 2m$.*

Proof. See Appendix B. ■

As was under FI, the *shape* of the mechanism ($D_1 = D_2 \equiv D; d_1 = d_2 \equiv d$) is determined by $l_1 = 0$;[21] that is driven by the trade-off between the auditing costs and the rent to X; and as the auditing costs are large enough, the optimal mechanism triggers auditing only in state 0 ($l_1 = 0$), but gives X net rent $q^2H - 2m$. Moreover, as under FI, the level of D and d are determined by (IR-Ih) and (IR-X) respectively: D is derived by substituting $l_0 = \frac{D}{R}$ (derived from the binding (G10h)) into the binding (IR-Ih), d by noting that (IR-X) is not binding so that d is pushed down until (LL) is binding, namely $d = D$.

The difference from FI is that here $l_0 = \frac{2}{q^2R+2(1-q)S}$ decreases with R, but its counterpart under FI, as stated in Lemma 3, was equal to 1 and independent of R. This difference arises because the collateral is the pool of the projects under Conglomeration and the bank asset under FI. The value of the projects increases with R, but that of the bank asset is independent of R. The probability of auditing is inversely related to the value of the collateral, and therefore decreasing with R under Conglomeration, but independent of R under FI.

Lemma 6 *If $m \geq \frac{1}{2}q^2H$, (IR-X) is binding, and, compared to the case of (IR-X) nonbinding, the only change with the optimal mechanism is that now $d_1 = d_2 = H + (\frac{m}{q} - \frac{qH}{2})$.*

Proof. As the concern about the rent to X disappears, the trade-off between the rent and the auditing costs, as in the parallel case under FI, drives $l_1 = 0$, to minimize the auditing costs. It follows, as in the nonbinding case, that $D_1 = D_2 \equiv D; d_1 = d_2 \equiv d$. D and l_0 are unchanged: They are decided by the simultaneous equations of binding (G10h) and (IR-Ih), both unrelated to m. However, d is now pinned down by the binding (IR-X) (instead of (LL)): $d = H+(\frac{m}{q}-\frac{qH}{2})$; the increment over H, $\frac{m}{q} - \frac{qH}{2}$, covers the part of the monitoring costs beyond the ex post rent, q^2H. ■

Note that for this case, l_0 is independent of m here, whereas its counterpart under FI strictly decreases with m. This difference is again derived from the inverse relationship between the value

[21] As under FI, $l_1 = 0$ implies $D_1 = D_2$ by the binding (G21h), and then $d_1 = d_2$ by the binding (P2h).

of the collateral and the probability of auditing. When (IR-X) is binding, the payment from a successful entrepreneur to X increases with m in order to satisfy her IR constraint, which raises the value of the collateral for investors under FI (namely the bank asset), but affects nothing of the collateral under Conglomeration (namely the projects).

Put the two cases together, $d = H + \max(0, \frac{m}{q} - \frac{qH}{2})$, $V^H \equiv \max(0, q^2H - 2m)$, and the conglomerate finances the two units of investment capital by issuing debt with face value H, which gives rise to book return rate of $\frac{H}{2}-1$. Let $C^H \equiv 2(1-q)^2Cl^H$. Then, $V^H = \max(0, q^2 \cdot \frac{2+C^H}{2q-q^2} - 2m)$, because by binding (IR-Ih), $H = \frac{2+C^H}{2q-q^2}$. To summarize.

Proposition 2 *Let* $H \equiv \frac{2R}{q^2R+2(1-q)S}$ *and* $l^H \equiv \frac{2}{q^2R+2(1-q)S}$. *The optimal mechanism under Conglomeration is:* $D_0 = d_0 = 0, D_1 = D_2 = H, d_1 = d_2 = H + \max(0, \frac{m}{q} - \frac{qH}{2})$, $l_0 = l^H$, *and* $l_1 = l_2 = 0$. *That is, the conglomerate is financed by issuing debt, of which the (book) rate of return is* $\frac{H}{2} - 1$; *a successful entrepreneur outlays* $H + \max(0, \frac{m}{q} - \frac{qH}{2})$; *the conglomerate is audited only in state 0, with each project audited with probability* l^H; *and X obtains net rent* $V^H = \max(0, q^2\frac{2+C^H}{2q-q^2} - 2m)$, *where* C^H *denotes the expected auditing costs under Conglomeration. Moreover, both* H *and* l^H *are decreasing with* R *and independent of* m.

3.5 Comparisons and the Equilibrium Mode

So far I have not examined mode 5 (the Mixed Mode) yet, which I will show soon is dominated by Conglomeration. For the present, let me compare between the four modes which I have examined: IF, Joint Liability without Monitoring, FI, and Conglomeration. First, by Lemma 2, the first two are equivalent. Second, by Assumption 3, $m < \frac{(1-q)q^2RC}{(q^2R+2(1-q)S)S}$, which implies $H + \max(0, \frac{m}{q} - \frac{qH}{2}) < \frac{R}{S}$, that is, $d^H < d^I$. Then, Conglomeration dominates IF. Finally, I am thus left to compare Conglomeration with FI only. FI strictly dominates Conglomeration if and only if $d^B < d^H$.

Lemma 7 $d^B < d^H$ *if and only if* $C^B < C^H$

Proof. $2qd^K = (2 + C^K) + (2m + V^K)$ for $K = H$ and B, which is an accounting equation: The total outlays of the entrepreneurs are decomposed as the repayment to the investors plus that to X: The former covers the investment costs (2) and the auditing costs (C^K), the latter covering the monitoring costs ($2m$) and the net rent to X (V^K). It follows that $d^B - d^H = V^B - V^H + C^B - C^H$. By Propositions 1 and 2, $V^K = \max(0, q^2 \frac{2+C^K}{2q-q^2} - 2m)$, which implies that $V^B - V^H$ has the same sign as $C^B - C^H$. Therefore, $d^B - d^H$ has the same sign as $C^B - C^H$. ■

Therefore, the choice between FI and Conglomeration, by the lemma, follows the social efficiency, although it is not what concerns the entrepreneurs.

As $C^B = (1-q)^2 C_2 l^B(m)$ and $C^H = 2(1-q)^2 C l^H$, $C^B < C^H$ if and only if $\frac{l^B(m)}{l^H} \le \frac{2C}{C_2}$. As $l^H = \frac{2}{q^2 R + 2(1-q)S}$, the inequality is equivalent to:

$$\frac{2C}{C_2} > \frac{l^B(m)(q^2 R + 2(1-q)S)}{2} \qquad \text{(FI-C)}$$

Lemma 8 *If $\frac{2C}{C_2} \le 1$, (FI-C) can never hold true. That is, if FI does not have any number advantage, it can never strictly dominate Conglomeration.*

Proof. See Appendix B. ■

In fact, for FI to dominate Conglomeration, its number advantage, measured by $\frac{2C}{C_2}$, must beyond threshold $T(m, R) \equiv \frac{l^B(m)(q^2R+2(1-q)S)}{2}$. The threshold measures the *Collateral Advantage* of Conglomeration over FI.

$\frac{\partial T}{\partial R} > 0$: The larger is R, the higher is the value of the projects, and hence the bigger the relative collateral advantage of Conglomeration.

$\frac{\partial T}{\partial m} < 0$, as $\frac{dl^B(m)}{dm} < 0$ by Proposition 1: The larger is m, the more is paid to X (the bank) in order to satisfy her IR constraint, the higher is the value of the bank asset, and hence the smaller is the relative collateral advantage of Conglomeration.

Now consider the Mixed Mode. It is dominated by Conglomeration, by the following intuitive argument. The collateral for investors under the Mixed Mode, composed of one project plus the intermediary asset within the other, is worth less than the pool of the two projects. Moreover, in case of default, the investors still have to audit two assets. Therefore, the Mixed Mode, compared to Conglomeration, has neither the collateral advantage nor the number advantage and is dominated by Conglomeration.

Therefore, the real race is between FI and Conglomeration. The equilibrium mode is FI if and only if (FI-C) holds true. To summarize:

Proposition 3 *Conglomeration also materializes the benefit of diversification, as FI. The equilibrium mode is either FI or Conglomeration. It is FI, only if FI has certain number advantage, and if and only if this number advantage, measured by* $\frac{2C}{C_2}$*, is beyond threshold* $T(m,R) \equiv \frac{l^B(m)(q^2R+2(1-q)S)}{2}$*. Moreover,* $\frac{\partial T}{\partial R} > 0$ *and* $\frac{\partial T}{\partial m} < 0$.

The comparative static result that $\frac{\partial T}{\partial m} < 0$ give rise to the following two predictions, which are on the contrary to what would be expected from cost or price considerations.

First, if the monitoring cost is directly measurable, the comparative static predicts:

Empirical Prediction A: The prevalence of bank financing increases with the cost of monitoring service provided by banks.

Most probably, the monitoring cost cannot be directly measured, but only be proxied. It could be proxied by the spread of the average return rate of bank loans over the average market rate of direct finance. In the paper, the former is $d^B(m)-1$ by Proposition 1, the latter is $\frac{H}{2}-1$ by proposition 2, and thus the spread is $d^B(m)-\frac{H}{2}$, increasing with m. When m is proxied by the spread, the comparative static predicts the following cross industry comparison:

Empirical Prediction B: The higher the spread between the average rate of bank loans and the average rate of publicly placed debt, the more prevalent of bank financing.

4 Extension: The Case of a Large Number of Entrepreneurs

In this section I compare FI with Conglomeration for the case of a large number of entrepreneurs. This exercise serves two purposes, to check the robustness of Proposition 3, and to compare the paper more closely with Diamond (1984).

Diamond (1984) shows that the incentive costs of FI approach 0 if the number of entrepreneurs go to infinity. Intuitively, in the same way do large numbers benefit Conglomeration. In the two-entrepreneur case, under both modes, there are two costs associated with the benefit of diversification: In state 0, auditing still happens since the liability to investors is not repaid, and in state 2, X obtains rent, since she is then paid by two successful entrepreneurs while one payment suffices to clear her liability to the investors. Both costs are worn away when the number of entrepreneurs goes to infinity: By the law of large numbers (LLN), the average number of successes is almost fixed, and hence both the probability of the liability not being fully repaid and the probability of X obtaining rent go to 0. This argument is equally applicable to both FI and Conglomeration.

Below a strict analysis is presented, which confirms the intuition above, and the two modes are compared in the speed of convergence.

First, re-state briefly the model with respect to the case of N entrepreneurs. Each of them has an identical and independent project. X, the provider of monitoring service, is accommodated as either the bank that now finances the N projects, or the headquarters of the conglomerate that consists of the N divisions. The costs of auditing the bank is denoted by C_N. We assume $C_N = zN^\alpha$ for some $z > 0$ and $\alpha \leq 1$. The economy has $N+1$ states, state $s = 0, 1...N$, defined by the number of successful projects and occurring with probability $p_N^s = C_N^s q^s (1-q)^{N-s}$.

Passing on to examine FI and Conglomeration in order, let me remark that what happens under IF is independent of the number of entrepreneurs: A successful entrepreneur still outlays

$d^I = \frac{R}{S}$, as in the case of $N = 2$.

4.1 FI for the Case of a Large Number of Entrepreneurs

Consider FI equipped with a mechanism of the following kind. $d_s = d$, $D_s = \begin{Bmatrix} sd, \text{for } s < k \\ kd, \text{for } s \geq k \end{Bmatrix}$, and $l_s = \begin{Bmatrix} 1, \text{for } s < k \\ 0, \text{for } s \geq k \end{Bmatrix}$, for some k. That is, a successful entrepreneur pays the bank with d, independent of s (the overall state), and the bank passes what it receives to the investors up to kd and is audited certainly in case of default, namely failure to fully repay kd. The liability contract is thus debt, with total face value $F = kd$. The mechanism that was considered by Diamond (1984) is of this kind. The mechanisms of the kind are certainly proof against grand collusion. They are also proof against partial collusion, because the outlay of a successful entrepreneur is independent of the overall state, the report of which X can manipulate.

The optimal mechanism of FI is of this kind. First, as we will show soon, X gets no net rent when N is large enough. Therefore, the optimal mechanism shall minimize the expected auditing costs, which implies that there exists some k such that $l_s = 0$ for $s \geq k$. Second, $l_s = 1$ for $s < k$, because the collateral under FI is the bank asset: Thus if the entrepreneurs want to lower l_s in these states, they have to fill more into the bank asset, which they dislike. The arguments so far pin down the arrangement on the liability side of the bank. Lastly, on the asset side, $d_s = d$ in order to be partial collusion proof.

The lemma below bears on the bindingness of (IR-X) (the IR constraint to X).

Lemma 9 *Let $C_N = zN^\alpha$. If (IR-X) is not binding, the gross rent to X under FI is* $V_N = \begin{Bmatrix} o(1) & & \alpha < 0.5 \\ O(\sqrt{N}) & \text{if} & \alpha = 0.5 \\ O(\sqrt{(2\alpha - 1)N \log N}) & & \alpha > 0.5 \end{Bmatrix}$ *with the optimal mechanism.*[22]

Proof. See Appendix B. ■

[22] Hereinafter, the notation $y \approx x$, sometimes denoted as $y = x + o$, means $\frac{y-x}{x} \to 0$ if $x \neq 0$, and $y \to 0$ if $x = 0$. The notation $y = O(x)$ means $y \approx \lambda x$ for some $\lambda > 0$, and $y = o(x)$ means $\frac{y}{x} \to 0$.

The gross rent for the bank is thus not sufficient to cover the total monitoring costs, Nm, for any given $m > 0$, when N is large enough. Therefore, (IR-X) is always binding. It follows that each entrepreneur's expected outlay covers exactly the investment cost (1), the monitoring costs (m), and the average auditing costs. That is,

$$qd = 1 + m + \frac{C_B}{N} \tag{5}$$

Here $C^B = P_d C_N$ is the expected auditing costs under FI, with P_d being the probability of default, namely, $P_d = \Pr(s < k)$.

Suppose P_d goes to 0 (to be verified soon). This, together with $C_N \leq zN$, implies that $\frac{C^B}{N} \approx 0$, and hence by (5), $d \approx \frac{1+m}{q}$. The IR constraint to the investors is $(1 - P_d)F + P_d Q = N + P_d C_N$, where Q is the expected repayment conditional on default, so $Q < F$. As $P_d \to 0$, $F \approx N$. In this mechanism, $F = kd$. It follows that $k = \frac{F}{d} \approx \frac{N}{d} \approx \frac{qN}{1+m}$. Then, by the Central Limit Theorem (CLT), the default probability $P_d = \sum_{s \leq k-1} p_N^s \approx \Phi(\frac{k-Nq}{\sqrt{Nq(1-q)}}) \approx \Phi(\frac{-m}{1+m}\sqrt{\frac{qN}{1-q}})$, which indeed approaches 0 quickly. Hence indeed $d \approx \frac{1+m}{q}$, which is the core result of Diamond (1986), namely that on average diversification wipes out the incentive costs and leaves only technical costs (the investment costs and the monitoring costs). On this ground, FI dominates IF: $\frac{1+m}{q} < \frac{R}{S}$ by Assumption 3.

The auditing costs under FI is $C^B \approx \Phi(\frac{-m}{1+m}\sqrt{\frac{qN}{1-q}})C_N$.

4.2 Conglomeration for the case of a Large Number of Entrepreneurs

Move on to examine Conglomeration. Given a mechanism $\{d_s, D_s; l_s\}_{s=0,1,2,\ldots,N}$, in state s, the conglomerate outlays D_s if honoring the contracts. If it lies that the state is $t < s$, then each of the $s - t$ actually successful but reportedly failed projects is audited with probability l_t. If none of them is audited, which happens with probability $(1 - l_t)^{s-t}$, the lie is not uncovered and the conglomerate outlays D_t according to the reported state, t; if any one of them is audited, the lie is uncovered, and then all the projects, worth sR in the state, are appropriated. Therefore, the according grand collusion proof constraint is

$$(1-(1-l_t)^{s-t})\cdot sR+(1-l_t)^{s-t}D_t\geq D_s \tag{Gsth}$$

Note that even if l_t is small, $(1-(1-l_t)^{s-t})$ could still approach 1 if $s-t$ is large, which represents another advantage of Conglomeration for the case of a large N, called the "*Spread Advantage*": Since the collateral of Conglomeration spreads across the N projects, a fraud of lying about the outcomes of many projects is vulnerable to be detected. The larger the number of lies involved, the bigger the spread advantage.

I consider the following kind of mechanisms, which make the most use of the spread advantage. $d_s=d$, $D_s=\begin{Bmatrix}0,\text{for } s<k\\ kd,\text{for } s\geq k\end{Bmatrix}$, and $l_s=\begin{Bmatrix}l_s,\text{for } s<k\\ 0,\text{for } s\geq k\end{Bmatrix}$, for some k. The arrangement between the entrepreneurs and X is the same as that under FI and is proof against partial collusion. For states $t\geq k$, the arrangement to the investors is also the same as that under FI. However, for states $t<k$, to utilize the spread advantage to the max, D_t is set to 0, so that for any $t<k$ (Gkth) is binding and moreover l_t is minimized: By the binding (Gkth), $l_t=1-(\frac{kR-D_k}{kR-D_t})^{\frac{1}{k-t}}$, decreasing with D_t.

Whether the optimal mechanism under Conglomeration is of the above kind is an open question. But even if it is not, FI has to dominate Conglomeration with a mechanism of the kind if it can dominate Conglomeration equipped with the optimal mechanism at all. By considering mechanisms of that kind, therefore, I can at least draw a necessary condition for FI to be viable.

The optimal k, as will be shown soon, is pinned down by the binding IR constraint to the investors. For any $t<k$, the binding (Gkth) implies $l_t=1-(1-\frac{d}{R})^{\frac{1}{k-t}}$. These l_t will be difficult to handle. But note that $\frac{d}{R}\frac{1}{k-t}\leq 1-(1-\frac{d}{R})^{\frac{1}{k-t}}\leq\log(1-\frac{d}{R})^{-1}\frac{1}{k-t}$.[23] As I am interested in finding a necessary condition for FI to dominate Conglomeration, I take $l_t=\log(1-\frac{d}{R})^{-1}\frac{1}{k-t}$. As happened under FI and will be shown, $d\to\frac{1+m}{q}$. Thus $\log(1-\frac{d}{R})^{-1}\to\log(1-\frac{1+m}{qR})^{-1}\equiv\gamma>0$.

[23] That is because $(1-x)\mu\leq 1-x^{\mu}\leq-\mu\log x$ for $0<x,\mu\leq 1$. For the former inequality, let $f(x)=1-x^{\mu}-(1-x)\mu$. $f(1)=0$, and $f'(x)=-\mu(x^{\mu-1}-1)<0$ for $x,\mu<1$. Therefore, $f(x)>0$ if $x<1$. For the latter, let $\beta=x^{\mu}$ and $g(\beta)=-\log\beta-(1-\beta)$. $g(1)=0$, and $g'=1-\frac{1}{\beta}<0$ for $\beta<1$. Therefore, $g(\beta)>0$ if $\beta<1$.

In any state $t < k$, $N - t$ projects fail, each audited with probability $l_t = \frac{\gamma}{k-t}$. Therefore, the expected auditing costs is $C^H = C\sum_{t\leq k-1} p_N^t (N-t)\frac{\gamma}{k-t} < \gamma NCP_d$, where $P_d \equiv \sum_{t\leq k-1} p_N^t$ is the probability of default.

As was under FI, (IR-X) is binding.[24] Suppose P_d and hence the average auditing costs ($\frac{C^H}{N}$) go to 0 (to be verified). Then, by the same arguments as those under FI, $d \approx \frac{1+m}{q}$ and $k \approx \frac{N}{d} \approx \frac{qN}{1+m}$. Then, as under FI, indeed $P_d \approx \Phi(\frac{-m}{1+m}\sqrt{\frac{qN}{1-q}})$, and goes to 0 quickly, and so does $\frac{C^H}{N} < \gamma CP_d$. As $d \approx \frac{1+m}{q}$, Conglomeration obtains the same benefit of full diversification, which confirms the intuition explained at the beginning of subsection.

Note that $C^H < \gamma NC\sum_{s\leq k-1} p_N^s \frac{1}{k-s} = \gamma NCP_d \cdot E_s(\frac{1}{k-s}|s \leq k-1) \equiv \overline{C^H}$. As for the conditional expectation, we have:

Lemma 10 *$E_s(\frac{1}{k-s}|s \leq k-1) \to \frac{m}{1-q}\log\frac{1-q+m}{m}$, when $N \to \infty$ and $\frac{k}{N} \to \frac{q}{1+m}$.*

Proof. See Appendix B. ■

4.3 The Comparison between the two Modes for the case of a Large Number of Entrepreneurs

FI could dominating Conglomeration only if C^B is no bigger than $\overline{C^H}$, or equivalently, $C_N \leq \gamma NC \cdot \frac{m}{1-q}\log\frac{1-q+m}{m}$ by Lemma 7. Substitute $\gamma(m,R) = \log(1-\frac{1+m}{qR})^{-1}$ and do rearrangement. The inequality is equivalent to

$$\frac{NC}{C_N} \geq \log(1-\frac{1+m}{qR})[\frac{m}{1-q}\log\frac{1-q+m}{m}]^{-1} \quad \text{(FI-C-N)}$$

That is, FI dominates Conglomeration and arises in equilibrium, only if the Number Advantage is beyond the threshold $T_L(m,R) = \log(1-\frac{1+m}{qR})[\frac{m}{1-q}\log\frac{1-q+m}{m}]^{-1}$. $\frac{\partial T_L}{\partial R} > 0$ and $\frac{\partial T_L}{\partial m} < 0$.[25]

[24] Actually, the gross rent is of order $\sqrt{N\log\log N}$, if (IR-X) is not binding. The proof is available upon request.

[25] $\log(1-\frac{1+m}{qR})$ decreases with m, and $m\log\frac{1-q+m}{m}$ increases with it: $\{m\log\frac{1-q+m}{m}\}' = \log\frac{1-q+m}{m} - \frac{1-q}{1-q+m} \equiv \log(1+x) - \frac{x}{1+x} > 0$, where $x = \frac{1-q}{m}$. For the last inequality, let $f(x) \equiv \log(1+x) - \frac{x}{1+x}$; then $f(0) = 0$ and $f' = \frac{1}{1+x} - \frac{1}{(1+x)^2} > 0$ for $x > 0$.

The two signs confirm the robustness of the comparative statics of the two-entrepreneur case as was summarized by Proposition 3. Moreover, note that $\lim_{m \to 0} T_L = \infty$, which means that FI has no chance to dominate Conglomeration when m is very small, provided that the FI's number advantage is not incredibly big, in the sense that the costs of auditing the bank increase linearly with its asset scale ($C_N = zN$), though the slope z could be tiny.[26] To sum up,

Proposition 4 *For the case of a large number of entrepreneurs, both Conglomeration and FI obtain the benefit of diversification; FI dominates Conglomeration only if its number advantage is bigger than the threshold on the RHS of (FI-C-N), which increases with R and decreases with m; and the chance of FI dominating Conglomeration goes to 0 when m approaches 0, provided the costs of auditing the bank increase linearly with its asset scale ($C_N = zN$).*

5 Conclusion

The paper, based on the allocation of liability, compares all the possible modes of organizing finance and delegated monitoring service on a level playing field. The paper finds that a mode of direct finance, *Conglomeration*, also accommodates monitoring, and thereby also materializes the benefit of diversification. Therefore, the materialization of the benefit does not necessarily make *the mode of financial intermediation* (FI) viable, as the literature of banking following Diamond (1984) claims. Actually, the paper finds that FI has any chance of dominating Conglomeration, thus to be viable, only if it has the number advantage, that is, banking technologies generate certain savings of auditing costs. Moreover, the paper finds that the cost of monitoring favors FI, as opposite to what that literature predicts.

Moreover, the exercise of the paper provides a new angel to consider the regulation of banks.

[26] However, it can be shown that when $m = 0$, FI dominates Conglomeration only if C_N is in the order of at most $\sqrt{N}$. The proof is available upon request. $m = 0$ is a very special case in which (IR-X) is never binding and then the optimal k is decided by the trade-off between the auditing costs and the rent to Ms X, as for the two-entrepreneur case.

One core message of the paper is that we should not take for granted that it has to be banks that provide certain services; on the contrary, we should realize that banking is only one particular mode in which the atoms of human capital (for services), physical capital, and productive projects are structured into molecules and that these atoms could be structured into other kinds of molecules. If the mode of banking suffers some problems (which are beyond this paper), we should discourage this particular mode and to encourage alternative modes by some measures, for example, a specific levy on banks. The effects of this levy on the choice of the equilibrium mode, it seems, has not been even noticed yet.

Appendix A: Discussion of Assumptions

This appendix checks the robustness of Proposition 3 with respect to the three assumptions, namely, those of stochastic auditing (Assumption 1), of the auditing costs being high ($2(1-q)C \geq qR$) (Assumption 2), and of monotonic securities (Assumption 3) The proposition basically makes four points.

A: Conglomeration has a positive chance of dominating IF and is thus an economically sensible mode to materialize the benefit of diversification.

B: Not because of this benefit, but because of the number advantage, is FI viable.

C: The chance of FI arising in equilibrium decreases with R.

D: This chance increases with m.

Let us check the robustness of each of the four points with respect to alternation of the two assumptions.

What if Only Deterministic Auditing is Allowed?

We saw in Subsection 3.2 that if only deterministic auditing is allowed, Joint Liability without Monitoring (Mode 2) obtains the benefit of diversification without resorting to monitoring if $(1-q)S \geq 1$; this mode then dominates all the other modes, and the whole Proposition 3 collapses. In the following, I consider the case where $(1-q)S < 1$ and hence mode 2 keeps equivalent to IF.

Under IF, there is only one option of mechanism: Auditing has to be exercised when failure is reported. The face value of debt, d, namely the entrepreneur's outlay in case of success, is determined by the binding IR constraint to investors: $qd = 1 + (1-q)C$. So $d^I = \frac{1+(1-q)C}{q}$.

Under Conglomeration, there are two options of mechanism: Auditing is exercised in state 0 only or in state1 as well. The latter cannot be optimal if Conglomeration has a chance of dominating IF: It replicates under Conglomeration the outcome of IF, as it renders a failed project always audited, but it engenders additionally the monitoring costs. For the purpose of comparing between the modes, below I only consider the mechanism with $l_1 = 0$, which implies that $D_1 = D_2 \equiv D$. D, the face value of the debt, is decided by the binding (IR-Ih): $(2q-q^2)D = 2+2(1-q)^2C$, which gives rise to $D = H' \equiv \frac{2+2(1-q)^2C}{2q-q^2}$. When $2m \leq q^2H'$, (IR-X) is not binding, and the outlay of a successful entrepreneur equals H'. When $2m > q^2H'$, the outlay increases to satisfy the binding (IR-X): $2q(1-q)(d-H') + q^2(2d-H') = 2m$, which gives $d = H' + \frac{m}{q} - \frac{qH'}{2}$. Overall, a successful entrepreneur outlay $d^H = H' + \max(\frac{m}{q} - \frac{qH'}{2}, 0)$.

Under FI, when (IR-X) is not binding, the alteration to deterministic auditing does not affect the optimal mechanism of FI, which, by Lemma 3, already has $l_0 = 1$. Then, a successful entrepreneur outlays $B \equiv \frac{2+(1-q)^2C_2}{2q-q^2}$, by the lemma. When (IR-X) is binding, similar to the case of Conglomeration, the outlay is increased to satisfy the constraint. Overall, $d^B = B + \max(\frac{m}{q} - \frac{qB}{2}, 0)$.

Note that $d^H \leq d^I \Leftrightarrow m \leq (1-q)qC$, which is tighter than the bound set in Assumption 3; that is, Conglomeration still has a positive, though smaller, chance to dominate IF. $d^B \leq d^H \Leftrightarrow B \leq H' \Leftrightarrow C_2 \leq 2C$. Therefore, FI dominates Conglomeration so long as FI has the number advantage, however small, and the chance of FI arising in equilibrium is independent of R and m.

To summarize, if only deterministic auditing is allowed, points A and B of Proposition 3 keep holding true, but the comparative statics of points C and D do not. The latter two points are driven by the trade-off between the number advantage of FI and the collateral advantage of Conglomeration, but the latter is assumed away, however, by the assumption of deterministic auditing: The assumption vanishes the inverse relationship between the value of the collateral and the probability of auditing, as now the probability is always 1, whenever auditing is needed.

What if $2(1-q)C < qR$?

If $2(1-q)C < qR$, Conglomeration is dominated by IF even when $m = 0$, though it still dominates mode 5, the mixed mode (the argument for the dominance does not depend on this assumption). We saw in the discussion of Lemma 4 that the optimal mechanism of Conglomeration is driven by the trade-off between the auditing costs and the rent to X. The balance tilts in the favor of reducing the former at the expense of the latter if the auditing cost is important enough, in the sense that $2(1-q)C \geq qR$. If this inequality does not hold, however, the balance tilts to the opposite direction, in the favor of reducing the rent at the expense of auditing costs (see the proof of Lemma 4). The simplest way of getting rid of the rent is not to use monitoring service at all, which brings us to mode 2, Joint Liability without Monitoring, a mode equivalent to IF. These arguments give an intuition as to why FI dominates Conglomeration if $2(1-q)C < qR$, even when $m = 0$.

For a strict proof of the dominance, imagine each mode as a machine that pumps income to investors at the technical costs (namely the auditing costs as 0 monitoring cost has been assumed) plus the incentive costs (namely X's rent). The efficiency of the two machines, IF and Conglomeration, can be compared in the costs of pumping out a given amount of income to investors.

As for IF, to pump out net income $q \cdot l^I R - (1-q) \cdot l^I C = l^I S$, it incurs technical costs $C^I = (1-q)l^I C$ and 0 incentive costs.

As for Conglomeration, l_1 is minimized by having (G21h) binding, but the minimum l_0 is derived from either binding (G10h) or binding (G20h).[27] The proof for the case of (G10h) binding is given below, the proof for the case of (G20h) binding put in Appendix B. Then given $\{l_0, l_1\}$, $D_1 = l_0 R$ and $D_2 = l_1 \cdot 2R + (1-l_1)D_1 = 2l_1 R + (1-l_1)l_0 R$, and Conglomeration pumps to the investors per project a net income of $N^H = \frac{1}{2}[2q(1-q)D_1 + q^2 D_2] - (1-q)l_e C = l_e S + \frac{1}{2}q^2(1-l_1)l_0 R$, where $l_e = (1-q)l_0 + ql_1$ is the expected probability of a failed project being audited. It incurs the auditing costs of $(1-q)l_e C$. Moreover, X gets a rent of $m_2 = D_2 - 2l_1 R$ in state 2 (see the proof of Lemma 4). Thus per project incentive costs are $\frac{1}{2}q^2 m_2 = \frac{1}{2}q^2(1-l_1)l_0 R$. Therefore, the overall costs of pumping out N^H are $C^H = (1-q)l_e C + \frac{1}{2}q^2(1-l_1)l_0 R$.

If IF pumps out the same amount of income, N^H, namely if $l^I S = N^H$, which gives rise to $l^I = \frac{N^H}{S} = l_e + \frac{1}{2S}q^2(1-l_1)l_0 R$, then it incurs costs $C^I = (1-q)C[l_e + \frac{1}{2S}q^2(1-l_1)l_0 R]$. The costs are smaller than C^H, the costs of Conglomeration, if and only if $(1-q)C[l_e + \frac{1}{2S}q^2(1-l_1)l_0 R] < (1-q)l_e C + \frac{1}{2}q^2(1-l_1)l_0 R \Leftrightarrow \frac{(1-q)C}{S} < 1 \Leftrightarrow (1-q)C < qR - (1-q)C$, which holds true if $2(1-q)C < qR$. Therefore, Conglomeration is dominated by IF if $2(1-q)C < qR$.

With the Mixed Mode dominated by Conglomeration, which is in turn dominated by IF, FI is left to be compared with IF only; that is the scenario in which the approach of Diamond (1984) is applicable. For this comparison, first note that as monitoring is provided not under IF but under FI, the monitoring cost (m) disadvantages FI, opposite to point D.

Second, as for point C, I only consider the case in which $m = 0$. We saw in the discussion following Lemma 3 (or its proof) that the optimal mechanism under FI has $l_1 = 0$ if $C_2(1-q)(3-q) \geq 2$; otherwise $l_1 = 1$. The former gives rise to $d^B = \frac{2+(1-q)^2 C_2}{2q-q^2}$, the latter to $d^B = \frac{2+(1-q^2)C_2}{2q}$. FI dominates IF ($d^B \leq d^I = \frac{R}{S}$), if and only if, for the former case, $\frac{C_2}{2C} \leq \frac{1-q-q^2\frac{R}{2C}}{S(1-q)^2}$, and for the latter case, $\frac{C_2}{2C} \leq \frac{1}{S(1+q)}$.[28] $\frac{1}{S(1+q)} < 1$ certainly; $\frac{1-q-q^2\frac{R}{2C}}{S(1-q)^2} < 1$ as well if $2(1-q)C < qR$:

[27] Subsection 3.4 showed that it is (G10h) that is binding, if $2(1-q)C \geq qR$. But that may not be true if the inequality does not hold.

[28] For the former, $\frac{2+(1-q)^2 C_2}{2q-q^2} \leq \frac{R}{S} \Leftrightarrow 2S + (1-q)^2 S C_2 \leq 2qR - q^2 R \Leftrightarrow (1-q)^2 S C_2 \leq 2(1-q)C - q^2 R = 2C(1-q-q^2\frac{R}{2C})$. For the latter, $\frac{2+(1-q^2)C_2}{2q} \leq \frac{R}{S} \Leftrightarrow 2S + S(1-q^2)C_2 \leq 2qR \Leftrightarrow S(1-q^2)C_2 \leq 2(1-q)C \Leftrightarrow S(1+q)C_2 \leq 2C$.

$\frac{1-q-q^2\frac{R}{2C}}{S(1-q)^2} < \frac{1-q-q^2 \cdot \frac{1-q}{q}}{S(1-q)^2} = \frac{1}{S} \leq 1$. Moreover, both $\frac{1}{S(1+q)}$ and $\frac{1-q-q^2\frac{R}{2C}}{S(1-q)^2}$ decreases with R. Therefore, FI dominates IF and thus arises in equilibrium only if it has the number advantage, and the chance of FI arising in equilibrium decreases with R. This comparative static is because IF's collateral advantage increases with R as $l^I = \frac{1}{S}$ decreases with R.

To summarize, if $2(1-q)C < qR$, point A of Proposition 3 fails to hold and point D goes to the opposite, both because Conglomeration is dominated by IF. However, point B and C keeps holding true. Indeed, the race is now between FI and IF, determined by balance of FI's number advantage versus IF's collateral advantage.

What If Assumption 3 Is Removed?

Assumption 3 in essence commands that $D_2 \geq D_1$ and consequently restricts the feasible set of mechanisms under modes 2, 3, 4, and 5, which all feature joint liability. The equivalence of mode 2 (Joint Liability without Monitoring) to IF, as we saw, does not depend on the assumption and carries on holding true in its removal And as is to be shown, its removal affects nothing of Conglomeration, and hence changes nothing of the mode's dominance over mode 5 (the Mixed Mode) by Conglomeration. Below I set off to examine the effects of its removal on the optimal mechanisms under FI and Conglomeration.

If Assumption 3 is removed, I have to consider mechanisms $\{d_s, D_s; l_s\}_{s=0,1,2}$ in which $D_2 < D_1$. With these mechanisms, the entity liable to investors have no incentive to misreport state 2 as state 1, leading to $l_1 = 0$, but have incentives to misreport state 1 as state 2. Under FI, in state 1, the bank outlays to investors D_1 if telling the truth; if it lies that the state is 2, then with probability l_2 it is audited by the investors and the bank loses the whole asset which is worth d_1 in the state, and with probability $1 - l_2$ the bank is not audited and outlays to the investors D_2 based on its reported state. The grand collusion proof constraint for the bank not to misreport state 1 as state 2 is thus:

$$D_1 \leq l_2 d_1 + (1 - l_2) D_2 \tag{G12b}$$

Similarly, under Conglomeration, in state 1 the conglomerate outlays D_1 if telling the truth. If it lies that the state is 2, then each of the two projects is audited with probability l_2; so with this probability the actual failed project is audited, which uncovers the fraud and causes the conglomerate loses all its projects which is worth R in state 1, and with probability $1 - l_2$ the project is not audited, the fraud uncovered, and the conglomerate outlays D_2. The grand collusion proof constraint for the conglomerate not to misreport state 1 as state 2 is thus:

$$D_1 \leq l_2 R + (1 - l_2) D_2 \tag{G12h}$$

Furthermore, as now l_2 could be positive, the IR for investors changes. Under FI, it becomes:

$$2 + [(1-q)^2 \cdot l_0 + 2q(1-q)l_1 + q^2 \cdot l_2]C_2 \leq q^2 D_2 + 2q(1-q)D_1 \tag{IR-Ib'}$$

And under Conglomeration, it becomes:

$$2 + [(1-q)^2 \cdot 2l_0 + 2q(1-q)l_1 + q^2 \cdot 2l_2]C_2 \leq q^2 D_2 + 2q(1-q)D_1 \tag{IR-Ih'}$$

Note again that the difference between (IR-Ib') and (IR-Ih') is that the coefficients before l_0 and l_2 are 1 in the former and 2 in the latter. This difference captures the number advantage of FI – only one asset needs to be audited under it, whereas two under Conglomeration.

Now given $\Psi \equiv \{d_s, D_s\}_{s=0,1,2}$ offered by the entrepreneurs, the investors commit to a strategy $\{l_s\}_{s=0,1,2}$ that minimize the auditing costs subject to all the grand collusion proof constraints (G10), (G20), (G21), and moreover (G12) as above. Except this change and the change in the IR for investors, the entrepreneurs' problem under FI is the same as Problem 1 of Subsection 3.3, and their problem under Conglomeration is the same as Problem 2 of Subsection 3.4. The solutions to these two problems, as before, depend on whether the IR for the expert is binding or not.

Lemma 11 *Suppose Assumption 3 is removed. If the IR for the expert is not binding, then:*

(a) The optimal mechanism under Conglomeration does not change and is the same as was specified in Lemma 5 of Subsection 3.4.

(b) As for FI, there are two possibilities. Let $J \equiv 2q(1-q)R - [q^2 + (1-q)^2]C_2$. *(i) If* $J \geq 2$ *and* $C_2 \leq C^0 \equiv \frac{2}{4q-2q^2-1}$[29], *or if* $J \leq 2$ *and* $q(1-q)R \geq \frac{2-2q}{2-q} + \frac{(2-q)q^2+(1-q)^3}{2-q}C_2$*, the optimal mechanism is:* $d_1 = \min\{\frac{2+[(1-q)^2+q^2]C_2}{2q(1-q)}, R\}$ *and* $d_2 = \max\{0, \frac{2-J}{q^2}\}$*;* $D_1 = d_1$ *and* $D_2 = d_2$*;* $l_0 = l_2 = 1; l_1 = 0$. *(ii) Otherwise, the optimal mechanism is the same as was specified in Lemma 3.*

If the IR for the expert is binding, then the optimal mechanisms are unchanged by the removal of Assumption 3, namely, that under FI is the same as specified in Lemma 4 and that under Conglomeration is the same as in Lemma 6.

Proof. See Appendix B. ■

To intuitively understand the lemma, let us go back to the intuition as for the optimal mechanism specified in Lemma 3. The mechanism, as expounded after the statement of that lemma, is driven by the trade-off between the auditing costs and the rent to X: The more the auditing exercised, the less the informational advantage X enjoys, and hence the less the informational rent to her. When l_2 is presupposed, essentially by Assumption 3, to be 0, the balance always tilts to reducing the auditing costs, leading to the mechanism in which auditing is exercised only in state 0 and X gets rents. However, Lemma 11 above says that if l_2 could be manipulated, the balance could tilt to the other direction under FI, to exercise auditing more often, now in state 2 ($l_2 = 1$) as well as in state 0, so as to reduce the rent to X, which now equals $q^2 d_2 = q^2 \max\{0, \frac{2-J}{q^2}\}$ and is positive only when the limited liability for entrepreneurs is binding in state 1 ($d_1 = R$).

When the IR for the expert is binding due to a high m, the concern of the rent evaporates and the sole driving force of the optimal mechanism is to reduce the auditing costs. Therefore, auditing is exercised only in state 0 and the optimal mechanism under FI is as specified in Lemma 4 and that under Conglomeration as in Lemma 6, namely, the removal of Assumption 3 has no bite at all.

Note that the mechanism in (bi) could feature that $D_2 = d_2 = 0$ and $D_1 = d_1 > 0$, namely,

[29] If $-2q^2 + 4q - 1 \leq 0 \Leftrightarrow q \leq \frac{2-\sqrt{2}}{2}$, we set $C_v^0 = \infty$ and the inequality hold true always.

the investors and the bank get paid only in state 1, but get nothing in state 2, when both projects succeed. Note that the mechanism would collapse if the investors could collude with X (namely the bank): Investors could agree to receive $d_1 - \varepsilon < d_1$ in either state 1 or 2, which incentivizes the bank to pick d_1 from one entrepreneur in state 2 by declaring that only his project has succeeded. But this kind of collusion has no bite on the mechanism specified in Lemma 3.

Note also that (bi) is the case only if $C_2 \leq C^0 \equiv \frac{2}{4q-2q^2-1}$.[30] On the other hand, $C_2 \geq C$ and by Assumption 2 $C \geq \frac{1}{1-q}$. Therefore, (bi) could be the case only if $\frac{1}{1-q} \leq \frac{2}{4q-2q^2-1} \Rightarrow q \leq \frac{3-\sqrt{3}}{2}$. That is, if $q > \frac{3-\sqrt{3}}{2} \approx 0.63$, which seems more realistically relevant, (bi) can never be the case and Assumption 3 has no bite.

Even where the removal of the assumption has bite, namely, where the IR for X is non-binding and the optimal mechanism under FI is as (bi), the above listed four points A, B, C, and D of Proposition 3 still hold. The removal of the assumption, as we saw, affects nothing of Conglomeration, hence nothing of point A which is concerned with the mode. Points C and D are driven by the trade-off between the number advantage and the collateral advantage, of which nothing is affected by the removal of the assumption. Only point B needs to be checked, namely, FI can dominate Conglomeration only if it has the number advantage, that is, $\frac{C_2}{2C} \leq 1$. To check that, note that when (bi) prevails, the best case for the entrepreneurs is that where the limited liability constraint is not binding, hence $d_2 = 0$ and $d_1 = \frac{2+[(1-q)^2+q^2]C_2}{2q(1-q)}$, by which a successful entrepreneur outlays $qd_2 + (1-q)d_1 = \frac{2+[(1-q)^2+q^2]C_2}{2q}$. FI with the mechanism of (bi) dominates Conglomeration only if this amount of outlay is no bigger than the outlay under Conglomeration, $\frac{2R}{q^2R+2(1-q)S}$ by Lemma 5, that is, only if $\frac{2+[(1-q)^2+q^2]C_2}{2q} \leq \frac{2R}{q^2R+2(1-q)S} \Leftrightarrow \frac{C_2}{2C} \leq \frac{1}{(1-q)^2+q^2} \cdot \frac{2(1-q)^2+q^2\frac{R}{C}}{q^2R+2(1-q)S} \equiv \kappa$. By Assumption 2, $\frac{R}{C} \leq \frac{2(1-q)}{q}$ and $q^2R + 2(1-q)S \geq 2$.[31] Therefore, the upper bound $\kappa \leq \frac{1-q}{(1-q)^2+q^2}$, which is no bigger than 1 when $q \geq 0.5$ and no bigger than $\frac{2+\sqrt{2}}{3} \approx 1.138$ all the time. Hence, FI with the mechanism of (bi) dominates Conglomeration only if it has the number advantage when $q \geq 0.5$, or only if $\frac{C_2}{2C} \leq 1.138$, and point B keeps

[30] In $R - C_2$ plane, line $q(1-q)R = \frac{2-2q}{2-q} + \frac{(2-q)q^2+(1-q)^3}{2-q}C_2$ intersects with line $J = 2$ at $C_2 = C^0$.

[31] Assumption 2 implies that $qR \geq 2$ and $S \geq 1$. Therefore $q^2R + 2(1-q)S = q \cdot qR + (1-q) \cdot 2S \geq 2$.

holding true with Assumption 3 removed, at least qualitatively.

To summarize, if Assumption 3 is removed, nothing is changed to Conglomeration, and nothing is changed to FI if $q > 0.63$, or if the IR for the expert is binding; and even where the removal has bite, points A, C and D of Proposition 3 are not affected and point B keeps holding true at least qualitatively.

Appendix B: The Proofs

The Proof of Lemma 3:

To prove the lemma, it suffices to show that the mechanism given in the lemma is optimal even if $m = 0$, namely, if (IR-X) is never binding. First let us make up the two constraints that are unbinding and missed in the main context. The IC constraint for the bank not to mis-report state 2 as state 0 is

(G20b): $l_0 \cdot 2d_2 \geq D_2$.

And in state 1, if X arranges the partial collusion in which she gives ϵ to the failed entrepreneur to buy his silence and declares state 2, she is not audited ($l_2 = 0$) and obtains $d_2 - D_2 - \epsilon$, which should be no bigger than $d_1 - D_1$ for any $\epsilon > 0$, in order to disincentivize the collusion. The partial collusion proof in state 1 is thus

(P1b): $d_1 - D_1 \geq d_2 - D_2$

Note that (P2b) is binding in the optimization. Otherwise, consider the mechanism $D_1 = D_2 = d_1 = 2d_2; l_0 = 1, l_1 = l_2 = 0$. It implements the benefit of cross subsidization, but gives no rent to X, which is impossible to achieve by Lemma 2. That is, this mechanism is infeasible. But it satisfies all the constraints except (P2b). Thus (P2b) must be binding, and so becomes:

(A1): $2d_2 - D_2 = (1 - l_1)(d_1 + d_2 - D_1)$.

It follows that (P1b) is not binding: $d_1 - D_1 \geq |_{d_1 \geq D_1 \text{ by } (LL)} (1 - l_1)(d_1 - D_1) = |_{(\text{A1})} d_2 - D_2 + l_1 d_2 \geq d_2 - D_2$.

As for the investors' problem of finding the optimal auditing strategy, (G21b) is binding; otherwise, they would lower l_1, which tightens (P2b), but they do not care. The binding (G21b) is

(A2) $D_2 = 2l_1d_2 + (1 - l_1)D_1$.

The entrepreneurs will choose d_1 such that (G10b) is binding; otherwise, they would lower d_1, which loosens (IR-Ib) and (P2b), though it tightens (P1b) which we saw is not binding. The binding (G10b) is

(A3): $D_1 = l_0d_1$.

Lastly, (IR-Ib) is binding:

(A4): $q^2D_2 + 2q(1 - q)D_1 = 2 + C_2[1 - q)^2l_0 + 2q(1 - q)l_1]$.

(G20b) is to be verified not binding. The entrepreneurs' problem is to minimize $d^B = qd_2 + (1 - q)d_1$, s.t. (A1)-(A4).

Substituting into (A1) D_2 and D_1 from (A2) and (A3) respectively, we get a link between d_1 and d_2: $(1-l_1)(d_2-d_1) = 0$. Then, two cases arise, either $l_1 = 1$, or $d_2 = d_1$. They are examined one by one, and the latter first.

If $d_2 = d_1 = d^B$, by (A2) $D_1 = l_0d^B$ and by (A3) $D_2 = (2l_1 + l_0 - l_0l_1)d^B$. Substitute these into (A4), we have $d^B = \frac{1}{q}\frac{2+C_2[(1-q)^2l_0+2q(1-q)l_1]}{2ql_1+(2-q)l_0-ql_0l_1}$. Applying $(\frac{a+bx}{c+dx})' = \frac{bc-ad}{(c+dx)^2}$, we have $\frac{\partial d^B}{\partial l_0} = \frac{1}{q}\frac{-2(2+q+ql_1)+2q(1-q)C_2l_1(ql_1-1)}{[2ql_1+(2-q)l_0-ql_0l_1]^2}$. Obviously $-2(2 + q + ql_1) \leq 0$ and $ql_1 - 1 < 0$. Therefore, $\frac{\partial d^B}{\partial l_0} < 0$ and the optimal $l_0 = 1$. Similarly, $\frac{\partial d^B}{\partial l_1}|_{l_0=1} = \frac{-2+(1-q)C_2(3-q)}{[2ql_1+(2-q)l_0-ql_0l_1]^2} > 0$, where $C_2(1 - q)(3 - q) > 2$ because $C_2 \geq C \geq \frac{1}{1-q}$, the last inequality derived from Assumption 3: $2(1 - q)C - (1 - q)C \geq qR-(1-q)C \geq 1$. Therefore, the optimal $l_1 = 0$. Substitute $l_0 = 1$ and $l_1 = 0$ into the formula of d^B, and we have $d_2 = d_1 = d^B = \frac{2+C_2(1-q)^2}{2q-q^2} \equiv B$, and into (A2) and (A3) we have $D_1 = D_2 = B$.

If $l_1 = 1$, by (A2) $D_2 = 2d_2$. Since $D_2 \leq 2l_0d_2$ ((G20b)), we get $l_0 = 1$. By (A3) $D_1 = d_1$. That is, all the asset paid-in is passed to investors. Then the LHS of (A4) equals $2qd^B$. As $l_0 = l_1 = 1$, the RHS of (A4) equals $2 + C_2[(1 - q)^2 + 2q(1 - q)]$. Therefore, $d^B = \frac{2+C_2(1-q^2)}{2q}$ in this case.

Compare the two cases. $\frac{2+C_2(1-q^2)}{2q} > \frac{2+C_2(1-q)^2}{2q-q^2} \Leftrightarrow C_2(1 - q)(3 - q) > 2$, which has been proved true. Thus, the first case gives the solution of the minimization problem. Therefore, the optimal mechanism, when I ignore (IR-X), is $d_2 = d_1 = D_1 = D_2 = B \equiv \frac{2+C_2(1-q)^2}{2q-q^2}$, $l_0 = 1$ and $l_1 = 0$. With this mechanism, (G20b) is satisfied and X obtains $2d_2 - D_2 = B$ in state 2; therefore, if $q^2 B > 2m$, (IR-X) is indeed unbinding.

Q.E.D.

The Proof of Lemma 5

As was in the proof of Lemma 3, it suffices to show that the optimal solution without taking into account (IR-X) gives X expected gross rent of $q^2 H$; therefore, if $q^2 H > 2m$, (IR-X) is indeed unbinding.

First we make up the neglected constraints. If the conglomerate reports state 0 in state 2, with probability $2l_0 - l_0^2$, one project at least is audited, and then the whole collateral, worth $2R$, is appropriated by the investors (here the implicit off-equilibrium assumption is that whenever the investors uncover a lie about one project, they commit to auditing the rest of the collateral in order to appropriate the whole of it). The IC constraint for the conglomerate not to mis-report state 2 as state 0 is:

(G20h): $(2l_0 - l_0^2) \cdot 2R \geq D_2$.

The partial collusion proof in state 1 is the same as (P1b):

(P1h): $d_1 - D_1 \geq d_2 - D_2$.

We first pin down $l_0^H(D_1, D_2)$ and $l_1^H(D_1, D_2)$. (G21h) is binding, as the investors minimize l_1, which pins down $l_1^H = \frac{D_2 - D_1}{2R - D_1}$. l_0 is present in both (G10h) and (G20h), which implies $l_0 \geq \frac{D_1}{R}$ and $l_0 \geq 1 - \sqrt{1 - \frac{D_2}{2R}}$ respectively. Thus the minimum $l_0^H = \max(\frac{D_1}{R}, 1 - \sqrt{1 - \frac{D_2}{2R}})$. Thus:

(B1): $(l_0^H, l_1^H) = (\max(\frac{D_1}{R}, 1 - \sqrt{1 - \frac{D_2}{2R}}), \frac{D_2 - D_1}{2R - D_1})$.

We then move on to pin down X's rent. Let $m_1 = d_1 - D_1$ (the ex post rent to X in state 1), $m_2 = 2d_2 - D_2$ (the rent in state 2), and $V = 2q(1-q)m_1 + q^2 m_2$ (the ex ante gross rent). Using these notations, (P2h) becomes $m_2 \geq 2(1 - l_1)m_1 + D_2 - 2l_1 R$, and (P1h) becomes $m_1 \geq \frac{m_2 - D_2}{2}$. m_i is nonnegative by (LL). To minimize V, $m_1 = 0$ and $m_2 = D_2 - 2l_1 R$. That is, (P2h) is binding and (P1h), equivalent to $0 \geq -l_1 R$, is unbinding. The m_2 so obtained is nonnegative: By (B1), $D_2 - 2l_1 R = \frac{(2R-D_2)D_1}{2R-D_1} \geq 0$. Thus:

(B2): $(m_1, m_2) = (0, D_2 - 2l_1 R)$.

Lastly, the (IR-Ih) is binding. Substitute (B1) into the binding (IR-Ih),

(B3): $q^2 D_2 + 2q(1-q)D_1 = 2 + 2[(1-q)^2 l_0^H(D_1, D_2) + q(1-q)l_1^H(D_1, D_2)]C$.

(B3) implicitly defines a function $D_2(D_1)$. $\{(D_1, D_2)|D_2 = D_2(D_1)\}$ is then the set of all feasible securities. If the repayment in state 1 (D_1) decreases, as compensation, the repayment in state 2 (D_2) has to increase, that is, $D_2' < 0$. Here, and for the rest of the proof, "$'$" represents the full derivative with respect to D_1.

Last, let $C^H \equiv 2C[(1-q)^2 l_0 + q(1-q)l_1]$ be the total auditing costs. Then the total financial costs consist of the investment costs (2) and X's rent (V) and C^H. The entrepreneurs' problem becomes $\min_{D_1, D_2} V + C^H$, s.t. (B1)-(B3).

The lemma asserts that the minimization happens at $D_1 = D_2$. As $D_1 \leq D_2$ is assumed, to prove the lemma, it suffices to show that $(V + C^H)' < 0$ everywhere. As $l_0 = \max(1 - \sqrt{1 - \frac{D_2}{2R}}, \frac{D_1}{R})$, we consider two cases depending whether $1 - \sqrt{1 - \frac{D_2}{2R}} \leq \frac{D_1}{R}$ or not.

Consider first the case in which $1 - \sqrt{1 - \frac{D_2}{2R}} \leq \frac{D_1}{R}$ and thus $l_0 = \frac{D_1}{R}$. To get an explicit trade-off between the rent (V) and the auditing costs (C^H), notice that $V = q^2 m_2 = q^2 D_2 - 2q^2 R l_1 \Rightarrow q^2 D_2 = V + 2q^2 R l_1$. And $l_0 = \frac{D_1}{R} \Rightarrow D_1 = l_0 R$. Let $l \equiv q l_1 + (1-q) l_0$ and substitute these into (B3), we get $V + 2q^2 l_1 R + 2q(1-q) l_0 R = 2 + 2(1-q)Cl \Leftrightarrow V + 2qRl = 2 + 2(1-q)Cl \Leftrightarrow V + 2(qR - (1-q)C)l = 2$. As $l = \frac{C^H}{2(1-q)C}$, it follows that $V + \frac{qR-(1-q)C}{(1-q)C} C^H = 2$. Then $(V + C^H)' = \frac{qR-2(1-q)C}{qR-(1-q)C} V'$. By Assumption 2, $qR - 2(1-q)C \leq 0$. Thus $(V + C^H)' < 0 \Leftrightarrow V' > 0$. $V = q^2 m_2$ and $m_2 = D_2 - 2l_1 R = \frac{(2R-D_2)D_1}{2R-D_1}$. Then, $V' \propto m_2' > 0$, since $\frac{\partial m_2}{\partial D_1} > 0$, $\frac{\partial m_2}{\partial D_2} < 0$, and $D_2' < 0$. Hence $(V + C^H)' < 0$ in this case.

Consider the case where $1 - \sqrt{1 - \frac{D_2}{2R}} > \frac{D_1}{R}$ and thus $l_0 = 1 - \sqrt{1 - \frac{D_2}{2R}}$. Then, $\frac{dl_0}{dD_2} > 0$ and $l_0' = \frac{dl_0}{dD_2} D_2' < 0$. As $V + C^H = q^2 D_2 - 2q^2 R l_1 + 2C[(1-q)^2 l_0 + q(1-q) l_1] = q^2 D_2 - 2qS l_1 + 2C(1-q)^2 l_0$ (remember $S = qR - (1-q)C$), $(V + C^H)' = q^2 D_2' - 2qS l_1' + 2C(1-q)^2 l_0' < q^2 D_2' - 2qS l_1'$. In order to show $(V + C^H)' < 0$, it suffices to prove that $q^2 D_2' - 2qS l_1' < 0$. By

(B3), $q^2D_2' + 2q(1-q) = 2(1-q)^2Cl_0' + 2q(1-q)Cl_1'$. q^2D_2' is smaller than the LHS of this equation, and the RHS is smaller than $2q(1-q)Cl_1'$, as $l_0' < 0$. Therefore, $q^2D_2' < 2q(1-q)Cl_1'$. Then, $q^2D_2' - 2qSl_1' < 2q(1-q)Cl_1' - 2qSl_1' = 2q[(1-q)C - S]l_1' < 0$; as for the last inequality, we applies $(1-q)C - S = 2(1-q)C - qR \geq 0$ and $l_1' < 0$ ($l_1 = \frac{D_2-D_1}{2R-D_1}$ by (B1) so that $\frac{\partial l_1}{\partial D_1} < 0$ and $\frac{\partial l_1}{\partial D_2} > 0$). Therefore, $(V + C^H)' < 0$ in this case.

To sum up, the solution to the entrepreneurs' problem is $D_1 = D_2 = H$. Accordingly, by (B1), $l_1 = 0; l_0 = \frac{D}{R}$. Substituting all these into (B3), we have $[q^2 + 2q(1-q)]H = 2 + 2C(1-q)^2\frac{H}{R}$, which implies $H = \frac{2R}{q^2R+2(1-q)S}$. By (B2), $m_1 = 0; m_2 = D_2 = H$. Then $d_1 = D_1 + m_1 = H; d_2 = \frac{m_2+D_2}{2} = H$. Indeed the gross rent to X is q^2H; therefore, if $q^2H > 2m$, (IR-X) is unbinding and the optimal mechanism is as specified in the lemma.

Q.E.D.

The Proof of Lemma 8

The lemma is equivalent to:

Lemma B8: if $C_2 \geq 2C$, $l^B(m) \geq \frac{2}{q^2R+2(1-q)S} \equiv l^H$.

Proof: If $m < \frac{1}{2}q^2 \cdot \frac{2+(1-q)^2C_2}{2q-q^2}$, $l^B(m) = 1$. As $l^H \leq 1$ always, the assertion holds true.

If $m \geq \frac{1}{2}q^2 \cdot \frac{2+(1-q)^2C_2}{2q-q^2}$, $l^B(m)$ is the solution for l_0 of (*3*) and ((*4*), and is thus the solution of the equation: $(1-q)^2C_2 = f(l_0) \equiv \frac{2(2-q)m}{2-(2-q)l_0} - \frac{2}{l_0}$. Let $\widetilde{l}$ denote the solution of $f(l_0) = 2(1-q)^2C$. Since $f'(\cdot) > 0$ and $C_2 > 2C$, $l^B(m) > \widetilde{l}$. To prove Assertion B8, therefore, it suffices to prove that $\widetilde{l} \geq l^H$. Given $f'(\cdot) > 0$, the inequality is equivalent to $f(\widetilde{l}) \geq f(l^H) \Leftrightarrow 2(1-q)^2C \geq \frac{2(2-q)m}{2-(2-q)l^H} - \frac{2}{l^H} = \frac{2(2-q)m}{2-(2-q)l^H} - [q^2R + 2(1-q)S] \Leftrightarrow 2(1-q)^2C + q^2R + 2(1-q)S \geq \frac{2(2-q)m}{2-(2-q)l^H}|_{S=qR-(1-q)C} \Leftrightarrow qR(2-q) \geq \frac{2(2-q)m}{2-(2-q)l^H} \Leftrightarrow qR(1 - \frac{2-q}{q^2R+2(1-q)S}) \geq m$. By Assumption 2, $m < C\frac{(1-q)q^2R}{(q^2R+2(1-q)S)S}$. The last inequality of the chain above is thus implied by $qR(1 - \frac{2-q}{q^2R+2(1-q)S}) \geq C\frac{(1-q)q^2R}{(q^2R+2(1-q)S)S} \Leftrightarrow q^2R + 2(1-q)S - (2-q) \geq \frac{(1-q)qC}{S}|_{S\geq 1} \Leftarrow q^2R + 2(1-q)S - (2-q) \geq (1-q)qC|_{S=qR-(1-q)C} \Leftrightarrow (2-q)(qR-1) \geq (1-q)(2-q)C \Leftrightarrow qR - 1 \geq (1-q)C$, assumed in Assumption 2.

Q.E.D.

If $2(1-q)C < qR$, ***IF dominates Conglomeration when (G20h) Is Binding***:

Then, then by the binding (G20h), $D_2 = (4l_0 - 2l_0^2)R$ (see the proof of Lemma 5) under Conglomeration. D_1 is decided by the binding (G21h): $D_1 = \frac{D_2 - 2l_1 R}{1-l_1} = \frac{4l_0 - 2l_0^2 - 2l_1}{1-l_1}R$. The ex post rent $m_2 = D_2 - 2l_1 R = (4l_0 - 2l_0^2 - 2l_1)R$. Note that $m_2 \geq 0$ implies that $2l_0 - l_0^2 - l_1 \geq 0$, and that (G10h), namely $D_1 \leq l_0 R$, implies that $\frac{4l_0 - 2l_0^2 - 2l_1}{1-l_1} \leq l_0 \Leftrightarrow l_1 \geq \frac{3l_0 - 2l_0^2}{2-l_0}$. Therefore:

(*): $2l_0 - l_0^2 - l_1 \geq 0$ and $l_1 \geq \frac{3l_0 - 2l_0^2}{2-l_0}$.

Under Conglomeration, given $\{l_0, l_1)$, the overall costs are $C^H = (1-q)l_e C + \frac{1}{2}q^2 m_2 = (1-q)l_e C + q^2(2l_0 - l_0^2 - l_1)R$, where $l_e = (1-q)l_0 + ql_1$, and the net income per project to investors is $N^H = \frac{1}{2}[2q(1-q)D_1 + q^2 D_2] - (1-q)l_e C = 2q(1-q)\frac{2l_0 - l_0^2 - l_1}{1-l_1}R + q^2(2l_0 - l_0^2)R - (1-q)l_e C$. To have IF pump out this amount of income, $l^I = \frac{N^H}{S} = \frac{1}{S}[2q(1-q)\frac{2l_0 - l_0^2 - l_1}{1-l_1}R + q^2(2l_0 - l_0^2)R - (1-q)l_e C]$. Then the costs of IF are $C^I = (1-q)Cl^I = \frac{(1-q)C}{S}[2q(1-q)\frac{2l_0 - l_0^2 - l_1}{1-l_1}R + q^2(2l_0 - l_0^2)R - (1-q)l_e C]$.

$C^H > C^I \Leftrightarrow (1-q)l_e C + q^2(2l_0 - l_0^2 - l_1)R > \frac{(1-q)C}{S}[2q(1-q)\frac{2l_0 - l_0^2 - l_1}{1-l_1}R + q^2(2l_0 - l_0^2)R - (1-q)l_e C] \Leftrightarrow_{1+\frac{(1-q)C}{S} = \frac{qR}{S}} (1-q)l_e C\frac{qR}{S} + q^2(2l_0 - l_0^2 - l_1)R > \frac{(1-q)C}{S}[2q(1-q)\frac{2l_0 - l_0^2 - l_1}{1-l_1}R + q^2(2l_0 - l_0^2)R]$, times $\frac{S}{qR(1-q)C} \Leftrightarrow l_e + \frac{qS}{(1-q)C}(2l_0 - l_0^2 - l_1) > 2(1-q)\frac{2l_0 - l_0^2 - l_1}{1-l_1} + q(2l_0 - l_0^2)$, minus $ql_1 \Leftrightarrow (1-q)l_0 + \frac{qS}{(1-q)C}(2l_0 - l_0^2 - l_1) > 2(1-q)\frac{2l_0 - l_0^2 - l_1}{1-l_1} + q(2l_0 - l_0^2 - l_1) \Leftrightarrow (\frac{S}{(1-q)C} - 1)q(2l_0 - l_0^2 - l_1) + (1-q)l_0 > 2(1-q)\frac{2l_0 - l_0^2 - l_1}{1-l_1} \Leftarrow (1-q)l_0 > 2(1-q)\frac{2l_0 - l_0^2 - l_1}{1-l_1}$, because $2l_0 - l_0^2 - l_1 \geq 0$ by (*) and $\frac{S}{(1-q)C} - 1 = \frac{qR - 2(1-q)C}{(1-q)C} > 0$ if $2(1-q)C < qR$.

$(1-q)l_0 > 2(1-q)\frac{2l_0 - l_0^2 - l_1}{1-l_1} \Leftrightarrow l_0 > 2\frac{2l_0 - l_0^2 - l_1}{1-l_1} = 2(1 - \frac{(1-l_0)^2}{1-l_1}) \Leftrightarrow f(l_1) \equiv l_0 + 2\frac{(1-l_0)^2}{1-l_1} \geq 2$. Note that $f(\cdot)$ is increasing and that by (*) $l_1 \geq \frac{3l_0 - 2l_0^2}{2-l_0}$. Therefore, $f(l_1) \geq f(\frac{3l_0 - 2l_0^2}{2-l_0}) = l_0 + 2(1-l_0)^2(1 - \frac{3l_0 - 2l_0^2}{2-l_0})^{-1} = l_0 + 2(1-l_0)^2(\frac{2(1-l_0)^2}{2-l_0})^{-1} = l_0 + 2 - l_0 = 2$.

Q.E.D.

Hereinafter, p_N^s is denoted as p_s for simplicity, when without confusion.

The Proof of Lemma 9

The IR constraint for the investors is binding, as follows:

(IR-I): $(k\sum_{s\geq k} p_s + \sum_{s\leq k-1} sp_s)d = N + C_N \sum_{s\leq k-1} p_s$.

(IR-I) determines a function $d(k)$. When (IR-X) is not binding and thus not taken into account, the optimal k that minimizes d is decided by the trade-off between the auditing costs and the rent to X: Auditing happens in states $s \leq k-1$ and X receives ex post rent in states $s > k$.

To find the first order condition of the minimization, let us simplify (IR-I). Divide both sides by N. By the Central Limit Theorem (CLT), $\frac{s-Nq}{\sqrt{Nq(1-q)}} \sim N(0,1)$, with the dense and cumulative distribution functions being $\phi(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ and $\Phi(x)$ respectively. And let $k = Nq + h\sqrt{Nq(1-q)}$. Then, $d(k)$ leads the following function.

(C1): $d(h) = \frac{1+\frac{C_N}{N}\Phi(h)}{q+\sqrt{\frac{q(1-q)}{N}}(h(1-\Phi(h))-\phi(h))}$.

Given N, the optimal h satisfies the first order condition

(C2): $\frac{C_N}{N}\phi(h)[q+\sqrt{\frac{q(1-q)}{N}}(h(1-\Phi(h))-\phi(h))] = [1+\frac{C_N}{N}\Phi(h)]\sqrt{\frac{q(1-q)}{N}}(1-\Phi(h))$.

Remember $C_N = zN^\alpha$ for some $\alpha \in (0,1]$. Suppose $\frac{C_N}{N}\Phi(h) = o(1)$, which is obvious for $\alpha < 1$ and to be verified for $\alpha = 1$. Then, the RHS of (C2) $\approx \sqrt{\frac{q(1-q)}{N}}(1-\Phi(h))$. Suppose $\frac{h}{\sqrt{N}} = o(1)$ (to be verified later), which implies that the LHS $\approx \frac{C_N}{N}\phi(h)q$. (C2) asymptotically becomes:

(C3): $\frac{qz}{\sqrt{q(1-q)}}N^{\alpha-0.5} = \frac{1-\Phi(h)}{\phi(h)}$.

Claim: The solution of (C3) is $h^* = \left\{\begin{array}{ccc} \frac{\sqrt{q(1-q)}}{qz}N^{0.5-\alpha}+o & & \alpha < 0.5 \\ \widehat{h} & \text{if} & \alpha = 0.5 \\ -\sqrt{(2\alpha-1)\log N}+o & & \alpha > 0.5 \end{array}\right\}$, where $\widehat{h}$ is a constant.

Proof: If $\alpha = 0.5$, (hb) becomes $\frac{qz}{\sqrt{q(1-q)}} = \frac{1-\Phi(h)}{\phi(h)}$. $\lim_{h\to-\infty}\frac{1-\Phi(h)}{\phi(h)} = \infty$ and by L'Hospital's rule, $\lim_{h\to+\infty}\frac{1-\Phi(h)}{\phi(h)} = \lim_{h\to+\infty}\frac{-\phi(h)}{-\phi(h)h} = 0$. $\frac{1-\Phi(h)}{\phi(h)}$ is decreasing, since $\{\frac{1-\Phi(h)}{\phi(h)}\}' = \frac{(1-\Phi(h))h-\phi(h)}{\phi(h)}$ and $(1-\Phi(h))h-\phi(h) = -\int_h^\infty(t-h)\phi(t)dt < 0$. Therefore the equation has a unique solution, $\widehat{h}$.

If $\alpha < 0.5$, (C3) implies that when $N \to \infty$, $\frac{1-\Phi(h)}{\phi(h)} = \frac{qz}{\sqrt{q(1-q)}}N^{\alpha-0.5} \to 0$. Thus $h \to +\infty$. It follows that $\frac{1-\Phi(h)}{\phi(h)} \approx \frac{1}{h}$, as $\lim_{h\to+\infty}h(\frac{1-\Phi(h)}{\phi(h)}-\frac{1}{h}) = \lim_{h\to+\infty}\frac{h(1-\Phi(h))}{\phi(h)}-1|_{\text{L'Hospital}} = \lim_{h\to+\infty}\frac{1-\Phi(h)}{-h\phi(h)} = \lim_{h\to+\infty}\frac{1}{-h}\frac{1-\Phi(h)}{\phi(h)} = 0$. Therefore, $\frac{qz}{\sqrt{q(1-q)}}N^{\alpha-0.5} \approx \frac{1}{h} \Rightarrow h \approx \frac{\sqrt{q(1-q)}}{qz}N^{0.5-\alpha}$.

If $\alpha > 0.5$, (C3) implies that when $N \to \infty$, $\frac{1-\Phi(h)}{\phi(h)} \to \infty$, and thus $h \to -\infty$, which implies $1 - \Phi(h) \to 1$. Therefore, $\frac{qz}{\sqrt{q(1-q)}} N^{\alpha - 0.5} = \phi(h)^{-1} = \sqrt{2\pi} e^{\frac{h^2}{2}}$, as $\phi(h) = \frac{1}{\sqrt{2\pi}} e^{-\frac{h^2}{2}}$. Take log operation on both sides$\Rightarrow \frac{h^2}{2} \approx (\alpha - 0.5) \log N \Rightarrow h \approx -\sqrt{(2\alpha - 1)\log N}$.

q.e.d.

Let $x = \frac{C_N}{N} \Phi(h^*)$ and $y = \sqrt{\frac{q(1-q)}{N}} (h(1 - \Phi(h)) - \phi(h))|_{h=h^*}$, so $d(h^*) = \frac{1+x}{q+y}$ by (C1). Then, $x = o(1)$ indeed; for $\alpha = 1$, $x = O(\Phi(h^*)) = |_{\text{the claim}} O(\Phi(-\sqrt{\log N})) = o(1)$. And by the claim, $y = O(\frac{h}{\sqrt{N}}) = o(1)$. Therefore, $d \to \frac{1}{q}$.

The gross rent to X is $V_N = \sum_{s \geq k} d(s-k) p_s$. Apply the CLT, $d \approx \frac{1}{q}$ and $k = Nq + h\sqrt{Nq(1-q)}$, and let $s = Nq + t\sqrt{Nq(1-q)}$. We have $V_N \approx \frac{\sqrt{Nq(1-q)}}{q} \int_{h^*}^{\infty} (t - h^*) \phi(t) dt$. The integration equals $\phi(h^*) - h^*(1 - \Phi(h^*))$. It converges to $\phi(\widehat{h}) - \widehat{h}(1 - \Phi(\widehat{h}))$, for $\alpha = 0.5$, and $\approx -h^*$ for $\alpha > 0.5$ ($h^* \to -\infty$); if $a < 0.5$, as $h^* > 0$, the integration is smaller than $\phi(h^*)$, which multiplied by $\sqrt{N}$ goes to 0, as $h^* = O(N^{0.5-\alpha})$. Therefore, $V_N = \left\{ \begin{array}{ccc} o(1) & & \alpha < 0.5 \\ O(\sqrt{N}) & \text{if} & \alpha = 0.5 \\ O(\sqrt{(2\alpha-1) N \log N}) & & \alpha > 0.5 \end{array} \right\}$. This shows that the gross rent to X is at most in the order of $\sqrt{N \log N}$, which happens for $\alpha > 0.5$.

Q.E.D.

The Proof of Lemma 10

As $\frac{k}{N} \to \frac{q}{1+m}$, it suffices to prove the lemma for $N = k\frac{1+m}{q}$. Let $\theta = \frac{q}{1+m} = \frac{k}{N}$. An intuitive proof of the lemma is as follows. $E(\frac{1}{k-s}|s \leq k-1) = \frac{p_{k-1} \cdot 1 + p_{k-2} \cdot \frac{1}{2} + \ldots + p_0 \frac{1}{k}}{p_{k-1} + p_{k-2} + \ldots + p_0} = \frac{1 + \frac{p_{k-2}}{p_{k-1}} \cdot \frac{1}{2} + \ldots \frac{p_0}{p_{k-1}} \cdot \frac{1}{k}}{1 + \frac{p_{k-2}}{p_{k-1}} + \ldots \frac{p_0}{p_{k-1}}}$. Given N, for any $i \geq 2$, $\frac{p_{k-i}}{p_{k-1}} = \frac{C_N^{k-i} q^{k-i} (1-q)^{N-k+i}}{C_N^{k-1} q^{k-1} (1-q)^{N-k+1}} = (\frac{1-q}{q})^{i-1} \frac{(k-1)\cdot(k-2)\cdot\ldots(k-i+1)}{(N-k+i)\cdot(N-k+i-1)\cdot\ldots(N-k+2)}$, where $C_N^i = \frac{N!}{i!(N-i)!}$ is the number of combinations. Given that k and N are large, $\frac{k-1}{N-k+i} \approx \frac{k-2}{N-k+i-1} \approx \ldots \frac{k-i+1}{N-k+2} \approx \frac{k}{N-k} = \frac{\theta}{1-\theta}$. Then $\frac{p_{k-i}}{p_{k-1}} \approx (\frac{(1-q)\theta}{q(1-\theta)})^{i-1} \equiv \lambda^{i-1}$, where $\lambda \equiv \frac{(1-q)\theta}{q(1-\theta)} < 1$ as $\theta < q$. Then, $E(\frac{1}{k-s}|s \leq k-1) \approx \frac{1 + \lambda \cdot \frac{1}{2} + \ldots \lambda^{k-1} \cdot \frac{1}{k}}{1 + \lambda + \ldots \lambda^{k-1}} = \frac{1-\lambda}{\lambda} \frac{\int_0^{\lambda} \frac{1-t^k}{1-t} dt}{1-\lambda^k} \to \frac{1-\lambda}{\lambda} \log \frac{1}{1-\lambda}$ when $k \to \infty$.

For a strict proof, we are going to show that for the comparison between $\lim_{N\to\infty} E(\frac{1}{k-s}|s \leq k-1)$ and $\frac{1-\lambda}{\lambda} \log \frac{1}{1-\lambda}$, both "$\geq$" and "$\leq$ ” hold true, so that they must be equal. For the ” $\geq$ ”

part, note that $\frac{p_{k-i+1}}{p_{k-i}} = \frac{q}{1-q} \cdot \frac{k-i+1}{N-k+i} < \frac{qk}{(1-q)(N-k)} = \lambda$ for any $i = 2, 3...k$. The following lemma is useful to establish " $\geq$ ".

Lemma A2: If $\frac{a_{i+1}}{a_i} < \lambda$, then $\frac{a_1 + a_2 \frac{1}{2} + ... a_k \frac{1}{k}}{a_1 + a_2 + ... a_k} \geq \frac{1 + \lambda \cdot \frac{1}{2} + ... \lambda^{k-1} \cdot \frac{1}{k}}{1 + \lambda + ... \lambda^{k-1}}$.

Proof: By mathematical induction with respect to k. For $k = 1$, the inequality surely holds true. Assume that the lemma holds true for a given k. Consider the case for $k+1$. Let $V_k = \frac{a_1 + a_2 \frac{1}{2} + ... a_k \frac{1}{k}}{a_1 + a_2 + ... a_k}$ and $W_k = \frac{1 + \lambda \cdot \frac{1}{2} + ... \lambda^{k-1} \cdot \frac{1}{k}}{1 + \lambda + ... \lambda^{k-1}}$. By the induction assumption $V_k \geq W_k$. Both V_k and W_k are a convex combination of $1, \frac{1}{2}, ..., \frac{1}{k}$, and are thus bigger than $\frac{1}{k+1}$. Notice that $\frac{a_{k+1}}{a_1 + a_2 + ... a_{k+1}} < \frac{\lambda^k}{1 + \lambda + ... \lambda^k}$, as it $\Leftrightarrow \frac{a_1}{a_{k+1}} + \frac{a_2}{a_{k+1}} + ... 1 > \lambda^{-k} + \lambda^{-(k-1)} + ... 1$, which is true because $\frac{a_i}{a_{k+1}} = \frac{a_i}{a_{i+1}} \cdot \frac{a_{I+1}}{a_{i+2}} ... \frac{a_k}{a_{k+1}} > (\frac{1}{\lambda})^{k+1-i}$ for any $i = 1, 2, .., k$. Then $V_{k+1} = \frac{a_1 + a_2 + ... a_k}{a_1 + a_2 + ... a_{k+1}} V_k + \frac{a_{k+1}}{a_1 + a_2 + ... a_{k+1}} \frac{1}{k+1} > \frac{1 + \lambda + ... \lambda^{k-1}}{1 + \lambda + ... \lambda^k} V_k + \frac{\lambda^k}{1 + \lambda + ... \lambda^k} \frac{1}{k+1} \geq \frac{1 + \lambda + ... \lambda^{k-1}}{1 + \lambda + ... \lambda^k} W_k + \frac{\lambda^k}{1 + \lambda + ... \lambda^k} \frac{1}{k+1} = W_{k+1}$, where for the first inequality we apply $V_k > \frac{1}{k+1}$ and $\frac{a_{k+1}}{a_1 + a_2 + ... a_{k+1}} < \frac{\lambda^k}{1 + \lambda + ... \lambda^k}$. q.e.d.

By Lemma A2, $E(\frac{1}{k-s} | s \leq k-1) = \frac{p_{k-1} \cdot 1 + p_{k-2} \cdot \frac{1}{2} + ... + p_0 \frac{1}{k}}{p_{k-1} + p_{k-2} + ... + p_0} \geq \frac{1 + \lambda \cdot \frac{1}{2} + ... \lambda^{k-1} \cdot \frac{1}{k}}{1 + \lambda + ... \lambda^{k-1}}$. Let N and $k = \frac{q}{1+m} N$ both go to infinity on both sides of the inequality, and we have $\lim_{N \to \infty} E(\frac{1}{k-s} | s \leq k-1) \geq \frac{1-\lambda}{\lambda} \log \frac{1}{1-\lambda}$.

To prove $\lim_{N \to \infty} E(\frac{1}{k-s} | s \leq k-1) \leq \frac{1-\lambda}{\lambda} \log \frac{1}{1-\lambda}$, let us restore notations p_N^s and disuse p_s. For any $L < k$, $\frac{p_N^{k-1} \cdot 1 + p_N^{k-2} \cdot \frac{1}{2} + ... + p_N^0 \frac{1}{k}}{p_N^{k-1} + p_N^{k-2} + ... + p_N^0} < \frac{p_N^{k-1} \cdot 1 + p_N^{k-2} \cdot \frac{1}{2} + ... + p_N^{k-L} \frac{1}{L}}{p_N^{k-1} + p_N^{k-2} + ... + p_N^{k-L}}$, because the former is the convex combination of the latter and the smaller terms, $\frac{1}{L+1}, \frac{1}{L+2} ... \frac{1}{K}$. For this inequality, keep L fixed and let N (and $k = \frac{q}{1+m} N$) go to infinity. Then the left hand side goes to $\lim_{N \to \infty} E(\frac{1}{k-s} | s \leq k-1)$. The right hand side goes to $\frac{1 + \lambda \cdot \frac{1}{2} + ... \lambda^{L-1} \cdot \frac{1}{L}}{1 + \lambda + ... \lambda^{L-1}}$, because $\frac{p_N^{k-i}}{p_N^{k-1}} = (\frac{1-q}{q})^{i-1} \frac{(k-1) \cdot (k-2) \cdot ... (k-i+1)}{(N-k+i) \cdot (N-k+i-1) \cdot ... (N-k+2)} \to \lambda^{i-1}$ for any i no bigger than the given L. Therefore, for any given L, $\lim_{N \to \infty} E(\frac{1}{k-s} | s \leq k-1) \leq \frac{1 + \lambda \cdot \frac{1}{2} + ... \lambda^{L-1} \cdot \frac{1}{L}}{1 + \lambda + ... \lambda^{L-1}}$. Let L go to infinity, $\lim_{N \to \infty} E(\frac{1}{k-s} | s \leq k-1) \leq \frac{1-\lambda}{\lambda} \log \frac{1}{1-\lambda}$.

Therefore, $\lim_{N \to \infty} E(\frac{1}{k-s} | s \leq k-1) = \frac{1-\lambda}{\lambda} \log \frac{1}{1-\lambda}$. Substitute $\lambda = \frac{(1-q)\theta}{q(1-\theta)} = \frac{(1-q)\frac{q}{1+m}}{q\frac{1+m-q}{1+m}} = \frac{1-q}{1-q+m}$. Then $\frac{1-\lambda}{\lambda} \log \frac{1}{1-\lambda} = \frac{m}{1-q} \log \frac{1-q+m}{m}$.

Q.E.D.

The Proof of Lemma 11

To find the optimal mechanisms with Assumption 3 removed when the IR for the expert is not-binding, I only need to find the optimal mechanisms out of those with $D_2 \leq D_1$, as Lemmas 3 and 5 have given rise to the optimal mechanisms out of those with $D_2 \geq D_1$.

To prove (a): I shall show that under Conglomeration, the optimal mechanism out of those with $D_2 \leq D_1$ also features $D_2 = D_1$ and hence is the same as was specified in Lemma 5.

If $D_2 \leq D_1$, then $l_1 = 0$; (G12h) is binding, which pins down $l_2 = \frac{D_1 - D_2}{R - D_2}$; and (G10h) implies (G20h): $D_2 \leq D_1 \leq l_0 R < (2l_0 - l_0^2) \cdot 2R$, and hence (G10h) is binding, which pins down $l_0 = \frac{D_1}{R}$. Substituting these l_0, l_1, and l_2 as functions of (D_1, D_2) into the binding (IR-Ih), as in the proof of Lemma 5, implicitly determines D_2 as a function of D_1:

(E1): $q^2 D_2 + 2q(1-q)D_1 = 2 + 2C[(1-q)^2 \cdot \frac{D_1}{R} + q^2 \cdot \frac{D_1 - D_2}{R - D_2}]$

The function $D_2(D_1)$, we saw, is decreasing.

Furthermore, as before, (CP2h) is binding, which implies (CP1h) and gives $m_2 = D_2 - 2l_1 R = D_2$ (see (B2) of the Proof of Lemma 5). Hence the total financial costs to the entrepreneurs are $V + C^H = q^2 D_2 + C^H$, where $C^H = 2C[(1-q)^2 l_0 + q^2 l_2]$ for this case. To show the optimization occurs at $D_2 = D_1$, I prove that so long as $D_2 < D_1$, the costs $V + C^H$ are diminished if D_2 is increased by a little, that is, $d(V + C^H) < 0$ with $dD_2 > 0$, which causes $dD_1 = \frac{dD_2}{D_2(D_1)} < 0$.

By the binding (IR-Ih), $q^2 D_2 + 2q(1-q)D_1 = 2 + C^H \Rightarrow dC^H = q^2 dD_2 + 2q(1-q)dD_1$. Hence $d(V + C^H) = 2q^2 dD_2 + 2q(1-q)dD_1$. Hence $d(V + C^H) < 0 \Leftrightarrow q dD_2 < (1-q)(-dD_1)$, which is proved true below.

Function $D_2(D_1)$ is determined by (E1). Total differentiate the equation on both sides: $q^2 dD_2 + 2q(1-q)dD_1 = \frac{2C(1-q)^2}{R} dD_1 + 2Cq^2[\frac{dD_1}{R - D_2} - \frac{R - D_1}{(R - D_2)^2} dD_2] \Leftrightarrow [\frac{q^2}{2C} + \frac{q^2(R - D_1)}{(R - D_2)^2}]dD_2 = [\frac{(1-q)S}{RC} - \frac{q^2}{R - D_2}](-dD_1)$. Notice that $D_2 < D_1 \leq R$ (the limited liability for the entrepreneurs) and hence $\frac{q^2}{R - D_2} > 0$. Thus the right hand side of the last equation is strictly smaller than $\frac{(1-q)S}{RC}(-dD_1)$ and the left hand side is bigger than $\frac{q^2}{2C} dD_2$. Therefore, from that equation we have $(1-q)(-dD_1) > \frac{RC}{S} \cdot \frac{q^2}{2C} dD_2 = dD_2 \cdot q \cdot \frac{qR}{2S} \geq q \cdot dD_2$, where the last inequality applies $qR \geq 2S$ which is equivalent to $2(1-q)C \geq qR$, laid down in Assumption 2. Hence I have proved that $q dD_2 < (1-q)(-dD_1)$ and thus $\frac{d(V + C^H)}{dD_2} < 0$ if $D_2 < D_1$. That is, the optimal mechanism out of those with $D_2 \leq D_1$ features $D_2 = D_1$.

To prove (b): If $D_2 \leq D_1$, then $l_1 = 0$. The binding (CP2f) becomes:

(E2): $d_2 - D_2 = d_1 - D_1$.

(E2) implies $d_2 - d_1 = d_2 - D_1 \leq 0$ and (G20f), which is then not binding: $D_2 = d_2 - d_1 + D_1|_{D_1 \leq l_0 d_1} \leq d_2 - (1-l_0)d_1|_{d_2 \leq d_1} \leq d_2 - (1-l_0)d_2 = l_0 d_2 < 2l_0 d_2$. Then l_0 only appears in (G10f) and is pinned down by it being binding:

(E3): $D_1 = l_0 d_1$.

From (E2) and (E3), I solve D_2 as a function of d and l:

(E4): $D_2 = d_2 - (1 - l_0)d_1$

And l_2 is pinned down by (G12f) being binding:

(E5): $l_2 d_1 + (1 - l_2)D_2 = D_1$.

Substitute (E3) and (E4) into (E5):

(E6): $(1 - l_2)d_2 = (1 - 2l_2 + l_0 l_2)d_1$.

Note that $d_2 \geq 0 \Leftrightarrow 1 - 2l_2 + l_0 l_2 \geq 0 \Leftrightarrow l_2 \leq \frac{1}{2-l_0}$. Hence, if $l_2 = 0$, then $l_1 = 0$.

Finally, substitute (E3) and (E4) into the binding (IR-If):

(E7): $q^2 d_2 + ((2q - q^2)l_0 - q^2)d_1 = 2 + C_2((1 - q)^2 l_0 + q^2 l_2)$.

The entrepreneurs' problem is then to $\min(1 - q)d_1 + qd_2$ s.t. (E6) and (E7). To solve the problem, consider two cases. First, the case where $l_2 < 1$.

The Case where $l_2 < 1$:

For this case, I solve $d_2 = \frac{1-2l_2+l_0 l_2}{1-l_2} d_1 \equiv t(l_0, l_2)d_1$. (E7) becomes:

(E8): $q^2 \frac{l_0-l_2}{1-l_2} d_1 + 2q(1 - q)l_0 d_1 = 2 + (1 - q)^2 C_2 l_0 + q^2 C_2 l_2$.

It is easy to show the optimization occurs at $l_0 = 1$, $l_2 = 0$, and $d_1 = d_2 = F \equiv \frac{2+C_v(1-q)^2}{2q-q^2}$. The Lagrangian is $L = (qt+(1-q))d_1+\lambda[2+(1-q)^2 C_2 l_0+q^2 C_2 l_2-q^2 \frac{l_0-l_2}{1-l_2} d_1-2q(1-q)l_0 d_1]$, where the multiplier $\lambda > 0$. Then, $\frac{\partial L}{\partial l_0} = qd_1 \frac{l_2}{1-l_2} + \lambda((1-q)^2 C_2 - 2q(1-q)d_1 - \frac{q^2 d_1}{1-l_2}) < 0$ at $l_2 = 0$, and $d_1 = d_2 = F \equiv \frac{2+C_v(1-q)^2}{2q-q^2}$. Hence the optimal $l_0 = 1$. And $\frac{\partial L}{\partial l_2} = -qd_1 \frac{1-l_0}{(1-l_2)^2} + \lambda(q^2 C_2 + q^2 d_1 \frac{1-l_0}{(1-l_2)^2}) > 0$ if $l_0 = 1$, and hence the optimal $l_2 = 0$. This solution is the same as was given by Lemma 3 and gives rise to the outlay of a successful entrepreneur being $F = \frac{2+C_v(1-q)^2}{2q-q^2}$.

The Case where $l_2 = 1$.

For this case, (E6) implies that $l_0 = 1$. (E7) becomes:

(E9): $q^2 d_2 + 2q(1-q)d_1 = 2 + C_2((1-q)^2 + q^2)$.

To minimize $(1-q)d_1 + qd_2$ s.t. (E9), the entrepreneurs want d_2 as small as possible, as its relative weight in the objective $\frac{q}{1-q}$ is bigger than that in the constraint, $\frac{q^2}{2q(1-q)}$. If $d_2 = 0$, then (E9) implies $d_1 = \frac{2+[(1-q)^2+q^2]C_v}{2q(1-q)}$. For this d_1, the limited liability constraint $d_1 \leq R$ is satisfied if and only if $J \geq 2$, where $J = 2q(1-q)R - [q^2 + (1-q)^2]C_2$. If $J < 2$, then the solution is $d_1 = R$ and then (E9) implies $d_2 = \frac{2-J}{q^2}$. Hence overall, $d_1 = \min\{\frac{2+[(1-q)^2+q^2]C_2}{2q(1-q)}, R\}$ and $d_2 = \max\{0, \frac{2-J}{q^2}\}$. Moreover, substituting $l_0 = l_2 = 1$ into (E3) gives $D_1 = d_1$ and into (E4) gives $D_2 = d_2$. Comparing the objective so obtained, $(1-q)d_1 + qd_2$, with the objective obtained in the former case, F, finishes the proof of part (b) of the lemma.

Q.E.D.

References

[1] Admati, A. and P. Pfleiderer (1990), "Direct and Indirect Sale of Information", *Econometrica*, 58, 901-928.

[2] Biais, B. and L. Germain (2002), "Incentive-Compatibility Contracts for the Sale of Information", *Review of Economic Studies*, 15 (4), 987-1003.

[3] Bond, P. (2004), "Bank and Nonbank Financial Intermediation", *Journal of Finance*, LIX (5) 2489-2529.

[4] Brennan, M. and T. Chordia, (1993), "Brokerage Commission Schedules," *Journal of Finance*, 48, 1379-1402.

[5] Constantinides, M. and B. Grundy (1989), "Optimal Investment with Stock Repurchase and Financing as Signals", *Review of Financial Studies*, 2 (4), 445-465.

[6] Diamond, D. (1984), "Financial Intermediation and Delegated Monitoring", *Review of Economic Studies*, 51 (3) 393-414.

[7] Diamond. D (1996), "Financial Intermediation and Delegated Monitoring: A Simple Case", *Economic Quartly*, Federal Reserve Bank of Richmond, 82.

[8] Diamond, D. and R. Rajan (2001), "Liquidity Risk, Liquidity Creation and Financial Fragility: A Theory of Banking", *Journal of Political Economy,* 109, 287–327.

[9] Gertner, R., D. Scharfstein, and J. Stein (1994), "Internal Versus External Capital Markets", *Quarterly Journal of Economics*, 109, 1211-1230.

[10] Gorton, G. and A. Winton, Financial Intermediation, mimeo, 2002, Wharton Financial Institutions Center, working paper 02-28. http://fic.wharton.upenn.edu/fic/papers/02/0228.pdf .

[11] Hart, O. (1995), *Firms, Contracts, and Financial Structure*, Oxford, UK: Oxford University Press.

[12] Hellwig, M. (2000), "Financial Intermediation with Risk Aversion", *Review of Economic Studies*, 67 (4), 719-742.

[13] Innes, R.(1990), "Limited liability and incentive contracting with ex-ante action choices", *Journal of Economic Theory*, 52, 45-67.

[14] Krasa, S. and A. P. Villamil (1992), "Monitoring the Monitor: An Incentive Structure for a financial intermediary", *Jounal Economic Theory,* 57 (1), 197-221.

[15] Lamont, O. (1997), "Cash Flow and Investment: Evidence From Internal Capital Markets," *Journal of Finance* 52, 83–109.

[16] Mookherjee, D. and I. Png (1989), "Optimal Auditing, Insurance, and Redistribution," *Quarterly Journal of Economics*, 104 (2), 399-415.

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

[18] Williamson, S. D. (1986), "Costly Monitoring, Financial Intermediation, and Equilibrium Credit Rationing", *Journal of Monetary Economics*, 18 (2), 159-179.

[19] Winton, A. (1995), "Delegated Monitoring and Bank Structure in a Finite Economy", *Journal of Financial Intermediation,* 4, 158-187.