

# Bank and Nonbank Financial Intermediation

PHILIP BOND\*

## ABSTRACT

Conglomerates, trade credit arrangements, and banks are all instances of financial intermediation. However, these institutions differ significantly in the extent to which the projects financed absorb aggregate intermediary risk, in whether or not intermediation is carried out by a financial specialist, in the type of projects they fund and in the type of claims they issue to investors. The paper develops a simple unified model that both accounts for the continued coexistence of these different forms of intermediation, and explains why they differ. Specific applications to conglomerate firms, trade credit, and banking are discussed.

CONGLOMERATES, TRADE CREDIT ARRANGEMENTS, AND BANKS are all instances of financial intermediation. In each case, the conglomerate headquarters/supplier/bank obtains funds by selling financial securities, while in turn providing funds in exchange for a claim on project cash flows.<sup>1</sup> However, in spite of the fundamental similarity between these forms of financial intermediation, important differences exist between them. In particular:

- (I) What happens to project financing when the financial intermediary as a whole performs badly? Projects financed by conglomerates are adversely affected, in the sense that resources available to each division for investment are curtailed.<sup>2</sup> On the other hand, large bank borrowers are not much affected by a decline in the fortunes of their lending bank.<sup>3</sup>
- (II) Who performs the intermediation function? Both in the case of conglomerates and trade credit, intermediation is carried out in conjunction with real economic activity. Historically, at least some commercial banks have

\*Philip Bond is at The Wharton School, University of Pennsylvania. I thank seminar audiences at the AFE and Gerzensee, Douglas Diamond, Michael Fishman, Arvind Krishnamurthy, Philip Strahan, Robert Townsend, and especially Richard Green (editor) and an anonymous referee for some very helpful comments. I am grateful to the Institute for Advanced Study for hospitality and financial support (in conjunction with Deutsche Bank) over the academic year 2002–2003. Any remaining errors are, of course, my own.

<sup>1</sup>Freixas and Rochet (1997, p. 15) suggest that a financial intermediary is “an economic agent who specializes in the activities of buying and selling (at the same time) financial contracts and securities.”

<sup>2</sup>See, for example, Lamont (1997) and Shin and Stulz (1998).

<sup>3</sup>The experience of small borrowers from small banks is in some respects similar to that of conglomerate divisions—see the discussion of credit crunch-like phenomena in the text below.

also fitted this pattern.<sup>4</sup> In contrast, modern commercial banks are run by financial specialists.

- (III) What types of project does an intermediary finance? On the one hand, commercial banks finance only relatively low-risk projects (or at least the low-risk component of cash flows). This is not the case for conglomerates.
- (IV) What sort of liabilities does a financial intermediary issue to fund itself? Different types of financial intermediary issue different mixes of financial securities: A large fraction of the claims issued by commercial banks are very low risk, while this is not the case for conglomerates.

In this paper, I develop a unified model (based on a single friction) that explains how these four features of financial intermediation are linked. By doing so, I account for the continuing coexistence of different forms of intermediation. In the model, the viability of all forms of financial intermediation mentioned depends on the advantages stemming from diversification. At the same time, the model accounts for why, given this shared origin, the questions of how much aggregate intermediary risk the projects financed should absorb, and who should actually intermediate, are resolved so differently in different forms of intermediation. The model's main prediction is that financial intermediaries fall into one of two broad categories. First, there are intermediaries resembling conglomerates. Intermediaries in this category finance high-risk/low-quality projects (III).<sup>5</sup> Consequently, the liabilities they issue to investors are also relatively high risk (IV). Because investors are left exposed to a substantial amount of risk, it is worthwhile to reduce this exposure by having the projects funded absorb some of each other's cash flow fluctuations (I).

Second, there are intermediaries that broadly resemble banks. Institutions of this type fund comparatively low-risk/high-quality projects (III). This allows them to issue mostly low-risk liabilities, such as bank deposits and low-risk bonds (IV). Since the liabilities are already low risk, there is then little to gain by having borrowers absorb some of each other's risk (I).<sup>6</sup>

Moreover, within the latter category we can distinguish between the cases in which intermediation is performed by a financial specialist, such as a modern

<sup>4</sup> See, for instance, Lamoreaux's (1994) study of 19th-century New England banking, in which she describes how banks were run largely by leading local merchants, with many of their loans going to these same individuals (i.e., "insider lending").

<sup>5</sup> Evidence suggests that divisions that form part of conglomerates are less profitable than comparable nonconglomerate firms (see, e.g., Graham, Lemmon, and Wolf (2002) and Campa and Kedia (2002)). Consequently, conglomerates will tend to trade at a discount relative to stand-alone firms (the well-established "diversification discount").

<sup>6</sup> One point of clarification is worth making here. As we will see, it is often the case that the intermediary runs a project himself, as well as financing other projects. The intermediary's own project is, of course, always exposed to the cash flow fluctuations of these other projects. The difference between conglomerate-like and bank-like intermediaries lies in whether or not projects *not* run directly by the intermediary are exposed to the cash flow fluctuations of other projects.

bank, and those in which intermediation is performed by a nonspecialist, such as trade credit arrangements and early forms of banking (II). The model predicts that when the intermediary obtains funding from a relatively small number of investors, then intermediation by a nonspecialist is preferable. Specialized financial intermediaries such as modern banks emerge as the number of investors rises. And within trade credit relationships the model predicts that funds will flow from the goods supplier to the goods purchaser. That is, it is trade credit rather than prepayment that is the predominant form of interfirm finance.

The key friction in the model is that information is expensive to share. This implies that low-risk securities are generally preferable to higher-risk ones: They are less information sensitive, and so entail less costly transmission of information. Optimal financial arrangements are essentially those in which the fewest possible resources are expended on information transmission. This objective is achieved by finding ways to make as many claims as possible as low risk as possible.

To understand the model's main results, start by observing that financial intermediaries are financed by claims of a variety of different risk levels and seniorities. Commercial banks raise financing by taking deposits, issuing other forms of debt, and often by issuing equity as well. With the exception of deposits, the same is true for conglomerates. Intermediaries can and do fail, and so not all of these different claims are low risk.

When an intermediary's income from its investments is low, this income shortfall must be absorbed by someone. There are three choices: the intermediary itself, the recipients of intermediary finance, and the intermediary's investors. When the income shortfall can be absorbed entirely by the intermediary itself, this will generally be the most efficient option, since the intermediary is the only party to directly observe its portfolio realization.

Now, consider the case in which larger income fluctuations are absorbed by intermediary investors. (As discussed, this is the case for large modern banks.) What happens if in this case we instead make the transfers from the projects financed back to the intermediary contingent on the intermediary's overall performance, with larger transfers when performance is poor? (Essentially, this is what happens in conglomerates.) The advantage is that the intermediary's investors must now bear less risk, so some of the higher risk and more junior claims can be transformed into lower risk and more senior claims. The disadvantage is, of course, that the projects funded will be exposed to more risk, which is itself costly in terms of the information transmission required.

Because most bank assets are relatively low risk, banks in turn are able to raise financing primarily from issuing very low-risk claims. Consequently, there is little to gain by reducing the risk of the relatively small number of high-risk claims. In other words, it is precisely because banks finance low-risk investments that it is efficient to have bank investors absorb the cost of low portfolio realizations.

In contrast, conglomerate assets are generally much riskier—and so in comparison to banks a larger fraction of their financing is derived from equity and

risky debt, and a lower fraction is derived from low-risk debt.<sup>7</sup> So imposing more risk on conglomerate divisions will be worthwhile: There are plenty of high-risk claims issued by the conglomerate for which the consequent reduction in risk will be beneficial.

The above argument accounts for the links between the type of project financed (III), the type of claims issued by the intermediary (IV), and the extent to which intermediary risk is borne by the projects financed (I). We now turn to the question of whether intermediation should be carried out by a specialist institution, or by a party who in any case needs to raise funding for itself (II).

Here, the paper's clearest predictions all relate to the case in which the projects funded are shielded from aggregate intermediary risk. As we argued above, arrangements of this sort only arise when the intermediary is able to finance itself without issuing high-risk claims. This in turn implies that the marginal claim issued by the intermediary is lower risk than the marginal claim issued by each project. So by acting as an intermediary, a project owner is able to reduce the risk of the claims he issues, leading to an increase in overall efficiency. As discussed, early forms of banking and trade credit arrangements are leading examples of this kind of arrangement. However, the prediction is overturned when both specialization in information production is possible and intermediary-issued claims are widely held—two conditions that are consistent with the rise of modern banks.

Although the model is couched in terms of information transmission being costly, the key elements needed for the results are that introducing contingencies into transfers between economic agents is costly, with the costs being higher when contingencies are invoked more frequently, and when more bilateral transfers are made contingent. Models of costly enforcement, costly collateral seizure, and costly renegotiation would all share these broad features, and so would lead to similar results (although of course the details of the arguments would differ).

Along with the institutional predictions discussed above, the paper also makes a more technical point. Most previous theoretical papers that have dealt with financial intermediation have focused on a relatively simple form in which intermediary borrowers repay the intermediary, which in turn repays its investors. In this paper, I consider a wider range of possible financial arrangements. In particular, I allow borrowers to hold offsetting claims in the intermediary, which gives rise to a form of joint liability among borrowers. The paper shows that while there are some parameter values for which standard intermediation arrangements remain optimal within this larger class, there are others for which intermediation with a degree of joint liability is strictly preferred.

<sup>7</sup> Note that under the case of close-to-perfect diversification discussed in the existing literature, this relationship does not hold: Claims on the intermediary will be close-to-riskless independent of the properties of the projects financed.

The current paper is clearly closely related to previous work on conglomerates, trade credit, and banking. In Section V, I discuss representative contributions to the study of each of these three institutions. Formally, the model developed is most closely related to the strand of the financial intermediation literature that has accounted for intermediation as a form of delegated monitoring. Diamond (1984), Williamson (1986), Krasa and Villamil (1992a), and Hellwig (2000) all fall within this class. As discussed in detail in Section III, the current paper differs in that it establishes the viability of intermediation without assuming that the probability of intermediary default is arbitrarily low. Aside from being of some interest in its own right, this property of the model is important because it allows us to address questions of the allocation of risk and the identity of the intermediary.<sup>8</sup>

The model employed is essentially a multiagent generalization of Townsend's (1979) costly state verification model. That is, an agent's output is private information unless a verification cost is incurred to disclose it to another agent. Krasa and Villamil (1992a) use a multiagent model of this sort to demonstrate that intermediation will emerge whenever the probability of intermediary default is close enough to zero, or equivalently whenever the degree of diversification possible is sufficiently high. However, because intermediary default is essentially nonexistent in this case, questions (I) to (IV) are hard to address. As discussed in Section III, the need to let the probability of default approach zero stems from assuming that all intermediary investors hold the same type of claim. In contrast, this paper uses Winton's (1995a) analysis of optimal seniority in a costly state verification setting to demonstrate that changes in an intermediary's income process that lead to second-order stochastic dominance will reduce the costs of monitoring the intermediary. This result is then enough to show that even with only two projects in the economy (i.e., very limited diversification and no way to eliminate intermediary default risk), the benefits of intermediation outweigh the costs. The default-prone intermediary can then be analyzed to study the consequences of different institutional responses to questions (I) to (IV).

The paper is organized as follows. Section I specifies the economic environment to be analyzed. Section II replicates Winton's (1995a) results on seniority

<sup>8</sup> Cerasi and Daltung (2000) and Krasa and Villamil (1992b) present models of intermediaries as delegated monitors in which perfect diversification is not possible. Both papers assume that the per-depositor costs of monitoring a bank are increasing in bank size, and so there is a size at which the increase in diversification provided by a larger bank is outweighed by the increase in monitoring costs. That is, there is an optimal bank size, and at that size there is a positive risk of failure. However, in order to establish the viability of intermediation, both papers must assume that the optimal bank size is large enough and that the corresponding probability of bank failure is close enough to zero. Moreover, both papers concentrate on the size of the bank; in contrast, the current paper explores the determinants of other properties of the financial intermediary. Finally, also related is the work of Winton (1995b). He establishes that with free entry into banking, there exist equilibria in which multiple banks exist, and each is of finite size with a positive probability of default. Again, the focus of the current paper is on the distinct issues of which agents intermediate, and whether or not entrepreneurs absorb intermediary risk.

in the context of the current model, and derives the result that second-order stochastic dominance is associated with a lowering of monitoring costs. Section III establishes the existence of financial intermediaries when only partial diversification is possible. Section IV establishes results concerning the optimal form of intermediation, which Section V then applies to derive predictions concerning conglomerates, banks, and trade credit arrangements. Section VI concludes.

## I. The Model

### A. The Agents

To keep the model as transparent as possible, we consider an economy with just two projects, where these projects represent the only sources of uncertainty. The projects 1 and 2 are run by entrepreneurs 1 and 2, respectively. We will typically use  $h$  to represent a generic project/entrepreneur. Each of projects  $h = 1, 2$  has a probability  $q_h \geq 1/2$  of “succeeding” and returning an amount  $H > 0$ , and a probability  $1 - q_h$  of “failing” and returning  $L \in [0, H)$ .<sup>9</sup> We write  $\omega_h$  for the random variable corresponding to the output of project  $h$ , and  $\bar{\omega}_h$  for its mean. The outcomes of the two projects are potentially correlated. The entrepreneurs have no income outside of the project returns.

In addition to the two entrepreneurs, there are  $2n$  investors, with typical member  $i$ . A total of  $n$  investors are needed to provide financing for each of the two entrepreneurs’ projects. As compensation for financing the entrepreneurs, each investor requires an expected payoff of  $\rho_n \equiv \rho/n$ —that is,  $\rho_n$  is the product of the funds provided by each investor,  $1/n$ , and the market interest rate,  $\rho$ . Investors have no income other than what the entrepreneurs transfer to them ( $\omega_i = 0$  for all investors  $i$ ).

We write  $I$  for the set of the  $2n$  investors, and  $K = I \cup \{1, 2\}$  for the set of all agents in the economy, with  $j, k$  generic agents. We make the following assumptions throughout:

**ASSUMPTION 1** (Both projects profitable): *Both projects are profitable in the absence of any financing frictions, that is, for  $h = 1, 2$ ,  $\bar{\omega}_h > \rho$ .*

**ASSUMPTION 2** (Both projects essential): *Some portion of the output from the high payoff of both entrepreneurs is needed in order to provide a payoff of  $\rho_n$  to each of the  $2n$  investors, that is,  $2\rho > \max\{\bar{\omega}_1 + L, \bar{\omega}_2 + L\}$ .*

The realization of each entrepreneur’s payoff  $\omega_h$  is privately observed by that entrepreneur. However, each entrepreneur  $h$  can disclose the realization of  $\omega_h$  to any second individual  $k \in K \setminus \{h\}$  at an effort cost  $c$ . Additionally, any agent  $k$  can disclose to any other agent  $j$  information he has previously acquired from the entrepreneurs 1 and 2. The cost of these “disclosures of disclosures” is also

<sup>9</sup> The assumption that  $q_h \geq 1/2$  is not essential, but simplifies the analysis.

$c$ .<sup>10</sup> Thus, the basic information structure is that of a generalized costly state verification model. In addition to disclosing endowment realizations, agents can also disclose information acquired from prior disclosures.<sup>11</sup>

All agents are risk neutral over nonnegative amounts of consumption  $x$ , and over nonnegative effort exertion  $e$ . That is, preferences are given by  $u(x, e) = x - e$ . Note that it is this limited liability constraint on consumption that makes the risk allocation problem nontrivial.

*B. Timing and Contracts*

There are three dates, labeled  $s = 0, 1, 2$ . The timing is as follows:

- $s = 0$ : Agents write contracts  $t$  (see below). The entrepreneur payoffs  $\omega_1, \omega_2$  are realized (after the contracts have been written).
- $s = 1$ : Each entrepreneur  $h = 1, 2$  can disclose his project realization to any subset of other agents. Following the disclosure of this information, transfers are made as contractually specified.
- $s = 2$ : Each agent  $j \in K$  can disclose to any subset of other agents the disclosures he received at date 1. Following these disclosures, further transfers are made, again as contractually specified.

All contracts are bilateral, and specify payments as follows. At each of dates 1 and 2, the transfer made between agent  $j$  and agent  $k$  can depend only on information that both possess—that is, either on information that agent  $j$  has disclosed to agent  $k$ , or vice versa. Thus the portion of the contract relating to the date 1 payment from agent  $j$  to agent  $k$  is just

$$t_{jk}^1(d_{jk}^1, d_{kj}^1), \tag{1}$$

where  $d_{jk}^1$  is the disclosure made by agent  $j$  to agent  $k$  at date 1. Similarly, at date 2 the transfer to be made from agent  $j$  to agent  $k$  is specified by

$$t_{jk}^2(d_{jk}^1, d_{kj}^1, d_{jk}^2, d_{kj}^2). \tag{2}$$

We obviously impose that at both dates  $s = 1, 2$ ,

$$t_{jk}^s = -t_{kj}^s. \tag{3}$$

<sup>10</sup> It would obviously be straightforward to generalize the analysis to the case in which the cost of disclosing disclosures differed from  $c$ . The implications of the analysis would be qualitatively unaffected.

<sup>11</sup> As in Townsend (1979), the verification cost is borne by the agent disclosing the information. Note that with two rounds of information sharing, it is easier to think of the verification decision as being made by the verified agent rather than by the verifying agent. Doing so avoids the complexity of modeling the degree to which a date 2 verification policy can depend on information possessed by the verifying agent. It is for this reason that we will refer to “disclosure” in place of “verification” throughout.

If agent  $j$  does not disclose to agent  $k$  at date  $s$ , we write  $d_{jk}^s = \emptyset$ . At date 1, only the two entrepreneurs 1, 2 have anything to disclose—so  $d_{jk}^1 = \emptyset$  if  $j \in I$ , while  $d_{hk}^1 \in \{\emptyset, \omega_h\}$  for  $h = 1, 2$ . At date 2, disclosures are made as to the vector of disclosures received at date 1. Thus  $d_{jk}^2 \in \{\emptyset, (d_{1j}^1, d_{2j}^1)\}$ .<sup>12</sup> Notationally, to capture the possibility of an entrepreneur disclosing his own endowment at date 2, we write  $d_{hh}^1 = \omega_h$ .

The set of bilateral contracts  $t \equiv \{t_{jk}^s : s = 1, 2 \text{ and } j, k \in K\}$  defines a game in which actions are disclosures. Each agent is restricted to choose from among strategies that give him nonnegative consumption,<sup>13</sup> independent of other agents' strategies.<sup>14</sup> Any pure-strategy equilibrium of this game induces a mapping from the state space  $\Omega$  to the transfers and disclosures:

$$\delta_{jk}^1 : \Omega \rightarrow \mathfrak{R} \cup \{\emptyset\}, \tag{4}$$

$$\delta_{jk}^2 : \Omega \rightarrow (\mathfrak{R} \cup \{\emptyset\})^2 \cup \{\emptyset\}, \tag{5}$$

$$\tau_{jk}^s : \Omega \rightarrow \mathfrak{R}. \tag{6}$$

We will refer to any particular set of mappings  $(\delta, \tau) \equiv \{\delta_{jk}^s, \tau_{jk}^s : j, k \in K, s = 1, 2\}$  as an *arrangement*. We say that an arrangement  $(\delta, \tau)$  is *incentive compatible* if the mappings  $\{\delta_{jk}^s, \tau_{jk}^s : j, k \in K, s = 1, 2\}$  arise as a pure-strategy equilibrium given contracts  $t$ .

Let  $\gamma_j(\omega; \delta, \tau)$  denote the total disclosure costs of agent  $j$  in state  $\omega$  under an arrangement  $(\delta, \tau)$ , that is,

$$\gamma_j(\omega; \delta, \tau) \equiv c \sum_{s=1,2} \sum_{k \in K \setminus \{j\}} \mathbf{1}_{\delta_{jk}^s(\omega) \neq \emptyset}(\omega), \tag{7}$$

where  $\mathbf{1}_{\delta_{jk}^s \neq \emptyset}(\omega)$  is the indicator function taking the value 1 whenever  $\delta_{jk}^s(\omega) \neq \emptyset$  and 0 otherwise. Let  $y_j^s(\omega; \delta, \tau)$  denote the resources of agent  $j$  at the end of period  $s$  in state  $\omega$ , that is,

$$y_j^s(\omega; \delta, \tau) \equiv \omega_j + \sum_{\bar{s}=1}^s \sum_{k \in K \setminus \{j\}} \tau_{kj}^{\bar{s}}(\omega). \tag{8}$$

So the utility  $u_j(\omega; \delta, \tau)$  of agent  $j$  in state  $\omega$  under arrangement  $(\delta, \tau)$  is simply

$$u_j(\omega; \delta, \tau) \equiv y_j^2(\omega; \delta, \tau) - \gamma_j(\omega; \delta, \tau). \tag{9}$$

<sup>12</sup> Allowing an agent to disclose only one of the disclosures, for example,  $d_{1j}^1$  and not  $d_{2j}^1$ , would have no qualitative effect on the results.

<sup>13</sup> This restriction is the two-period generalization of the assumption in the costly state verification literature that an agent cannot report an income of  $\tilde{\omega}$  that leads to no verification, but that triggers a required transfer in excess of his true income  $\omega$ . That is, there is an implicit assumption that there exists some central authority with enforcement capabilities that can punish an agent enough to deter this kind of behavior. Note that this central authority is required to act only out-of-equilibrium.

<sup>14</sup> We restrict attention to contracts  $t$  that possess such strategies.

Finally, let  $U_j(\delta, \tau)$  be the expected utility of agent  $j$  under arrangement  $(\delta, \tau)$ ,

$$U_j(\delta, \tau) \equiv E[u_j(\omega; \delta, \tau)]. \tag{10}$$

In the analysis that follows, we will explore the properties of constrained efficient incentive compatible arrangements. We are interested in arrangements that maximize the entrepreneurs' payoffs while delivering the market rate of return to the investors, that is,

$$U_i(\delta, \tau) \geq \rho_n \quad \text{for all } i \in I. \tag{I-IR}$$

The entrepreneur participation constraints are

$$U_h(\delta, \tau) \geq 0 \quad \text{for } h = 1, 2. \tag{E-IR}$$

Consider an arrangement  $(\delta, \tau)$  that satisfies both the investor (I-IR) and entrepreneur participation constraints (E-IR). We say that an arrangement  $(\hat{\delta}, \hat{\tau})$  *dominates*  $(\delta, \tau)$  if it gives (weakly) higher utility to both entrepreneurs and satisfies the investor participation constraints (I-IR).<sup>15</sup> Moreover, we will say that  $(\hat{\delta}, \hat{\tau})$  *strictly dominates*  $(\delta, \tau)$  if it dominates  $(\delta, \tau)$  and either strictly increases the utility of one of the entrepreneurs, or weakly increases the utility of all investors while strictly increasing the utility of at least one of them. An arrangement is *undominated* whenever it is not strictly dominated.

### C. Informational Insiders

The class of possible arrangements is very large. As we will see, a useful property of the arrangements to keep track of is the number of agents who pool information from multiple sources. Because of their privileged information, we refer to such agents as (informational) *insiders*. Formally, given an arrangement  $(\delta, \tau)$ , we will say that an agent is an insider either if he receives disclosures from at least two other agents, or if he is an entrepreneur and receives a disclosure from one other agent. That is, agent  $j$  is an insider either if  $j \in I$  and  $\exists \omega, \omega' \in \Omega, s, s' \in \{1, 2\}$  and  $k \neq l \in K \setminus \{j\}$  such that  $\delta_{kj}^s(\omega) \neq \emptyset$  and  $\delta_{lj}^{s'}(\omega') \neq \emptyset$ ; or if  $j \in \{1, 2\}$  and  $\exists \omega \in \Omega, s \in \{1, 2\}$  and  $k \in K \setminus \{j\}$  such that  $\delta_{kj}^s \neq \emptyset$ . Any agent who is not an insider is an *outsider*.

## II. Disclosure to Multiple Investors

As in Diamond (1984), intermediation of financial arrangements in the current setting lets an entrepreneur avoid disclosing to multiple agents (i.e., avoid duplication in monitoring), but introduces the delegation problem of keeping the intermediary honest. Diversification is the key to establishing that the former effect dominates, and overall disclosure costs are lower under intermediation. Previous research has focused on the advantages of almost perfect

<sup>15</sup> Note that this definition of domination is implied by, but does not imply, Pareto domination.

diversification (see the introduction): In this case the intermediary's income-per-investor is close to nonstochastic, so the intermediary is basically left with no information to misrepresent. In contrast, intermediation in the current paper depends on the benefits of a much less extreme form of diversification, namely the shift from financing one project to financing both. As we will see, the consequent reduction in the variance of the intermediary's income allows for the transformation of some of the more junior investor claims on the intermediary into more senior claims. Thus even a marginal increase in diversification leads to a reduction in delegation costs, which is enough to establish the viability of poorly diversified intermediaries.

Because the seniority structure of investor claims on the intermediary is central to this argument, we start by analyzing the seniority structure that arises when a single agent  $k$  discloses to some set  $J$  of outsider investors. This is the extension of the costly state verification problem studied by Winton (1995a). As is well known, with a single investor the optimal contract is debt-like, in the sense of involving costly verification (here, disclosure) only over some lower interval of the entrepreneur's income realization (see Townsend (1979) and Gale and Hellwig (1985)). Winton established that this property continues to obtain with multiple investors. Moreover, he showed that the optimal contract will feature multiple levels of seniority (in the sense that verification regions of the investors can be ordered), and that when all agents in question are risk neutral with limited liability, there will be as many seniority levels as there are investors. In this section I first map some of Winton's key results into the framework of the current paper, and then apply these results to quantify the size of each seniority class.

For the purposes of this paper, we need to be able to characterize the total expected disclosure costs of one individual  $k$  disclosing to a subset of investors  $J$  in the following two cases: (a) an entrepreneur disclosing directly to investors  $J$  and (b) an "intermediary," who could be either an investor or one of the entrepreneurs, and who receives transfers from the entrepreneurs and then in turn discloses to investors  $J$ . For this characterization we need to isolate the component of agent  $k$ 's income process that he either consumes (i.e.,  $y_k^2$ ), or transfers to the investors  $J$  (i.e.,  $\sum_{s=1,2} \sum_{i \in J} \tau_{ki}^s$ ). We denote this quantity by  $T_{k,J}(\omega; \delta, \tau)$ ,

$$T_{k,J}(\omega; \delta, \tau) \equiv y_k^2(\omega; \delta, \tau) + \sum_{s=1,2} \sum_{i \in J} \tau_{ki}^s(\omega). \quad (11)$$

All income in the economy originates with one of the two entrepreneurs,  $h = 1, 2$ . As such, the disclosing agent  $k$  will in general have the most resources available when both entrepreneurs succeed (state  $HH$ ) and the least available when both fail (state  $LL$ ), with the one-success-one-failure states  $LH, HL$  falling somewhere in between. All arrangements that we need to analyze in this paper do in fact satisfy this resource ordering across states. Moreover, since we can always change the naming of the two entrepreneurs, we can without loss assume

that the resource mapping  $T_{k,J}$  takes a higher value in state  $HL$  than  $LH$ . Formally, for the remainder of this section we assume:

$$T_{k,J}(LL; \delta, \tau) \leq T_{k,J}(LH; \delta, \tau) \leq T_{k,J}(HL; \delta, \tau) \leq T_{k,J}(HH; \delta, \tau). \quad (12)$$

Whenever inequality (12) holds, it will be useful to refer to state  $LL$  as being lower than state  $LH$ , which in turn we will refer to as lower than state  $HL$ , which in turn is lower than state  $HH$ .

As is standard, we will say that the subarrangement between agent  $k$  and the investors  $J$  is *debt-like* if agent  $k$  only discloses to each investor when his available income (given here by the mapping  $T_{k,J}$ ) falls below some critical level, and moreover does not disclose in any state in which the investor receives his maximal transfer.<sup>16</sup> Formally, we have the following definition.

**DEFINITION 1 (Debt-like):** *The subarrangement of  $(\delta, \tau)$  between agent  $k$  and investors  $J$  is said to be debt-like if for each  $i \in J$  the subset of states in which agent  $k$  discloses to investor  $i$  is one of  $\emptyset, \{LL\}, \{LL, LH\}, \{LL, LH, HL\}$ , and moreover agent  $k$  does not disclose to investor  $i$  in any state  $\omega \in \arg \max_{\tilde{\omega} \in \Omega} (\tau_{ki}^1(\tilde{\omega}) + \tau_{ki}^2(\tilde{\omega}))$ .*

When a subarrangement between agent  $k$  and investors  $J$  is debt-like, it is natural to speak of an investor  $i \in J$  as being *senior* to a second investor  $j \in J$  if investor  $i$  receives his maximal payment in strictly more states than investor  $j$ , or equivalently, if investor  $i$  is disclosed to in strictly fewer states than investor  $j$ . Given the resource ordering (12), the four possible *seniority classes* for the investor  $J$  are as follows. First, we have the most senior group  $N_{k,J}^\emptyset(\delta, \tau)$ , who agent  $k$  never discloses to. Second, we have the next most senior group  $N_{k,J}^{LL}(\delta, \tau)$ , who agent  $k$  discloses to only in the fail–fail state  $LL$ . The next in terms of seniority is the group  $N_{k,J}^{LL,LH}(\delta, \tau)$ , who agent  $k$  discloses to whenever entrepreneur 1 fails (i.e., states  $LL$  and  $LH$ ). Finally, the most junior group is  $N_{k,J}^{-HH}(\delta, \tau)$ , who agent  $k$  discloses to in all states *other than* the success–success state  $HH$ .<sup>17</sup>

What can we say about the size of these seniority classes? Agent  $k$  must be transferring a constant amount to investors in the most senior class  $N_{k,J}^\emptyset$ , since he never discloses to these investors. Moreover, the constant payments must be at least  $\rho_n$ , the amount investors demand in expectation. The expected resources agent  $k$  possesses to make these constant payments is simply  $T_{k,J}(LL)$ . So there can be at most  $\lceil T_{k,J}(LL)/\rho_n \rceil$  investors to whom agent  $k$  never discloses, where for the remainder of the paper  $\lfloor x \rfloor$  will be used to denote the largest integer weakly less than  $x$ .<sup>18</sup>

<sup>16</sup> It is common to speak of the constant maximal payment received in nondisclosure states as the “face value” of debt.

<sup>17</sup> Note that when the subarrangement between agent  $k$  and investors  $J$  is debt-like, there is never any disclosure in states in which the resource mapping  $T_{k,J}$  obtains its maximal value. So agent  $k$  will never disclose to any member of  $J$  in state  $HH$ .

<sup>18</sup> Observe that for any  $x, y, \lambda \in \mathfrak{N}_+$ , the following hold:  $\lfloor x \rfloor \in (x - 1, x]$ ,  $\lfloor x \rfloor + \lfloor y \rfloor \leq \lfloor x + y \rfloor$ ,  $\lfloor x - y \rfloor \leq \lfloor x \rfloor - \lfloor y \rfloor \leq \lfloor x - y + 1 \rfloor$ , and  $\lfloor \lambda x \rfloor \leq \lambda \lfloor x \rfloor$ .

For investors in the next seniority class  $N_{k,J}^{LL}$ , the transfer from agent  $k$  must be constant over the states  $LH, HL$ , and  $HH$ . So the aggregate transfer received by members of the two most seniority classes  $N_{k,J}^\emptyset$  and  $N_{k,J}^{LL}$  can be no more than

$$\Pr(LL) T_{k,J}(LL) + \Pr(LH, HL, HH) T_{k,J}(LH). \tag{13}$$

This expression corresponds to the expected resources agent  $k$  has available when he is restricted to access  $T_{k,J}(LH)$  or less in all states other than the state in which he discloses,  $LL$ .

Continuing in this manner implies that the size of the four seniority classes,  $N_{k,J}^\emptyset, N_{k,J}^{LL}(\delta, \tau), N_{k,J}^{LL,LH}$ , and  $N_{k,J}^{-HH}$ , must satisfy the following three inequalities:<sup>19</sup>

$$|N_{k,J}^\emptyset| \leq \min \left\{ |J|, \left\lceil \frac{T_{k,J}(LL)}{\rho_n} \right\rceil \right\} \tag{14}$$

$$\begin{aligned} &|N_{k,J}^\emptyset \cup N_{k,J}^{LL}| \\ &\leq \min \left\{ |J|, \left\lceil \frac{\Pr(LL)T_{k,J}(LL) + \Pr(LH, HL, HH)T_{k,J}(LH)}{\rho_n} \right\rceil \right\} \end{aligned} \tag{15}$$

$$\begin{aligned} &|N_{k,J}^\emptyset \cup N_{k,J}^{LL} \cup N_{k,J}^{LL,HL}| \\ &\leq \min \left\{ |J|, \left\lceil \frac{\Pr(LL)T_{k,J}(LL) + \Pr(LH)T_{k,J}(LH) + \Pr(HL, HH)T_{k,J}(HL)}{\rho_n} \right\rceil \right\}. \end{aligned} \tag{16}$$

Note for use below that the right-hand sides of the inequalities (14) through (16) are of the form

$$\min \left\{ |J|, \left\lceil \frac{1}{\rho_n} E_\omega[\min\{T_{k,J}(\omega), T_{k,J}(\zeta)\}] \right\rceil \right\} \tag{17}$$

for  $\zeta = LL, LH, HL$  respectively.

To give a corresponding lower bound for the size of the seniority classes we need to know more about the relationship between the agent  $k$  and the investors  $J$ . To this end, we define the following three additional properties that the subarrangement between these agents may possess. The first two are straightforward.

**DEFINITION 2 (Absolute priority):** *The subarrangement of  $(\delta, \tau)$  between agent  $k$  and investors  $J$  is said to feature absolute priority if it is debt-like, and whenever agent  $k$  discloses to agents in one seniority class in some state  $\omega$  then any agent  $i \in J$  who belongs to a more junior seniority class receives a zero transfer in that state, that is,  $\tau_{ki}^1(\omega) + \tau_{ki}^2(\omega) = 0$ .*

<sup>19</sup> Of course,  $|N_{k,J}^\emptyset \cup N_{k,J}^{LL} \cup N_{k,J}^{LL,HL} \cup N_{k,J}^{-HH}| = |J|$ .

Absolute priority implies that an investor who is junior to another investor receives no consumption (at least not from agent  $k$ ) whenever the more senior investor is disclosed to. In a similar vein, agent  $k$  is effectively *junior* to all the investors  $J$  if he himself does not receive any consumption in states in which he discloses:

**DEFINITION 3 (Agent  $k$  junior):** *The subarrangement of  $(\delta, \tau)$  between agent  $k$  and investors  $J$  is said to make agent  $k$  junior (to the investors  $J$ ) if agent  $k$  has zero consumption in any state  $\omega \in \Omega$  in which he discloses to at least one of the investors in  $J$ .*

Our third property corresponds to Winton's (1995a) Corollary 3, which states that in a continuous state-space setting there are as many seniority classes as investors. We term this property *maximal use of seniority*. Since in our discrete state space there can be at most four seniority classes, this property will clearly not hold in the same form here. Instead, we will say that maximal use of seniority holds if in any state, agent  $k$  concentrates all his transfers to disclosees in  $J$  to just one of these investors. Moreover, if this "preferred" investor is disclosed to in several other states, he should be the preferred investor in these states also. By maximizing the transfers to this preferred investor in disclosure states, agent  $k$  can lower the transfer the preferred investor receives in nondisclosure states, in turn potentially leading to an increase in the seniority of one of the other investors. Finally, the participation constraints of all investors in  $J$  should hold at equality—again, this frees up the resources to increase the seniority of all investors as much as possible. Formally,

**DEFINITION 4 (Maximal use of seniority):** *The subarrangement of  $(\delta, \tau)$  between agent  $k$  and investors  $J$  is said to make maximal use of seniority if the following conditions hold:*

- (1) *In any state  $\omega$ , there is at most one investor in  $J$  to whom agent  $k$  discloses and makes a strictly positive transfer.*
- (2) *Suppose agent  $k$  discloses to an investor  $i \in J$  in both states  $\omega$  and  $\omega'$ , where  $\omega$  is lower than  $\omega'$ . Then if the transfer from agent  $k$  to investor  $i$  is strictly positive in state  $\omega$ , it must be strictly positive in state  $\omega'$  also. That is, whenever state  $\omega$  is lower than state  $\omega'$ ,*

$$\delta_{ki}^2(\omega'), \delta_{ki}^2(\omega) \neq \emptyset \quad \text{and} \quad \tau_{ki}^1(\omega) + \tau_{ki}^2(\omega) > 0 \Rightarrow \tau_{ki}^1(\omega') + \tau_{ki}^2(\omega') > 0. \quad (18)$$

- (3) *Each investor  $i \in J$  receives exactly  $\rho_n$  in expectation.*

Our first result is then essentially a generalization of Winton's (1995a) analysis to a discrete state-space setting (although only for the case in which agents are risk neutral with limited liability), and establishes that the three properties just defined, plus debt-likeness, are optimal.

**PROPOSITION 1** (Basic properties): *Let us suppose an incentive compatible arrangement  $(\delta, \tau)$  that satisfies the investor participation constraints (I-IR) and involves an agent  $k$  disclosing to some set  $J$  of outsider investors, who themselves never disclose. Then there exists an incentive compatible arrangement  $(\hat{\delta}, \hat{\tau})$  that dominates  $(\delta, \tau)$  and in which the subarrangement between agent  $k$  and investors  $J$  is debt-like, features absolute priority, has agent  $k$  junior, and makes maximal use of seniority. Moreover, no agent discloses to an agent under  $(\hat{\delta}, \hat{\tau})$  to whom he did not disclose under  $(\delta, \tau)$ .*

*Proof:* The proof is omitted, but is available upon request from the author. The first part of the proof parallels that given in Winton (1995a); because attention is restricted to the case of risk neutrality with limited liability, the final step of establishing that the subarrangement is debt-like can be established more directly. Q.E.D

In light of Proposition 1, we make the following additional definition.

**DEFINITION 5** (Optimal seniority): *An incentive compatible arrangement  $(\delta, \tau)$  is said to feature optimal seniority between an agent  $k$  and a set of outsider investors  $J$  if the subarrangement between agent  $k$  and investors  $J$  is debt-like, features absolute priority, has agent  $k$  junior, and makes maximal use of seniority.*

This paper's main results stem from considering how the expected disclosure costs borne by some agent are affected by a change in the income process of that agent. The remainder of this section is devoted to showing conditions under which second-order stochastic dominance implies a reduction in expected disclosure costs. The first step is to apply Proposition 1 to complete our characterization of the size of the four seniority classes.

**LEMMA 1** (Number in each seniority class): *Let us suppose an incentive compatible arrangement  $(\delta, \tau)$  features optimal seniority between an agent  $k$  and a set of outsider investors  $J$ , and that inequality (12) holds. Then each of the three inequalities (14) to (16) holds at equality.*

*Proof:* See the Appendix.

The characterization of the size of the seniority classes that is provided by Lemma 1 is enough for us to establish the key result of this section, that is, a characterization of how disclosure costs change if we change the resource mapping  $T_{k,J} : \Omega \rightarrow \Re$  that determines the combined consumption of agent  $k$  and the total transfer to be made to outsider investors  $J$ .

Before proceeding to the formal result, consider the following simple numerical example in which agent  $k$  is entrepreneur 1 and is disclosing to three of the investors,  $J = \{i_1, i_2, i_3\}$ , say. Let the project success payoff be  $H = 120$  and the failure payoff be  $L = 0$ , with the probability of success equal to  $2/3$  and the two projects being stochastically independent. Since agent  $k$  is entrepreneur 1, he

succeeds in states  $\omega = HH, HL$  and fails in states  $\omega = LH, LL$ . Finally, let each investor's reservation utility be  $\rho_n = 20$ .

In this example, the resource mapping  $T_{k,J}$  that determines the resources available to agent  $k$  to consume and transfer to investors  $J$  is just  $T_{k,J}(LL, LH) = 0$  and  $T_{k,J}(LL, LH) = 120$ . So clearly the two most senior classes  $N_{k,J}^\emptyset$  and  $N_{k,J}^{LL}$  are empty. Intuitively, if agent  $k$  never discloses to an investor  $i$  in any state, then the investor knows it is possible that agent  $k$  has no resources and consequently must receive a zero transfer in all states. But this is inconsistent with meeting the investor individual rationality constraint (I-IR). It follows that all three of the investors in  $J$  must fall into the next seniority class,  $N_{k,J}^{LL,LH}$ . That is, agent  $k$  will disclose to them whenever his project fails (states  $LL, LH$ ). The total expected disclosure costs incurred by agent  $k$  are thus  $3c \Pr(LH, LL) = c$ .

Next, consider how expected disclosure costs change if we alter agent  $k$ 's project income so that his success payment is lowered to  $H = 110$ , while his failure payment is raised to  $L = 20$ . That is, even though agent  $k$ 's *expected* income is unchanged, he now has more resources available to consume and transfer to investors  $J$  in states  $LL$  and  $LH$ , while in states  $HL$  and  $HH$  he has fewer resources. The effect of this change in the distribution of  $T_{k,J}$  is that agent  $k$  is now able to transfer an amount 20 to one of the investors in every state, thus satisfying the individual rationality constraint (I-IR) for that investor. So we now have  $|N_{k,J}^\emptyset| = 1$ ,  $|N_{k,J}^{LL,LH}| = 2$ , with the other seniority classes being empty. Total disclosure costs from agent  $k$  to investors  $J$  are now  $2c \Pr(LL, LH) = 2c/3$ , that is,  $c/3$  less than before.

To keep this example as transparent as possible, we have directly changed agent  $k$ 's endowment process so as to reduce its variance, but the same effect could be achieved by altering the transfers received from other agents. The important point to note is that by changing agent  $k$ 's income process, or more generally his resource mapping  $T_{k,J}$ , in such a way that he has more resources in low resource states, but fewer resources in high resource states, we allow for some of the investors to become more senior. This leads to an overall reduction in disclosure costs. The following proposition generalizes these observations.

PROPOSITION 2 (Change of distribution): *Let  $(\delta, \tau)$  and  $(\hat{\delta}, \hat{\tau})$  be incentive compatible arrangements, both with optimal seniority between some agent  $k$  and some set of outsider investors  $J$ . Assume that inequality (12) is satisfied for both arrangements. Then agent  $k$ 's total disclosure costs to investor  $J$  are lower under  $(\hat{\delta}, \hat{\tau})$  than under  $(\delta, \tau)$  by an amount*

$$c(\Delta_{LL}\Pr(LL) + \Delta_{LH}\Pr(LH) + \Delta_{HL}\Pr(HL)), \tag{19}$$

where for  $\zeta = LL, LH, HL$

$$\begin{aligned} \Delta_\zeta = & \min \left\{ |J|, \left[ \frac{1}{\rho_n} E_\omega [\min\{T_{k,J}(\omega; \hat{\delta}, \hat{\tau}), T_{k,J}(\zeta; \hat{\delta}, \hat{\tau})\}] \right] \right\} \\ & - \min \left\{ |J|, \left[ \frac{1}{\rho_n} E_\omega [\min\{T_{k,J}(\omega; \delta, \tau), T_{k,J}(\zeta; \delta, \tau)\}] \right] \right\}. \end{aligned} \tag{20}$$

In particular, if for  $\zeta = LL, LH, HL$

$$E_\omega[\min\{T_{k,J}(\omega; \hat{\delta}, \hat{\tau}), T_{k,J}(\zeta; \hat{\delta}, \hat{\tau})\}] \geq E_\omega[\min\{T_{k,J}(\omega; \delta, \tau), T_{k,J}(\zeta; \delta, \tau)\}], \quad (21)$$

then the disclosure costs from agent  $k$  to the investors  $J$  are weakly lower under arrangement  $(\hat{\delta}, \hat{\tau})$ .

*Proof:* See the Appendix.

Note that condition (21) is implied by second-order stochastic dominance.

### III. Intermediation with the Risk of Default

In this section, we study arrangements involving intermediation, in the sense of there being an agent (either an investor or an entrepreneur) who acquires information about both the entrepreneurs' projects and then discloses this information to the remaining investors. Clearly, any form of intermediation needs at least one insider. So in more formal terms, this section is devoted to the study of arrangements featuring a single insider.

Figures 1 and 2 respectively display arrangements with no insiders, and with one of the entrepreneurs acting as an intermediary (i.e., one insider). For the purposes of exposition, in the main text we focus on the special case in which the

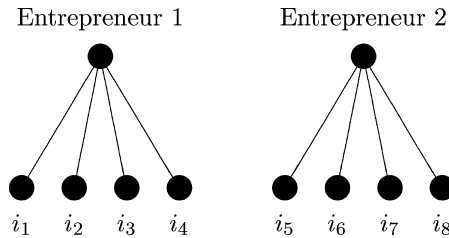


Figure 1. No insiders.

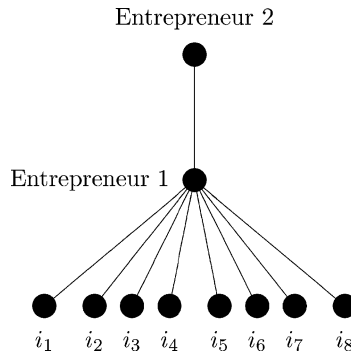


Figure 2. Simple intermediation by an entrepreneur.

two projects have the same probability of success (i.e.,  $q_1 = q_2$ ). Then proceeding somewhat loosely for the moment, disclosure costs without intermediation (Figure 1) are basically those that are incurred by a single entrepreneur with an income process  $2\omega_1$  who has to transfer an expected amount  $\rho_n$  to each of  $2n$  investors. That is, without intermediation (or more generally, without insiders), the fact that there are two projects in the economy is essentially irrelevant, and disclosure costs are the same as if we replaced the two entrepreneurs with a single entrepreneur twice the size of each.

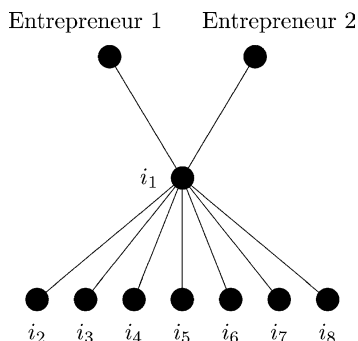
In contrast, the intermediary in Figure 2 will receive an income stream that is basically the sum of the two project realizations (i.e.,  $\omega_1 + \omega_2$ ). Diversification implies that the income stream  $\omega_1 + \omega_2$  second-order stochastically dominates (SOSD) the income stream  $2\omega_1$ . As Proposition 2 established though, second-order stochastic dominance implies a reduction in total disclosure costs. Moreover, the size of the reduction increases linearly in the number of investors  $n$  who are required to finance each entrepreneur. In contrast, the cost of introducing the intermediary is just the extra disclosure that the nonintermediary entrepreneur must now make to the intermediary. So for all  $n$  large enough intermediation will lead to a reduction in disclosure costs.

The viability of intermediation in this case stems from the fact that even a small amount of diversification allows for the transformation of some junior investor claims into more senior ones. It is natural to interpret the most senior claims as low-risk debt or bank deposits, while more junior claims would correspond to either risky debt or equity. Consequently, the framework predicts that a key characteristic of financial intermediaries is that they issue comparatively high levels of low-risk debt. Empirically, this is clearly true for commercial banks, while for conglomerates it is consistent with the general finding that cash flow volatility and leverage are negatively correlated (see, e.g., Harris and Raviv (1991, p. 334)).<sup>20</sup>

Proceeding more formally, we will describe an arrangement  $(\delta, \tau)$  as *simple intermediation by an entrepreneur* whenever one entrepreneur  $h$  discloses only to the other entrepreneur  $h'$  in period 1, with entrepreneur  $h$  neither making nor receiving any disclosures in period 2, and paying an amount  $R_h$  (respectively  $C_h$ ) to entrepreneur  $h'$  whenever  $\omega_h = H$  (respectively  $\omega_h = L$ ). If the contract between the intermediary and the entrepreneur is thought of as a debt contract, the success payment  $R_h$  corresponds to the face value of the loan, while the failure payment  $C_h$  is effectively the value of "collateral" that is recovered when the project fails.

Similarly, we will describe an arrangement  $(\delta, \tau)$  as *simple intermediation by an investor* whenever both entrepreneurs disclose only to a single investor,  $m \in I$ , in period 1, neither make nor receive any disclosures in period 2, and entrepreneur  $h = 1, 2$  pays  $R_h$  (respectively  $C_h$ ) when  $\omega_h = H$  (respectively  $\omega_h = L$ ). An arrangement featuring intermediation by an investor is illustrated graphically in Figure 3. For both types of intermediation, we will often make

<sup>20</sup> A caveat should be noted here: While the model predicts that intermediaries will issue more low-risk debt than stand-alone firms, they may issue *less* high-risk debt.



**Figure 3. Simple intermediation by an investor.**

reference to the *entrepreneur payments*  $(R_1, R_2, C_1, C_2)$ , where it is understood that if intermediation is by entrepreneur  $h$ , then  $R_h = H$  and  $C_h = L$  (i.e., an entrepreneur intermediary effectively transfers all his project income to himself).

The following proposition then formalizes the intuitive argument for the superiority of intermediation given above.

**PROPOSITION 3 (Intermediation):** *There exists an  $n^*$  such that provided  $n \geq n^*$ , the following is true: If  $(\delta, \tau)$  is an incentive compatible arrangement with no insiders that satisfies the investor participation constraints (I-IR), then there exists a simple intermediation by an entrepreneur arrangement  $(\hat{\delta}, \hat{\tau})$  that strictly dominates  $(\delta, \tau)$ . Moreover, under the arrangement  $(\hat{\delta}, \hat{\tau})$ , the entrepreneur payments satisfy  $C_h = L (h = 1, 2)$ , and the combined decrease in aggregate expected disclosure costs is at least  $\min_{\omega} \Pr(\omega)c$ .*

*Proof:* See the Appendix.

In the economy under consideration there is always a strictly positive probability that both entrepreneurs' projects will fail. Consequently, any intermediary must have a strictly positive risk of default, in the sense of needing to disclose to at least some investors that his income is low and they will be paid less than in other states. This feature makes the form of intermediation established by Proposition 3 fundamentally different from the forms studied by Diamond (1984) and Krasa and Villamil (1992a). In both these papers, an intermediary is shown to be viable only when it holds a portfolio that is arbitrarily well diversified and consequently the probability of defaulting on the investors is arbitrarily low. What accounts for this difference in results?

First, and in contrast to Diamond's model, the total cost of monitoring the intermediary to keep him honest is assumed to be increasing in the number of intermediary investors. Consequently the number of investors needed to finance each project ( $n$ ) will in general affect whether or not intermediation is viable. Second, and more importantly, this paper makes use of the fact that the

claims held by intermediary investors differ, both in practice (think of bank depositors, bond holders, and equity holders) and in theory (again, see Winton (1995a)). The heterogeneity of investor claims implies that changes in the intermediary's income distribution that lead to second-order stochastic dominance (Proposition 2) will tend to reduce the costs of monitoring the intermediary, since some of the relatively junior claims can be transformed into more senior claims. In contrast, Krasa and Villamil (1992a) also model monitoring costs as increasing in  $n$ , but do not allow for differentiation among investor claims. Under these assumptions, second-order stochastic dominance may actually increase rather than reduce monitoring costs, with the size of the increase growing in  $n$ .<sup>21</sup>

Finally, it is worth observing that the requirement of Proposition 3 that  $n$  be sufficiently high is not very stringent—it is required only to ensure that enough junior investors are made senior to compensate for the extra layer of agency associated with intermediation.

At this point we have tied the existence of a financial intermediary to the risk profile of claims it must issue to raise financing. This is the key property of the model that will allow us to link the type of projects financed by an intermediary to how its portfolio risk is allocated between investors and entrepreneurs, and to whether an investor or an entrepreneur (and if so, which one) should act as the intermediary. Sections IV and V take up these questions. But before proceeding, we conclude this section by observing that simple intermediation is not just better than any arrangement with no insiders, but is also at least weakly better than any other arrangement with a single insider.

Consider first an incentive compatible arrangement  $(\delta, \tau)$  with exactly one insider, where the insider is an investor. By definition he must receive disclosures from at least two other agents. On the one hand, if the insider is receiving information about both entrepreneurs, the expected cost must be at least  $(1 - q_1)c + (1 - q_2)c$ . But if we are going to incur these costs, we may as well have both entrepreneurs disclose to the insider at date 1, and channel all payments through the insider, that is, simple intermediation by an investor. On the other hand, if the insider is receiving information about only one of the entrepreneurs, then having an insider does not add anything at all.

Next, if the only insider of arrangement  $(\delta, \tau)$  is entrepreneur  $h$ , then either he must be receiving information about entrepreneur  $h'$ , in which case we may as well make entrepreneur  $h$  the intermediary and channel all payments through him, or else he is receiving information about himself from some investor, in which case this disclosure is pointless and are better off without any insiders.

<sup>21</sup> Consider the following simple example:  $L = 0, H = 5/2, q_1 = q_2 = 1/2, \rho = 1$ , independent projects. Since  $(1 - \Pr(HH))H = 15/8 < 2\rho$ , disclosure to investors is required in all states other than  $HH$ . But since Krasa and Villamil restrict all investors to belong to the same seniority class, this means that the intermediary discloses to all investors in every state except  $HH$ , and expected disclosure costs are  $2nc(1 - \Pr(HH)) = 3nc/2$ . In contrast, total disclosure costs under direct investment are just  $nc$ . Thus in this example, the fact that all investors are in the same seniority class implies that disclosure costs are always *increased* by intermediation.

Formally, we have the following lemma.

**LEMMA 2** (One insider  $\Rightarrow$  simple intermediation): *Let  $(\delta, \tau)$  be any arrangement with a single insider that satisfies the investor and entrepreneur participation constraints (I-IR) and (E-IR). Then  $(\delta, \tau)$  is dominated either by simple intermediation or by an arrangement with no insiders.*

*Proof:* The proof is omitted, but is available upon request from the author. The basic idea is straightforward: Whatever payments occur in the initial arrangement  $(\delta, \tau)$  are simply channeled through the single insider, who acts as the intermediary. The main difficulty lies in dealing with any investors who previously acted as “pseudo”-intermediaries, in the sense of receiving a disclosure from a single entrepreneur and then disclosing to some subset of other investors. In order to establish dominance by simple intermediation, the total transfers made to these pseudo-intermediaries must be decreased to offset their savings in disclosure costs. One then needs to check that the disclosure needed from the insider to the pseudo-intermediaries is no more than before. The formal proof takes care of this case. Q.E.D.

#### **IV. Reducing the Cost of Intermediary Default**

Above we saw that intermediation is preferable to unintermediated financial arrangements, even though intermediaries themselves default with a positive probability. We next proceed to consider three different ways in which the cost of intermediary default can be reduced.

First, are intermediary default costs lower when the collateral payments  $C_h$  are set as high as possible, or when the face value of debt  $R_h$  is set as high as possible? Second, is it worth incurring the costs associated with further information sharing in order to make the payments between the intermediary and each entrepreneur contingent on the outcome of the other entrepreneur’s project? Third, does an investor or an entrepreneur make the better intermediary?

##### *A. Payments from the Entrepreneurs*

First, holding the identity of the intermediary fixed for now, what is the optimal form of simple intermediation? Clearly the subarrangement between the intermediary and the investors should feature optimal seniority. Additionally, we have a choice to make as to whether an entrepreneur  $h$  who transfers resources to the intermediary should obtain most of his consumption in the failure state or the success state. On the one hand, concentrating consumption in the success state allows us to make the failure payment  $C_h$  relatively large, which helps to increase the number of investors the intermediary never has to disclose to. But on the other hand, concentrating consumption in the failure state allows the success payment  $R_h$  to be set at a high level, which potentially

decreases the number of investors the intermediary only needs to disclose to in the state where both entrepreneurs fail ( $\omega = LL$ ). Formally, we will say that entrepreneur payments  $(R_1, R_2, C_1, C_2)$  are *debt-like* if the entrepreneurs receive less consumption when their projects succeed than when they fail, that is, if  $R_h - C_h \leq H - L$  for  $h = 1, 2$ . The following result establishes that debt-like entrepreneur payments are indeed desirable—that is, the former of the above effects dominates and an entrepreneur’s consumption should be concentrated in the state in which his project succeeds.

LEMMA 3 (Simple intermediation): *Let  $(\delta, \tau)$  be a simple intermediation (by either an entrepreneur or an investor) arrangement with entrepreneur payments  $(R_1, R_2, C_1, C_2)$  and satisfying the investor participation constraints (I-IR). Then the arrangement  $(\delta, \tau)$  is dominated by a simple intermediation arrangement  $(\hat{\delta}, \hat{\tau})$  with optimal seniority between the intermediary and the investors, and debt-like entrepreneur payments  $(\hat{R}_1, \hat{R}_2, \hat{C}_1, \hat{C}_2)$ .*

*Proof:* See the Appendix.

### B. Additional Contingencies (Joint Liability)

Simple intermediation allocates all the burden of intermediary default to the investors, with the entrepreneurs being entirely unaffected. To see why having the investors absorb all the consumption risk of intermediary default may not be optimal, consider a simple intermediation arrangement in which the intermediary discloses to some investors not just in state  $LL$ <sup>22</sup> but also in state  $LH$ . That is, at least some of the claims issued by the intermediary to investors involve a risk of default that is at least as high as the probability of failure of entrepreneur 1.

The intermediary’s income in the four states  $LL, LH, HL$ , and  $HH$  is  $C_1 + C_2, C_1 + R_2, C_2 + R_1, R_1 + R_2$ . As far as the intermediary is concerned, he has too much income in state  $HH$ , but not enough income in state  $LH$ . If he could increase his state  $LH$  income, he could reduce the number of investors he has to disclose to in that state. He can achieve just such a change in the income distribution by increasing the success payment made by entrepreneur 2 by an amount  $b_2$ , while at the same time offering to pay that entrepreneur a “bonus” payment  $B_2$  in the case where both entrepreneurs succeed, that is, state  $HH$ . When the payments  $b_2$  and  $B_2$  are selected so as to leave the expected net transfer of the entrepreneur 2 unchanged, the effect is to replace the intermediary’s income process with one that SOSD it. By Proposition 2, this leads to a lowering of disclosure costs to the investors. This perturbation of simple intermediation amounts to introducing a degree of joint liability between the entrepreneurs: Entrepreneur 2’s final consumption is now linked to entrepreneur 1’s project realization. Several interpretations of this kind of arrangement are discussed in detail in Section V.

<sup>22</sup> Assumption 2 and the need to give each of  $2n$  investors an expected utility of  $\rho_n$  implies that disclosure to at least some investors in state  $LL$  is essential.

The payments  $b_2$  and  $B_2$  entail making the transfer between the intermediary and entrepreneur 2 contingent on the realization of entrepreneur 1's project. This is only possible if the intermediary now discloses to entrepreneur 2 whenever he succeeds but entrepreneur 1 fails. But when there are multiple investors per entrepreneur, this extra cost will be more than compensated for by the reduction in the intermediary's disclosure costs to the investors.

Increasing the intermediary's income in state  $LH$  is clearly beneficial only if the intermediary actually discloses to investors in that state. So intuitively, whether or not the above perturbation of simple intermediation actually improves matters depends on the probability of intermediary default. We will say that a simple intermediation arrangement  $(\delta, \tau)$  with optimal seniority is *low risk* if the intermediary only discloses to investors in state  $LL$ . Otherwise we say the simple intermediation arrangement is *high risk*.

Additionally, we will describe an arrangement  $(\delta, \tau)$  as *intermediation with joint liability* if it differs from a simple intermediation arrangement (with entrepreneur payments  $(R_1, R_2, C_1, C_2)$ ) only in that the intermediary discloses to at least one nonintermediary entrepreneur  $h$  whenever he has succeeded ( $\omega_h = H$ ) and the other entrepreneur  $h'$  has failed ( $\omega_{h'} = L$ ), while making a "bonus" payment  $B_h$  in state  $HH$ .

**PROPOSITION 4** (High-risk simple intermediation dominated): *There exists an  $n^*$  such that provided  $n \geq n^*$ , the following is true: Let  $(\delta, \tau)$  be a simple intermediation arrangement with optimal seniority, high-risk, debt-like entrepreneur payments, and satisfying the entrepreneur participation constraints (E-IR). Then  $(\delta, \tau)$  is dominated by an intermediation with joint-liability arrangement  $(\hat{\delta}, \hat{\tau})$ . Moreover, if under  $(\delta, \tau)$  the intermediary discloses to at least two investors in a state other than  $LL$ , the aggregate increase in welfare at least  $\min\{\Pr(LH), \Pr(HL)\}$ .*

*Proof:* See the Appendix.

Observe that in many joint-liability arrangements the entrepreneurs will actually be more junior than the most junior of the investors, who can be interpreted as holders of the intermediary's equity. That is, a jointly liable entrepreneur only receives his bonus payment when both entrepreneurs succeed, that is, state  $HH$ , while even the most junior investors may be paid in full in one of the states where only one of the projects succeeds, that is,  $LH$  or  $HL$ , as well as state  $HH$ .

### C. Choice of Intermediary

Among low-risk simple intermediation arrangements, is it better to have an investor or an entrepreneur be the intermediary? On the one hand, if an entrepreneur is the intermediary, then there are  $2n$  investors for the intermediary to deal with, but only one of the two entrepreneurs has to disclose to the

intermediary. On the other hand, if an investor is the intermediary, there are only  $2n - 1$  investors for him to deal with, but now both entrepreneurs must disclose to the intermediary. Entrepreneur disclosure occurs with probability  $\Pr(\omega_h = L)$ , while disclosure by a low-risk intermediary occurs with probability  $\Pr(LL) < \Pr(\omega_h = L)$ . It follows that disclosure costs are lower when an entrepreneur is the intermediary.

**LEMMA 4** (Low-risk simple intermediation by investor dominated): *There exists an  $n^*$  such that provided  $n \geq n^*$ , the following is true: Let  $(\delta, \tau)$  be a simple intermediation by an investor arrangement, with optimal seniority, low-risk, debt-like entrepreneur payments, and satisfying the entrepreneur participation constraints (E-IR). Then  $(\delta, \tau)$  is dominated (at least weakly) by a simple intermediation by an entrepreneur arrangement.*

*Proof:* See the Appendix.

## V. Interpretation of Results

As the previous section indicates, there are circumstances under which some form of intermediation of financial arrangement is optimal, but intermediaries do not resemble the modern banks most often discussed in the literature. From Proposition 4, in some circumstances it is preferable to have entrepreneurs absorb some of each other's risk (i.e., joint-liability intermediation). And as Lemma 4 implies, it is often best to have one of the entrepreneurs intermediate.

### A. Conglomerates

The contemporary institution that most closely resembles joint-liability intermediation is a conglomerate. Within a conglomerate, the headquarters are responsible for raising funds from capital markets and then disbursing them to the various divisions. It is well established that even within a diversified conglomerate, each division is affected by the performance of other divisions. In particular, empirical studies have shown that the investment of one division is related to the cash flow of the whole conglomerate (see, e.g., Lamont (1997) and Shin and Stulz (1998)). In the language of this paper, each division receives funds from the headquarters, transfers funds back to the headquarters when it performs well, and receives a bonus payment for future investment when other divisions also perform well. Note that in contrast to many other models, conglomerates emerge in the current framework even without assuming that conglomeration exogenously eases the frictions between agents within the conglomerate.

When does intermediation take this form? As discussed, the benefits of joint-liability intermediation over simple intermediation stem from the former arrangement's ability to transform junior investor claims into more senior ones. This advantage clearly only obtains when junior claims exist that can be

transformed, that is, when simple intermediation is high risk. And simple intermediation (with entrepreneur payments  $(R_1, R_2, C_1, C_2)$ ) will in turn be high risk if and only if

$$(C_1 + C_2)\Pr(LL) + \min\{C_1 + R_2, R_1 + C_2\}(1 - \Pr(LL)) < \rho_n \times (\# \text{ of nonintermediary investors}). \quad (22)$$

More formally, and in terms of the underlying parameters  $L, H$  and the probabilities  $\Pr(\omega)$ , we have:

COROLLARY 1 (Complex arrangements optimal): *Suppose that*

$$2L\Pr(LL) + (L + H)(1 - \Pr(LL)) < 2\rho. \quad (23)$$

*Then there exists an  $n^*$  such that provided  $n \geq n^*$ , the following is true: If  $(\delta, \tau)$  is an incentive compatible arrangement satisfying the participation constraints (I-IR) and (E-IR) and has one or no insiders, it is strictly dominated by an intermediation with joint-liability arrangement  $(\hat{\delta}, \hat{\tau})$ . Conversely, the arrangement  $(\hat{\delta}, \hat{\tau})$  is undominated by any arrangement with less than two insiders.*

*Proof:* See the Appendix.

Corollary 1 indicates that it is low-quality and/or high-risk projects that will be financed by conglomerate-like intermediation arrangements. That is, if the probability  $\Pr(LL)$  of both projects simultaneously failing is high, or if the failure payoff  $L$  is low, or if the success payoff  $H$  is low, then inequality (23) is more likely to hold. Moreover, note that when the projects are i.i.d., a mean-preserving spread<sup>23</sup> in the failure and success payoffs will reduce the left-hand side of inequality (23) and so raise the chances of it holding. In contrast, when projects are higher quality, then they may be financed by a simple intermediation arrangement (see Corollary 2).

Empirically, these predictions are consistent with a conglomerate discount: Conditional on observing a conglomerate we can infer that the underlying divisions are of lower quality than stand-alone projects financed by bank loans and/or trade credit. Graham et al. (2002) and Campa and Kedia (2002) provide evidence consistent with this prediction that it is firms with lower values that form conglomerates. In common with Fluck and Lynch (1999), this also implies that a merger announcement should lead to a positive share price response for the aggregate of the two merging firms (since conditional on the decision to merge conglomeration is more efficient). Likewise, a spin-off announcement should also lead to a positive price response, since the decision reveals that the project quality has risen enough for the financing arrangement to revert to simple intermediation. Empirical support for these predictions is discussed in

<sup>23</sup> That is, for the case where  $q_1 = q_2 = q$  and the projects are independent, consider decreasing the failure payoff  $L$  by  $q\varepsilon$  while increasing the success payoff  $H$  by  $(1 - q)\varepsilon$ , so that the expected output of each project is left unchanged. Then it is easily verified that the left-hand side of inequality (23) is reduced by  $\varepsilon\Pr(HH)$ .

detail by Fluck and Lynch. The main difference between this paper and theirs is that here the act of conglomeration does not eliminate the agency problems present in a direct financing arrangement. Rather, conglomeration emerges as a more efficient response to a common set of financing frictions.<sup>24</sup>

Although conglomerates and commercial banks are both forms of financial intermediation, for the most part these institutional arrangements have been analyzed entirely separately. Perhaps the most prominent exception is Gertner, Scharfstein, and Stein (1994), which explicitly compares conglomerates and banks. Similar to this paper, their emphasis is on the relative variability of the transfers made by conglomerate divisions to their headquarters, and the analogous transfers from bank borrowers to a bank. The main difference is that whereas the aforementioned authors stress the increase in investment productivity associated with the ex post reallocation of resources across divisions, the current paper focuses on the lowering of intermediary default risk that such transfers can engender. Put somewhat differently, the choice between extending a loan and taking an equity stake that is analyzed by Gertner, Scharfstein, and Stein would also be faced by a single large investor, while the trade-off analyzed here arises only for intermediaries acting on behalf of other investors.

### *B. Credit Crunches*

At first sight, the prediction that entrepreneurs funded by a financial intermediary will on occasion be called upon to absorb some of the intermediary's portfolio risk may not seem applicable in the context of the modern banking system. There is, however, one area of modern banking in which a phenomenon resembling the imposition of risk on bank borrowers is in fact observed: the so-called credit crunch.<sup>25</sup>

To see how the model can be used to account for credit crunch episodes, consider an extension of the model in which successful entrepreneurs subsequently have access to a second investment opportunity. If an investment of  $l$  is made, this second project returns  $sl$ , which is assumed to exceed the opportunity cost of funds. Assume moreover that this second investment can only be financed by the original intermediary.<sup>26</sup> Finally, we take as given that financing for this

<sup>24</sup> This argument is also related to that of Maksimovic and Phillips (2002), who argue that scarce managerial resources determine which firms join conglomerates.

<sup>25</sup> See, for example, Bernanke and Lown (1991). Existing explanations of this phenomenon typically take as given that banks must prioritize paying off their depositors above all other concerns. For example, explanations based on Myers' (1977) debt overhang effect assume that outstanding debt is noncontingent. Likewise, more recent explanations such as those of Holmström and Tirole (1997) and Diamond and Rajan (2000) assume that deposit claims are noncontingent and so are senior to the interests of bank borrowers. The emphasis here is instead on suggesting a model in which the same friction that leads to intermediation also accounts for senior deposit claims with limited contingencies. That is, at least in principle, it is easy to imagine that it would be optimal to shield bank borrowers from a decline in bank fortunes by having bank depositors absorb the shortfall in funds. The model of this paper explicitly allows for arrangements of this sort.

<sup>26</sup> For example, only the original intermediary knows if the second investment opportunity actually exists.

second project is provided in such a way that it gives both the entrepreneur and the intermediary a positive surplus. A credit crunch resembles the joint-liability intermediation arrangements discussed in this paper, in that the failure of lots of borrowers' projects may lead to a successful borrower's loss of financing for the second investment opportunity, and thus to a reduction in his welfare. However, there is also a potentially important distinction. Since the bank also loses some surplus from not financing the second investment project, no disclosure to the adversely affected entrepreneur is necessary to make this incentive compatible.<sup>27</sup> In contrast, in the pure joint-liability arrangements considered above, the intermediary must disclose in states *LH* and *HL* in order to justify reducing a successful entrepreneur's welfare.

In this setting, a close variant of the argument we have already made for the advantages of joint liability applies. In principle, the intermediary could make sure to have sufficient funds to finance the second round of investment opportunities, but only at the cost of having to default more often on the investors. This in turn would raise disclosure costs. Whenever the investors are sufficiently diffuse, these extra disclosure costs will exceed the extra surplus obtainable from funding the second round of entrepreneurial investments. And as discussed above, no disclosure from the intermediary to the entrepreneurs is necessary when refinancing does not take place, since the intermediary's welfare is also lower in such a case.

In addition to providing a potential explanation of the credit crunch phenomenon, the model presented also has predictions for which sectors of the banking system are likely to be the most afflicted. As discussed above, the joint-liability arrangements are most likely to be observed when the underlying projects are low quality and/or high risk. There is some evidence that this is indeed the case. First, bank-level studies such as that of Hancock and Wilcox (1998) suggest that the effects of bank capital on lending activity are most pronounced for small banks, the primary sources of funds for small businesses. Second, using individual loan data, Hubbard, Kuttner, and Palia (2002) find that the interest rate effect of poor bank performance is concentrated on small borrowers and borrowers without a bond rating.

### *C. Intermediation by Nonspecialists: The Origins of Banking, and Trade Credit*

The second main departure from modern banking that this paper predicts is that it is often efficient for intermediation to be carried out by an entrepreneur (as opposed to by an investor; see Lemma 4). Intermediation by an entrepreneur is potentially beneficial because it enables one of the entrepreneurs to raise finance on the terms available to a diversified intermediary. Whenever the claim held by the marginal investor in the intermediary is low risk, financing costs are then lowered.

<sup>27</sup> If financing is provided in such a way that the intermediary loses money from funding the second investment, then the standard version of the model is enough to account for the credit crunch effect.

Before turning to a discussion of how this result should be interpreted, we first tie up a loose end by establishing that there are indeed some circumstances under which intermediation by an entrepreneur is optimal (i.e., essentially the dual of Corollary 1). To state the result, we need one further definition. Let  $(\delta, \tau)$  be a simple intermediation arrangement with optimal seniority between the intermediary and the investors. Then we will say that the arrangement has *maximal nondisclosure* whenever there are  $\lfloor 2L/\rho_n \rfloor$  investors who are never disclosed to. Clearly any arrangement with entrepreneur payments satisfying  $C_h = L$  satisfies maximal nondisclosure, although the converse need not hold. We then have the following corollary.

**COROLLARY 2** (Simple intermediation optimal): *There exists an  $n^*$  such that provided  $n \geq n^*$ , the following is true: Let  $(\delta, \tau)$  be a simple intermediation by an entrepreneur arrangement, with optimal seniority between the intermediary and the investors, low risk, and maximal nondisclosure. Suppose, moreover, that  $\min_{\omega \in \Omega} \Pr(\omega) \geq q_m - q_{m'}$ , where  $m$  is the intermediary entrepreneur and  $m'$  is the other entrepreneur, that is, either the riskier entrepreneur is the intermediary, or the difference between the project success probabilities is not too great. Then if a second arrangement  $(\tilde{\delta}, \tilde{\tau})$  strictly dominates  $(\delta, \tau)$ , it must in turn be dominated by a simple intermediation by an entrepreneur arrangement  $(\hat{\delta}, \hat{\tau})$  with optimal seniority between the intermediary and the investors, low risk, and maximal nondisclosure.*

*Proof:* See the Appendix.

Banking services in the contemporary United States are provided by specialist financial institutions, in the sense that banks do not themselves have ownership stakes in productive real assets. But in many ways this is a relatively recent phenomenon. For example, Lamoreaux (1994) studies the development of banks in 19th-century New England. She finds that during this period many banks were owned and operated by entrepreneurs, to whom they then lent a sizeable fraction of their funds—a practice she terms “insider lending.” This form of intermediation resembles the simple intermediation arrangements discussed here.<sup>28</sup>

At least in the United States, the modern-day financing arrangement that most resembles simple intermediation by an entrepreneur is trade credit.<sup>29</sup> As the name implies, in trade credit relationships one firm raises financing surplus to its own needs, and in turn supplies financing to a second firm. Trade credit continues to be an important source of financing for many firms. Moreover, the

<sup>28</sup> Cerasi and Daltung (1998) also discuss business-owned banks. In their explanation, it is optimal for a firm to own the bank it gets financing from when the bank can be made sufficiently diversified to effectively leave the firm-bank the residual claimant on profits.

<sup>29</sup> A rough indication of the scale of trade credit in the United States can be obtained from the Federal Reserve’s Flow of Funds report. For example, in the third quarter of 2002, the total value of trade receivables owed to nonfinancial businesses was roughly twice the total value owed to banks by these same businesses.

provision of consumer credit to a firm's customers can itself be viewed as a special instance of trade credit.

Trade credit has a number of notable features. First and most obviously, the credit relationship exists between firms that also trade goods with each other. Second, credit flows in the same direction as the goods traded: The supplier of goods provides financing to the purchaser, and not the other way around (this would be prepayment). Third, trade credit is expensive (see, e.g., Petersen and Rajan (1997)).

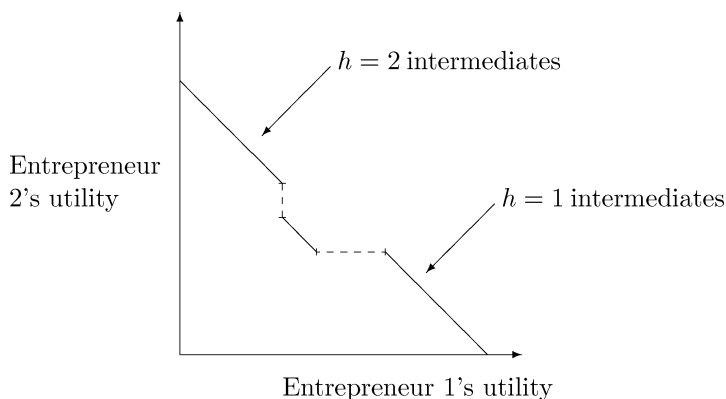
The model developed here can be easily adapted to provide an explanation of the first of these observations: In many circumstances, it would be realistic to assume that monitoring/disclosure costs are lower between trading partners.<sup>30</sup> In fact, the model can also be used to account for the other two observations as well.

The key point to note is that the model predicts that simple intermediation (which is how we are interpreting trade credit arrangements) will only be observed when it is low risk. From (22), it is clear that this in turn requires any entrepreneur  $h$  who is not an intermediary to make relatively large transfers,  $C_h$  and  $R_h$ , to the intermediary. So conditional on observing simple intermediation in the first place, we should expect to see borrowers paying a lot for their credit. That is, trade credit is expensive.

This same implication also potentially accounts for the direction of flow of trade credit. Consider two entrepreneurs who must somehow decide on a financing arrangement. This decision is *not* independent of the decision of how to split the surplus available. Specifically, delivering a high share of the surplus to entrepreneur 1 is not in general compatible with that entrepreneur being the borrower in a simple intermediation arrangement. The reason for this is that, as we saw above, the financing arrangement will only take the form of simple intermediation if it is low risk, and this requires entrepreneur 1 to make large transfers to the intermediary. Figure 4 graphically illustrates this point.

It follows that the entrepreneur with the stronger bargaining position will in general be the intermediating entrepreneur. So if the goods supplier is the more indispensable of the two parties, the model predicts that credit will flow from the goods supplier to the goods purchaser—exactly as is in fact observed. Petersen and Rajan (1997) report evidence consistent with this explanation: Firms with higher gross profit margins (which presumably reflect market power) provide more trade credit. (Also consistent is that in Lamoreaux's description of firm-owned banks in 19th-century New England, the owners of banks appear to be deriving substantial surplus from the arrangement—as is evidenced, for example, by the efforts that incumbent firm-owned banks devoted to preventing further entry into the banking sector.)

<sup>30</sup> See, for example, Biais and Gollier (1997) for an explanation of trade credit as stemming from a supplier's superior access to information. In a related vein, see, for example, Frank and Maksimovic (1998) for an account of trade credit based on the supplier's superior ability "to reclaim value from the repossessed good" (abstract). However, it is not clear whether either view can account for the fact that trade credit flows from (and not to) the supplier.



**Figure 4. Schematic Pareto frontier.** Allocations giving entrepreneur 1 a high share of the surplus are those in which he is the intermediary. Under some circumstances allocations in which the surplus is divided evenly are incompatible with simple intermediation; the extra disclosure costs entailed by more complex arrangements lower the joint surplus available to be divided.

Basically, we have just argued that the high cost and the direction of flow of trade credit can be explained by a constraint on how the surplus in simple intermediation (by an entrepreneur) arrangements can be divided. The following corollary provides the formal statement of this result.

**COROLLARY 3 (Division of surplus):** *Let  $(\delta, \tau)$  be a simple intermediation by an entrepreneur arrangement, with optimal seniority between the intermediary and the investors  $I$ , low-risk, and debt-like entrepreneur payments. Let entrepreneur  $h$  be the nonintermediary. Then entrepreneur  $h$ 's welfare can be no more than*

$$\bar{\omega}_h - (1 - q_2)c - \rho - (\rho - L - \Pr(\omega_h = L, \omega_{h'} = H)(H - L)). \tag{24}$$

*In particular, provided the correlation in project outcomes is not so negative that  $\Pr(HH) < \Pr(\omega_h = L, \omega_{h'} = H)$ , then entrepreneur  $h$ 's income must be less than  $\bar{\omega}_h - (1 - q_2)c - \rho$ .*

*Proof:* See the Appendix.

**D. Intermediation by Financial Specialists**

Lemma 4's prediction that simple intermediation is best carried out by an entrepreneur is consistent with historical banking arrangements and with trade credit. However, it is at first sight hard to square with modern banks, which are clearly specialist financial institutions. One possible response is to argue that in fact, not all modern banks represent forms of simple intermediation—see the discussion of the credit crunch above. But large banks in particular do appear close to simple intermediation.

However, we can account for simple intermediation arrangements in which the intermediary is an investor if we assume that one of the investors is more

skilled in the activity of information transfers. Expertise in handling information is widely regarded as a core banking activity.<sup>31</sup> The special investor with low information transmission costs is essentially a potential banker in this setting.<sup>32</sup>

Specifically, suppose that the cost of disclosing to this special investor is  $\tilde{c} < c$ , and likewise that this special investor also has disclosure costs to other investors of  $\tilde{c}$ . By making this small change to the model we can account not just for the existence of simple intermediation by an investor, but also for under what circumstances it is most likely to emerge.

The argument is as follows. The total disclosure costs in a simple intermediation by an entrepreneur arrangement (with optimal seniority, low risk, and maximal nondisclosure) are

$$\Pr(\omega_2 = L)c + \left(2n - \left\lceil \frac{2L}{\rho_n} \right\rceil\right) \Pr(LL)c, \quad (25)$$

where without loss we assume entrepreneur 1 is the intermediary. If instead we make our special investor the intermediary, total disclosure costs are

$$\Pr(\omega_1 = L)\tilde{c} + \Pr(\omega_2 = L)\tilde{c} + \left(2n - 1 - \left\lceil \frac{2L}{\rho_n} \right\rceil\right) \Pr(LL)\tilde{c}. \quad (26)$$

That is, both entrepreneurs must now disclose, but the intermediary has one less investor to disclose to. So total disclosure costs are lower when our special investor intermediates if and only if

$$\Pr(LH)\tilde{c} + \Pr(\omega_2 = L)(\tilde{c} - c) + \left(2n - \left\lceil \frac{2L}{\rho_n} \right\rceil\right) \Pr(LL)(\tilde{c} - c) < 0. \quad (27)$$

This condition is clearly more likely to be satisfied when  $n$ , the number of investors required to finance each entrepreneur, is large. This is consistent with the emergence of modern banks, in which the expansion of the depositor base has gone hand in hand with the specialization of the banking function.

## VI. Discussion

This paper has argued that several apparently distinct facets of financial intermediation are in fact jointly determined. In particular, the choices of which projects are funded, how an intermediary is financed, whether the recipients of finance should be isolated from each other's risk, and who intermediates, are all linked. On the one hand, intermediaries that fund relatively high-risk projects will tend to resemble conglomerates. On the other hand, intermediaries funding low-risk projects will resemble either firms engaged in trade credit or modern

<sup>31</sup> See, for instance, Petersen and Rajan (2002) for an empirical study of how changing costs of information processing have changed the practice of banking.

<sup>32</sup> It should be noted that the introduction of an investor with lower costs of information transmission will do nothing to change the finding that joint-liability intermediation dominates high-risk simple intermediation.

banks, depending on whether or not specialization brings with it advantages in the transmission of information.

In the context of this paper, modern banks are just one of several different possible forms of financial intermediary. They are “special” in that they possess particular properties—both assets and liabilities are low risk, and borrowers are largely insulated from each other. But these properties are consequences of the same basic friction that also underlies other forms of intermediation, such as conglomerates.

In this paper, I have restricted attention to the simplest environment in which intermediation emerges: There are just two entrepreneurs, and each entrepreneur has just two income realizations. To what extent are the results obtained special to this formulation?

Generalizing the analysis to the case of nonbinary project outcomes would be straightforward but messy; the results obtained would be qualitatively similar. The consequences of increasing the number of sources of uncertainty are in general a little harder to predict. However, some insight can be gained by considering what happens as the probability that both projects in the current model simultaneously fail tends to zero (i.e.,  $\Pr(LL) \rightarrow 0$ ). This case is analogous to the benefits of diversification that would be obtainable if an intermediary could increase its size beyond funding just two projects. It is straightforward to see that as  $\Pr(LL)$  converges to zero, it is much easier to obtain low-risk intermediation. Consequently, conglomerate-like intermediation becomes less prevalent, and simple intermediation more prevalent. So with more than two projects, we might see another distinguishing feature of banks, namely that their portfolios are large in comparison to conglomerates. A full analysis of this case is left for future research.

Finally, one case in which the implications of increasing the number of entrepreneurs in the model are straightforward is the case of a two-sector economy (where entrepreneurs in each sector are highly correlated).<sup>33</sup> Here, when projects in both sectors are high quality and/or low risk we will again observe simple intermediation. On the other hand, when the projects are of lower quality, we will again observe joint-liability-like arrangements, where entrepreneurs in one sector make larger repayments to the intermediary when the other sector has performed poorly.

In the paper, I have assumed that the cost of information transfer  $c$  is exogenously fixed. Perhaps somewhat surprisingly, the relative merits of different financing arrangements are unaffected by the magnitude of  $c$ . This is easily seen once it is noted that the cost of some financial arrangement is just the probability of information sharing being called for, multiplied by its cost. What is affected by  $c$  is the viability, or otherwise, of raising funding. There exists a range of  $c$  in which only low-risk/high-quality projects are viable, and will be funded by intermediaries resembling simple banks, with the intermediary

<sup>33</sup> I thank the referee for this suggestion. Also, note that the entrepreneurial projects within each sector must be less than perfectly correlated—for otherwise the problem of asymmetric information could be easily solved by just having a representative entrepreneur from each sector disclose.

being a nonspecialist. As  $c$  falls, higher-risk and lower-quality projects become viable, with more complex financial institutions such as conglomerates emerging to provide financing. Thus the model predicts that (if information costs fall over time) simple financial institutions will emerge before more complex ones. At the same time, further decreases in  $c$  may allow for greater diversification, which in turn will tend to push intermediaries back to simpler bank-like structures.<sup>34</sup>

**Appendix: Mathematical Proofs**

*Proof of Lemma 1:* We will establish this for inequality (15); the other two cases follow similarly. Suppose that contrary to the stated result, inequality (15) holds strictly. That is,  $|N_{k,J}^\emptyset \cup N_{k,J}^{LL}| < |J|$  and

$$\Pr(LL)T_{k,J}(LL) + \Pr(LH, HL, HH)T_{k,J}(LH) - \rho_n |N_{k,J}^\emptyset \cup N_{k,J}^{LL}| \geq \rho_n. \tag{A1}$$

By the property of maximal use of seniority, all investors in  $J$  each get exactly  $\rho_n$  in expectation. So  $\rho_n |N_{k,J}^\emptyset \cup N_{k,J}^{LL}| = E[\sum_{i \in N_{k,J}^\emptyset \cup N_{k,J}^{LL}} (\tau_{ki}^1(\omega) + \tau_{ki}^2(\omega))]$ . Consequently, in at least one of the states  $LL$  or  $LH$ , agent  $k$  must make a nonzero transfer to a member of either  $N_{k,J}^{LL, LH}$  or  $N_{k,J}^{-HH}$  (since agent  $k$  is junior, the surplus resources cannot be going to him). Moreover, by the property of maximal use of seniority, we know that it is exactly one member of these more junior classes who receives a nonzero transfer—say investor  $j$ . Since the aggregate transfer to the more senior investors  $N_{k,J}^\emptyset \cup N_{k,J}^{LL}$  is constant over states  $LH$ ,  $HL$ , and  $HH$ , by absolute priority investor  $j$ 's transfer in each of these states is bounded below by  $T_{k,J}(LH) - \sum_{i \in N_{k,J}^\emptyset \cup N_{k,J}^{LL}} (\tau_{ki}^1(LH) + \tau_{ki}^2(LH))$ . Moreover, since agent  $j$  is disclosed to in state  $LH$ , his transfer in at least one of states  $HL$  and  $HH$  must be strictly above this lower bound. Finally, investor  $j$ 's transfer in state  $LL$  is at least  $T_{k,J}(LL) - \sum_{i \in N_{k,J}^\emptyset \cup N_{k,J}^{LL}} (\tau_{ki}^1(LL) + \tau_{ki}^2(LL))$ . Together these observations imply that  $E[\tau_{kj}^1(\omega) + \tau_{kj}^2(\omega)] > \rho_n$ , a contradiction to the maximal use of seniority. Q.E.D.

*Proof of Proposition 2:* Agent  $k$ 's total expected cost of disclosure to investors in  $J$  under arrangement  $(\delta, \tau)$  is simply

$$\begin{aligned} &c |N_{k,J}^{LL}(\delta, \tau)| \Pr(LL) + c |N_{k,J}^{LL, LH}(\delta, \tau)| \Pr(LL, LH) \\ &+ c |N_{k,J}^{-HH}(\delta, \tau)| \Pr(LL, LH, HH). \end{aligned} \tag{A2}$$

Rewriting gives

$$\begin{aligned} &c |J| - c \Pr(LL) |N_{k,J}^\emptyset(\delta, \tau)| - c \Pr(LH) (|N_{k,J}^\emptyset(\delta, \tau)| + |N_{k,J}^{LL}(\delta, \tau)|) \\ &- c \Pr(HL) (|N_{k,J}^\emptyset(\delta, \tau)| + |N_{k,J}^{LL}(\delta, \tau)| + |N_{k,J}^{LL, LH}(\delta, \tau)|). \end{aligned} \tag{A3}$$

<sup>34</sup> And of course, as the cost of information disclosure approaches 0 it will cease to have any effect as a determinant of the structure of financial institutions.

The change in disclosure costs is then immediate from Lemma 1's result that (14)–(16) hold at equality. Q.E.D.

*Proof of Proposition 3:* First, note that since the arrangement  $(\delta, \tau)$  has no insiders, it is trivially (weakly) dominated by an incentive compatible arrangement in which the only disclosure is from entrepreneurs to investors. So without loss we can assume that  $(\delta, \tau)$  itself possesses this property. Partition the investors into two subsets,  $M_1$  and  $M_2$ , where the subset  $M_h$  contains all investors disclosed to by entrepreneur  $h$  (investors who are never disclosed to can be allocated in an arbitrary way). Without loss, we can assume that only entrepreneur  $h$  ever makes a nonzero transfer to investors  $M_h$ . By Proposition 1, we know that  $(\delta, \tau)$  must be at least weakly dominated by an alternative arrangement in which for  $h = 1, 2$ , the subarrangement between entrepreneur  $h$  and investors  $M_h$  features optimal seniority—so again, without loss we assume that  $(\delta, \tau)$  itself has these properties. Let  $\Upsilon_{hL}$  and  $\Upsilon_{hH}$ , respectively, denote the total transfer made from entrepreneur  $h$  to the investors  $M_h$  when his project fails and succeeds. Without loss assume that  $\Upsilon_{2H} \leq \Upsilon_{1H}$ . For use below, let  $\alpha$  be any constant in  $[0, 1]$ , such that both  $\alpha \frac{\Pr(HL)}{\Pr(LH)} \in (0, 1)$  and  $\alpha(1 + \frac{\Pr(HL)}{\Pr(LH)}) \leq \varepsilon$ , where  $\varepsilon > 0$  is such that  $\frac{1}{\Upsilon_{2H}} \min_{h=1,2}(\Upsilon_{hH} - \Upsilon_{hL}) \geq \varepsilon$  in any possible arrangement satisfying the investor participation constraints (I-IR).<sup>35</sup>

Next, we construct a simple intermediation (by an entrepreneur) arrangement  $(\hat{\delta}, \hat{\tau})$  as follows. Let entrepreneur 1 be the intermediary. Define the transfers from entrepreneur 2 by  $C_2 = L$ , with  $R_2$  chosen to give entrepreneur 2 the same welfare he had under  $(\delta, \tau)$ , that is,  $U_2(\delta, \tau) = \bar{\omega}_2 - q_2 R_2 - (1 - q_2)(L + c)$ . From the property of optimal seniority and the fact that Assumption 2 implies that entrepreneur 2 must certainly have been disclosing in the failure state in  $(\delta, \tau)$ , we know that  $U_2(\delta, \tau) = q_2(H - \tau_{21}^1 - \tau_{21}^2 - \Upsilon_{2H}) - E[\gamma_2(\omega; \delta, \tau)]$ , that is, his only consumption comes in the success state.<sup>36</sup> So

$$q_2(R_2 - \tau_{21}^1 - \tau_{21}^2 - \Upsilon_{2H}) = E[\gamma_2(\omega; \delta, \tau)] - (1 - q_2)c \geq 0. \tag{A4}$$

Entrepreneur 1's utility under the new arrangement  $(\hat{\delta}, \hat{\tau})$  is

$$\begin{aligned} U_1(\hat{\delta}, \hat{\tau}) &= \bar{\omega}_1 + q_2 R_2 + (1 - q_2)L - 2\rho - E[\gamma_1(\omega; \hat{\delta}, \hat{\tau})] \\ &= \bar{\omega}_1 + \bar{\omega}_2 - 2\rho - (1 - q_2)c - U_2(\delta, \tau) - E[\gamma_1(\omega; \hat{\delta}, \hat{\tau})]. \end{aligned} \tag{A5}$$

Since certainly

$$U_1(\delta, \tau) + U_2(\delta, \tau) \leq \bar{\omega}_1 + \bar{\omega}_2 - 2\rho - E[\gamma_1(\omega; \delta, \tau)] - E[\gamma_2(\omega; \delta, \tau)], \tag{A6}$$

it follows that

$$U_1(\hat{\delta}, \hat{\tau}) \geq U_1(\delta, \tau) + E[\gamma_1(\omega; \delta, \tau)] + E[\gamma_2(\omega; \delta, \tau)] - E[\gamma_1(\omega; \hat{\delta}, \hat{\tau})] - (1 - q_2)c. \tag{A7}$$

<sup>35</sup> Observe that  $(\Upsilon_{hH} - \Upsilon_{hL})/\Upsilon_{2H}$  must be bounded away from 0: Certainly  $\Upsilon_{2H} \leq L + H$ , and Assumption 2 ensures that  $\Upsilon_{hH} - \Upsilon_{hL}$  is bounded away from 0.

<sup>36</sup> Observe that  $\tau_{21}^1 + \tau_{21}^2$  must be constant across states.

That is, to show that the new arrangement strictly dominates the old one, it suffices to show that the disclosure costs of the new arrangement are strictly less than under the old arrangement.

To establish this, we make use of the characterization of the change in disclosure costs provided by Proposition 2. The argument is as follows:

1. The disclosure costs  $E[\gamma_1(\omega; \delta, \tau)] + E[\gamma_2(\omega; \delta, \tau)]$  under  $(\delta, \tau)$  are equal to those that *would* arise in an economy in which each entrepreneur  $h = 1, 2$  had a total income of  $\Upsilon_{hL}$  and  $\Upsilon_{hH}$ , respectively, when  $\omega_h = L, H$ , no transfer occurs between the two entrepreneurs, and the subarrangement between the entrepreneur  $h$  and investors  $M_h$  features optimal seniority.
2. Consider the minimal disclosure costs that could be achieved if entrepreneur 1 has income  $\Upsilon_{1L} + \Upsilon_{2L}$ ,  $\Upsilon_{1L} + \Upsilon_{2H}$ ,  $\Upsilon_{1H} + \Upsilon_{2L}$ ,  $\Upsilon_{1H} + \Upsilon_{2H}$  in states  $LL, LH, HL$ , and  $HH$  respectively, and has to deliver  $\rho_n$  in expectation to each of the  $2n$  investors. I claim that for  $n$  large enough, this minimal disclosure cost is strictly less than  $E[\gamma_1(\omega; \delta, \tau)] + E[\gamma_2(\omega; \delta, \tau)]$ :
  - (a) The entrepreneur could *artificially* split the above income process into two: One part, given by  $\Upsilon_{1L}, \lambda_{LH}(\Upsilon_{1L} + \Upsilon_{2H}), \lambda_{HL}(\Upsilon_{1H} + \Upsilon_{2L}), \Upsilon_{1H}$  in states  $LL, LH, HL$ , and  $HH$  respectively, to be paid out only to investors in  $M_1$ , and the remainder to be paid out to investors in  $M_2$ . The constants  $\lambda_{LH}$  and  $\lambda_{HL}$  lie in the interval  $(0, 1)$ , with their determination described below.
  - (b) The constants  $\lambda_{LH}$  and  $\lambda_{HL}$  are chosen to satisfy

$$\begin{aligned} & \Pr(LH)\lambda_{LH}(\Upsilon_{1L} + \Upsilon_{2H}) + \Pr(HL)\lambda_{HL}(\Upsilon_{1H} + \Upsilon_{2L}) \\ & = \Pr(LH)\Upsilon_{1L} + \Pr(HL)\Upsilon_{1H}, \end{aligned} \tag{A8}$$

$$\lambda_{LH}(\Upsilon_{1L} + \Upsilon_{2H}) > \Upsilon_{1L}, \tag{A9}$$

$$(1 - \lambda_{HL})(\Upsilon_{1H} + \Upsilon_{2L}) > \Upsilon_{2L}, \tag{A10}$$

$$\lambda_{HL}(\Upsilon_{1H} + \Upsilon_{2L}) \geq \lambda_{LH}(\Upsilon_{1L} + \Upsilon_{2H}), \tag{A11}$$

$$(1 - \lambda_{LH})(\Upsilon_{1L} + \Upsilon_{2H}) \geq (1 - \lambda_{HL})(\Upsilon_{1H} + \Upsilon_{2L}). \tag{A12}$$

Provided these four conditions are satisfied, the portion of the income process reserved for investors  $M_1$  is essentially a “smoothed out” version of the initial process considered,  $\Upsilon_{1L}, \Upsilon_{1L}, \Upsilon_{1H}, \Upsilon_{1H}$ . That is, the entrepreneur now has a higher income in state  $LH$ , a lower income in state  $HL$ , and the same expected income. More formally, these four conditions ensure that the income process  $\Upsilon_{1L}, \lambda_{LH}(\Upsilon_{1L} + \Upsilon_{2H}), \lambda_{HL}(\Upsilon_{1H} + \Upsilon_{2L}), \Upsilon_{1H}$  (in states  $LL, LH, HL$ , and  $HH$  respectively) SOSD<sup>37</sup> the process  $\Upsilon_{1L}, \Upsilon_{1L}, \Upsilon_{1H}, \Upsilon_{1H}$  (in states  $LL, LH, HL$ , and  $HH$  respectively), while the residual portion reserved for investors  $M_2$  SOSD the process

<sup>37</sup> Straightforward concavity arguments confirm that the former income process is preferred to the latter by any risk-averse agent.

$\Upsilon_{2L}, \Upsilon_{2H}, \Upsilon_{2L}, \Upsilon_{2H}$  (again in states  $LL, LH, HL,$  and  $HH$  respectively). Moreover, in both cases the ordering of the income process across the four states is unchanged.

- (c) Do there exist  $\lambda_{LH}$  and  $\lambda_{HL}$  satisfying conditions (A8) through (A12)? Define

$$\lambda_{LH} = \frac{\Upsilon_{1L} + \alpha \frac{\Pr(HL)}{\Pr(LH)} \Upsilon_{2H}}{\Upsilon_{1L} + \Upsilon_{2H}}, \tag{A13}$$

$$\lambda_{HL} = \frac{\Upsilon_{1H} - \alpha \Upsilon_{2H}}{\Upsilon_{1H} + \Upsilon_{2L}}. \tag{A14}$$

These definitions ensure that condition (A8) is satisfied, and given the definition of  $\alpha$  and the fact that  $\Upsilon_{2H} \leq \Upsilon_{1H}$  (see start of proof), we have  $\lambda_{LH}, \lambda_{HL} \in (0, 1)$ . Finally, conditions (A9) through (A12) can also easily be seen to be satisfied.

- (d) Moreover, note that since  $\alpha > 0$ , we have strictly increased the resources available for the investors  $M_1$  in state  $LH$  and for investors  $M_2$  in state  $HL$ . Combined with SOSD, Proposition 2 then implies that disclosure costs are strictly lower for  $n$  large enough.
- (e) Since we have been able to achieve a reduction in disclosure costs even under this artificial division of the income process  $\Upsilon_{1L} + \Upsilon_{2L}, \Upsilon_{1L} + \Upsilon_{2H}, \Upsilon_{1H} + \Upsilon_{2L}, \Upsilon_{1H} + \Upsilon_{2H}$ , the subclaim follows a fortiori.
3. In fact, we know that entrepreneur 1's income process under  $(\hat{\delta}, \hat{\tau})$  is  $2L, L + R_2, L + H,$  and  $H + R_2$  (in states  $LL, LH, HL,$  and  $HH,$  respectively). In states  $LL$  and  $HL$  this income is trivially higher than that under the process considered above. In state  $LH, \Upsilon_{1L} + \Upsilon_{2H} \leq L + (\tau_{21}^1 + \tau_{21}^2) + \Upsilon_{2H} \leq L + R_2$  (from (A4)), while similarly in state  $HH,$  we know  $\Upsilon_{1H} + \Upsilon_{2H} \leq H + (\tau_{21}^1 + \tau_{21}^2) + \Upsilon_{2H} \leq H + R_2$ . So a fortiori, the disclosure costs between entrepreneur 1 and the investors must be strictly lower under  $(\hat{\delta}, \hat{\tau})$  than the total disclosure costs in  $(\delta, \tau)$ .
4. Finally, note that the only other disclosure cost under  $(\hat{\delta}, \hat{\tau})$  is that of entrepreneur 2 disclosing to entrepreneur 1 when he fails. But for all  $n$  sufficiently large, this is outweighed by the savings in the cost of disclosure to the investors. Q.E.D.

*Proof of Lemma 3:* Denote the intermediary agent by  $m$ . A direct application of Proposition 1 implies that  $(\delta, \tau)$  is always dominated by a simple intermediation arrangement  $(\hat{\delta}, \hat{\tau})$  with the same intermediary and the same entrepreneur payments, in which the subarrangement between the intermediary  $m$  and the investors  $I_m \equiv I \setminus \{m\}$  features optimal seniority. If  $R_h - C_h \leq H - L$  for both entrepreneurs  $h = 1, 2$ , there is nothing more to show. So without loss suppose that  $R_1 - C_1 > H - L$ .

Consider first the case in which  $R_2 - C_2 \leq H - L$ , so that  $C_1 + R_2 < C_2 + R_1$ . Let  $(\hat{\delta}, \hat{\tau})$  be the simple intermediation arrangement with entrepreneur payments  $(\hat{R}_1, \hat{R}_2, \hat{C}_1, \hat{C}_2)$ , with agent  $m$  still being the intermediary, and with the subarrangement between the intermediary  $m$  and the investors  $I_m$  still

featuring optimal seniority. The payments from entrepreneur 2 are left unchanged, that is,  $(\hat{C}_2, \hat{R}_2) = (C_2, R_2)$ ; while  $\hat{C}_1$  is set higher than  $C_1$ , with  $\hat{R}_1$  lowered by an offsetting amount so as to leave the expected amount paid by entrepreneur 1 unchanged. Any change of this sort leaves the welfare of all nonintermediary entrepreneurs unaffected, and provided that the intermediary's income remains weakly higher in state *HL* than state *LH*, Proposition 2 implies that the intermediary's disclosure costs are at least weakly reduced, and thus his welfare is at least weakly increased as well. Given this, to establish the required result we need only choose  $\hat{C}_1$  so that either feasibility binds, that is,  $\hat{C}_1 = L$ , or the intermediary's income in states *LH* and *HL* is equalized, that is,  $\hat{R}_1 - \hat{C}_1 = R_2 - C_2$ . In either case, we end up with an arrangement that at least weakly dominates the initial one and in which the entrepreneurial payments are debt-like.

It remains to deal with the case in which  $R_2 - C_2 > H - L$  also. In this case, set the new entrepreneur payments under  $(\hat{\delta}, \hat{\tau})$  to  $\hat{R}_h = R_h - (1 - q_h)\varepsilon$ ,  $\hat{C}_h = C_h + q_h\varepsilon$  for both entrepreneurs  $h = 1, 2$ . The welfare of both entrepreneurs is unaffected. Moreover, it is easily shown that since  $q_h \geq 1 - q_h$ , this change increases the intermediary's income in all states other than *HH* and leaves his expected income unchanged, and so at least weakly reduces the intermediary's disclosure costs and increases his welfare. To complete the proof, just choose  $\varepsilon$  so that the feasibility of the failure payment binds for at least one of the two entrepreneurs, that is,  $\hat{C}_1 = L$  or  $\hat{C}_2 = L$ . So  $\hat{R}_h - \hat{C}_h \leq H - L$  for at least one of  $h = 1, 2$ . The proof is then completed by an application of the first case above. Q.E.D.

*Proof of Proposition 4:* Let  $(\delta, \tau)$  be a high-risk simple intermediation arrangement with optimal seniority, debt-like entrepreneur payments and satisfying the entrepreneur participation constraints (E-IR). Let agent  $m$  be the intermediary, and let  $(R_1, R_2, C_1, C_2)$  denote the entrepreneur payments.

On the one hand, if the intermediary is one of the entrepreneurs, assume without loss that it is entrepreneur 2 who is not the intermediary. Since the entrepreneur participation constraint (E-IR) holds, either  $C_2 \leq L - (1 - q_2)c$  or  $R_2 \leq H - (1 - q_2)c$ ; but if the former holds, then since the entrepreneur payments are debt-like, we have  $L + R_2 \leq C_2 + H$  and so again  $R_2 \leq H - (1 - q_2)c$ . On the other hand, if the intermediary is an investor, then without loss assume that  $C_1 + R_2 \leq C_2 + R_1$ . And by the same argument as above,  $R_2 \leq H - (1 - q_2)c$ .

Since the investor participation constraint (I-IR) must hold, Assumption 2 implies that  $R_h - C_h \geq \hat{\varepsilon}$  for some  $\hat{\varepsilon} > 0$  that depends only on the underlying entrepreneur project parameters  $L, H, q_1, q_2$ , and  $\rho$ . So there exists an  $\varepsilon > 0$ , again dependent only on  $L, H, q_1, q_2$ , and  $\rho$ , such that

$$R_2 + \varepsilon \leq H \quad \text{and}$$

$$R_1 + R_2 - \frac{\Pr(LH, HH)}{\Pr(HH)}\varepsilon \geq \max\{C_1 + R_2 + \varepsilon, R_1 + C_2\}. \quad (A15)$$

Now, consider an intermediation with joint-liability arrangement  $(\hat{\delta}, \hat{\tau})$ , in which the same agent  $m$  acts as the intermediary, optimal seniority holds between the intermediary and the investors, but in which entrepreneur 2 now pays  $R_2 + \varepsilon$  when he succeeds but gets a bonus  $B_2 = \frac{\Pr(LH,HH)}{\Pr(HH)}\varepsilon$  back when both entrepreneurs succeed. Note that entrepreneur 2's welfare is unchanged. The expected disclosure cost from the intermediary to the entrepreneur associated with making this bonus payment incentive compatible is  $\Pr(LH)$ , that is, the intermediary discloses when the bonus payment is not going to be made.

To complete the proof it suffices to show that the intermediary's expected cost of disclosing to investors  $I_m \equiv I \setminus \{m\}$  is lower under  $(\hat{\delta}, \hat{\tau})$  than under  $(\delta, \tau)$  by at least  $c\Pr(LH)$ . This is basically an application of Proposition 2. The main complication is that introducing joint liability may cause the intermediary's income to be strictly higher in state  $LH$  than in state  $HL$ .

To handle this possibility, define  $\varepsilon' = \min\{\varepsilon, (R_1 + C_2) - (C_1 + R_2)\}$ ,  $\varepsilon'' = \varepsilon - \varepsilon'$ , and let  $(\tilde{\delta}, \tilde{\tau})$  be the joint-liability arrangement in which entrepreneur 2 pays  $R_2 + \varepsilon'$  when he succeeds and gets back  $\varepsilon'/q_1$  when entrepreneur 1 also succeeds. Recall that  $T_{m,I_m}(\omega; \delta, \tau)$  denotes intermediary  $m$ 's total consumption plus transfer to agents  $I_m$ .

First we apply Proposition 2 for the change from  $(\delta, \tau)$  to  $(\tilde{\delta}, \tilde{\tau})$ . Then when  $\tilde{\Delta}_\zeta$  (where  $\zeta = LL, LH, HL$ ) is as defined in (20) of Proposition 2,  $\tilde{\Delta}_{LL} = 0$  while

$$\begin{aligned} \tilde{\Delta}_{LH} = & \min \left\{ |I_m|, \left[ \frac{1}{\rho_n} E[\min\{T_{m,I_m}(\omega; \tilde{\delta}, \tilde{\tau}), C_1 + R_2 + \varepsilon'\}] \right] \right\} \\ & - \min \left\{ |I_m|, \left[ \frac{1}{\rho_n} E[\min\{T_{m,I_m}(\omega; \delta, \tau), C_1 + R_2\}] \right] \right\} \end{aligned} \tag{A16}$$

and  $\tilde{\Delta}_{HL} \geq 0$ . Next, we consider the change from  $(\tilde{\delta}, \tilde{\tau})$  to  $(\hat{\delta}, \hat{\tau})$ . For this, we make use of a parallel version of Proposition 2 that holds when  $T_{m,I_m}(LH; \delta, \tau) \geq T_{m,I_m}(HL; \delta, \tau)$ . With  $\hat{\Delta}_\zeta$  again as defined by (20) we have  $\hat{\Delta}_{LL} = \hat{\Delta}_{HL} = 0$  while

$$\begin{aligned} \hat{\Delta}_{LH} = & \min \left\{ |I_m|, \left[ \frac{1}{\rho_n} E[\min\{T_{m,I_m}(\omega; \hat{\delta}, \hat{\tau}), C_1 + R_2 + \varepsilon' + \varepsilon''\}] \right] \right\} \\ & - \min \left\{ |I_m|, \left[ \frac{1}{\rho_n} E[\min\{T_{m,I_m}(\omega; \tilde{\delta}, \tilde{\tau}), C_1 + R_2 + \varepsilon'\}] \right] \right\}. \end{aligned} \tag{A17}$$

Thus

$$\begin{aligned} \tilde{\Delta}_{LH} + \hat{\Delta}_{LH} = & \min \left\{ |I_m|, \left[ \frac{1}{\rho_n} E[\min\{T_{m,I_m}(\omega; \hat{\delta}, \hat{\tau}), C_1 + R_2 + \varepsilon\}] \right] \right\} \\ & - \min \left\{ |I_m|, \left[ \frac{1}{\rho_n} E[\min\{T_{m,I_m}(\omega; \delta, \tau), C_1 + R_2\}] \right] \right\} \end{aligned} \tag{A18}$$

and the expected disclosure costs from intermediary  $m$  to investor  $I_m$  decrease by at least  $c\Pr(LH)(\tilde{\Delta}_{LH} + \hat{\Delta}_{LH})$ . Observe that

$$\begin{aligned}
 & E[\min\{T_{m,I_m}(\omega; \hat{\delta}, \hat{\tau}), C_1 + R_2 + \varepsilon\}] \\
 &= E[\min\{T_{m,I_m}(\omega; \delta, \tau), C_1 + R_2\}] + \varepsilon'(1 - \Pr(LL)) + \varepsilon''\Pr(LH, HH) \\
 &\geq E[\min\{T_{m,I_m}(\omega; \delta, \tau), C_1 + R_2\}] + \varepsilon\Pr(LH, HH). \tag{A19}
 \end{aligned}$$

So clearly there exists some  $n^*$  such that whenever  $n \geq n^*$ , we have

$$\begin{aligned}
 & \frac{1}{\rho_n} E[\min\{T_{m,I_m}(\omega; \hat{\delta}, \hat{\tau}), C_1 + R_2 + \varepsilon\}] \\
 & \geq \frac{1}{\rho_n} E[\min\{T_{m,I_m}(\omega; \delta, \tau), C_1 + R_2\}] + 2. \tag{A20}
 \end{aligned}$$

Since the arrangement  $(\delta, \tau)$  is high risk, we know that

$$\left[ \frac{1}{\rho_n} E[\min\{T_{m,I_m}(\omega; \delta, \tau), C_1 + R_2\}] \right] \leq |I_m| - 1 \tag{A21}$$

and so  $\tilde{\Delta}_{LH} + \hat{\Delta}_{LH} \geq 1$  for all  $n \geq n^*$ . Moreover, if the intermediary  $m$  discloses to at least two investors in a state other than  $LL$ , then inequality (A21) must hold strictly, and so  $\tilde{\Delta}_{LH} + \hat{\Delta}_{LH} \geq 2$  for all  $n \geq n^*$ . In both cases the welfare of all agents other than the intermediary is the same under arrangements  $(\delta, \tau)$  and  $(\hat{\delta}, \hat{\tau})$ , while the intermediary's welfare is weakly increased. In the latter case, the intermediary's welfare is increased by at least  $c\Pr(LH)$ . This completes the proof. Q.E.D.

*Proof of Lemma 4:* Let  $(\delta, \tau)$  be a low-risk simple intermediation by an investor arrangement, with optimal seniority, debt-like entrepreneur payments, and satisfying the entrepreneur participation constraints (E-IR). Let investor  $m$  be the intermediary, and let  $(R_1, R_2, C_1, C_2)$  denote the entrepreneur payments, where without loss we assume that  $C_1 + R_2 \leq R_1 + C_2$ . From the entrepreneur participation constraints (E-IR), we know that at least one of  $C_1 \leq L - (1 - q_1)c$  and  $R_1 \leq H - (1 - q_1)c$  holds, and so since the entrepreneur payments are debt-like,  $R_1 \leq H - (1 - q_1)c$  must hold for sure.

Next, construct a simple intermediation arrangement  $(\hat{\delta}, \hat{\tau})$  as follows: Let entrepreneur 1 be the intermediary, keep entrepreneur 2's payments to the new intermediary unchanged at  $C_2$  and  $R_2$ , and let the subarrangement between entrepreneur 1 and the investors  $I$  feature optimal seniority. Note that since the original entrepreneur payments are debt-like, we still have the same ordering of intermediary income across states, that is,

$$L + C_2 < L + R_2 \leq H + C_2 < H + R_2. \tag{A22}$$

Since the original arrangement  $(\delta, \tau)$  was low risk,

$$\Pr(LL)(C_1 + C_2) + \Pr(LH, HL, HH)(C_1 + R_2) \geq |I_m| \rho_n = (2n - 1) \rho_n. \tag{A23}$$

Then since  $H \geq R_1 + (1 - q_1)c$ , there exists an  $n^*$  such that for all  $n \geq n^*$

$$\Pr(LL)(L + C_2) + \Pr(LH)(L + R_2) + \Pr(HL, HH)(H + C_2) \geq 2n\rho_n. \tag{A24}$$

Consequently, we know that no disclosure will take place in states *HL* and *HH* in the new intermediation arrangement  $(\hat{\delta}, \hat{\tau})$ , and moreover the increase in intermediary disclosure costs between  $(\delta, \tau)$  and  $(\hat{\delta}, \hat{\tau})$  is no more than  $\Pr(LL, LH)$ .

Since making entrepreneur 1 the intermediary eliminates disclosure from that entrepreneur to the intermediary in states *LL* and *LH*, it follows that the new arrangement must have weakly lower aggregate disclosure costs than the old one. By construction, the welfare of each agent other than entrepreneur 1 and the old intermediary *m* is unchanged, while the welfare of the old intermediary is weakly decreased (to exactly  $\rho_n$ ). It follows that entrepreneur 2's welfare is now weakly higher, completing the proof. Q.E.D.

*Proof of Corollary 1:* Take  $n^*$  large enough so that Proposition 3 and Proposition 4 apply, and such that we can add  $3\rho_n$  to the left-hand side of (23) without violating the inequality. Let  $(\delta, \tau)^0$  be an incentive compatible arrangement with one or no insiders that satisfies the participation constraints (I-IR) and (E-IR). If  $(\delta, \tau)^0$  has no insiders, Proposition 3 implies that it is dominated by a simple intermediation arrangement  $(\delta, \tau)^1$ . Likewise, if  $(\delta, \tau)^0$  has one insider, Lemma 2 implies that it is dominated by a simple intermediation arrangement  $(\delta, \tau)^1$ . Next, take a sequence of simple intermediation arrangements, starting with  $(\delta, \tau)^1$ , with each one dominating the preceding arrangement and decreasing aggregate expected disclosure costs by at least  $\min\{\Pr(LH), \Pr(HL)\}c$ . The sequence clearly stops, say at  $(\delta, \tau)^Z$ . Lemma 3 implies that  $(\delta, \tau)^Z$  is dominated by a simple intermediation arrangement  $(\tilde{\delta}, \tilde{\tau})$  with optimal seniority between the intermediary and the investors, and with debt-like entrepreneur payments. By hypothesis, the arrangement  $(\tilde{\delta}, \tilde{\tau})$  must be high risk with the intermediary disclosing to at least two investors in a state other than *LL*. So Proposition 4 applies and  $(\tilde{\delta}, \tilde{\tau})$  is dominated by an intermediation with joint-liability arrangement  $(\hat{\delta}, \hat{\tau})$ , with the combined decrease in aggregate disclosure costs equal to at least  $\min\{\Pr(LH), \Pr(HL)\}c$ . Finally, observe that  $(\hat{\delta}, \hat{\tau})$  cannot in turn be dominated by an arrangement with one (respectively, no) insider, or else Lemma 2 (respectively, Proposition 3) would imply that we could find a further element of the sequence  $\{(\delta, \tau)^z : z = 0, 1, \dots, Z\}$ , a contradiction. This completes the proof. Q.E.D.

*Proof of Corollary 2:* Without loss assume that entrepreneur 1 is the intermediary. Total disclosure costs under arrangement  $(\delta, \tau)$  are

$$(1 - q_2)c + \left(2n - \left\lceil \frac{2L}{\rho_n} \right\rceil\right) \Pr(LL)c. \tag{A25}$$

Next, let  $(\tilde{\delta}, \tilde{\tau})$  be any incentive compatible arrangement satisfying the investor participation constraint (I-IR). By Assumption 2, both entrepreneurs must disclose. Moreover, at least  $2n - \lceil 2L/\rho_n \rceil$  investors must receive some disclosure.

If an arrangement  $(\tilde{\delta}, \tilde{\tau})$  has two or more insiders, total expected disclosure costs are at least

$$\min_{h=1,2} (1 - q_h)c + \min_{\omega \in \Omega} \Pr(\omega)c + \left(2n - \left\lceil \frac{2L}{\rho_n} \right\rceil\right) \Pr(LL)c, \tag{A26}$$

which by hypothesis is greater than expression (A25). So no two-insider arrangements can dominate  $(\delta, \tau)$ . If an arrangement  $(\tilde{\delta}, \tilde{\tau})$  has one insider, from Lemmas 2 and 3 it is dominated by a simple intermediation arrangement  $(\hat{\delta}, \hat{\tau})$  with optimal seniority between the intermediary and the investor. Total expected disclosure costs are at least

$$\min_{h=1,2} (1 - q_h)c + \left( 2n - \left\lceil \frac{2L}{\rho_n} \right\rceil \right) \Pr(LL)c, \quad (\text{A27})$$

which is strictly more than (A25) unless  $(\hat{\delta}, \hat{\tau})$  is low risk, has maximal nondisclosure, and has an entrepreneur as the intermediary. Finally, if an arrangement  $(\tilde{\delta}, \tilde{\tau})$  has no insiders, then Proposition 3 implies that it is dominated by a simple intermediation arrangement, and the above analysis just applies again. Q.E.D.

*Proof of Corollary 3:* Without loss, assume that entrepreneur 1 is the intermediary. Since the arrangement  $(\delta, \tau)$  is low risk, entrepreneur 2's payments  $C_2$  and  $R_2$  must satisfy

$$(L + C_2)\Pr(LL) + (L + R_2)\Pr(LH, HL, HH) \geq 2\rho, \quad (\text{A28})$$

or equivalently

$$C_2\Pr(LL) + R_2\Pr(LH, HL, HH) \geq 2\rho - L. \quad (\text{A29})$$

Together with the fact that entrepreneur 2's payments are debt-like, inequality (A29) implies that entrepreneur 2's welfare is

$$\begin{aligned} U_2(\delta, \tau) &= \bar{\omega}_2 - (1 - q_2)c - \Pr(LL, HL)C_2 - \Pr(HL, HH)R_2 \\ &= \bar{\omega}_2 - (1 - q_2)c - \Pr(LL)C_2 - \Pr(LH, HL, HH)R_2 \\ &\quad + \Pr(HL)(R_2 - C_2) \\ &\leq \bar{\omega}_2 - (1 - q_2)c - \rho - (\rho - L - \Pr(HL)(H - L)). \end{aligned} \quad (\text{A30})$$

Note that Assumption 2 requires that

$$\rho - L > \frac{1}{2}\Pr(HL, HH)(H - L). \quad (\text{A31})$$

So whenever  $\Pr(HH) \geq \Pr(HL)$ , then  $U_2(\delta, \tau) < \bar{\omega}_2 - (1 - q_2)c - \rho$ . Q.E.D.

## REFERENCES

- Bernanke, Ben S., and Cara S. Lown, 1991, The credit crunch, *Brookings Papers on Economic Activity* 204–239.
- Biais, Bruno, and Christian Gollier, 1997, Trade credit and credit rationing, *Review of Financial Studies* 10, 903–937.
- Campa, José Manuel, and Simi Kedia, 2002, Explaining the diversification discount, *Journal of Finance* 57, 1731–1762.
- Cerasi, Vittoria, and Sonja Daltung, 1998, Close relationships between banks and firms: Is it good or bad? *Research in Economics* 52, 233–253.

- Cerasi, Vittoria, and Sonja Daltung, 2000, The optimal size of a bank: Costs and benefits of diversification, *European Economic Review* 44, 1701–1726.
- Diamond, Douglas W., 1984, Financial intermediation and delegated monitoring, *Review of Economic Studies* 51, 393–414.
- Diamond, Douglas W., and Raghuram G. Rajan, 2000, A theory of bank capital, *Journal of Finance* 55, 2431–2465.
- Fluck, Zsuzsana, and Anthony W. Lynch, 1999, Why do firms merge and then divest? A theory of financial synergy, *Journal of Business* 72, 319–346.
- Frank, Murray, and Vojislav Maksimovic, 1998, Trade credit, collateral, and adverse selection, Working paper, University of British Columbia and University of Maryland.
- Freixas, Xavier, and Jean-Charles Rochet, 1997, *Microeconomics of Banking* (MIT Press, Cambridge).
- Gale, Douglas, and Martin Hellwig, 1985, Incentive-compatible debt contracts: The one-period problem, *Review of Economic Studies* 52, 647–663.
- Gertner, Robert H., David S. Scharfstein, and Jeremy C. Stein, 1994, Internal versus external capital markets, *Quarterly Journal of Economics* 109, 1211–1230.
- Graham, John R., Michael L. Lemmon, and Jack G. Wolf, 2002, Does corporate diversification destroy value? *Journal of Finance* 57, 695–720.
- Hancock, Diana, and James A. Wilcox, 1998, The “credit crunch” and the availability of credit to small business, *Journal of Banking and Finance* 22, 983–1014.
- Harris, Milton, and Artur Raviv, 1991, The theory of capital structure, *Journal of Finance* 46, 297–355.
- Hellwig, Martin F., 2000, Financial intermediation with risk aversion, *Review of Economic Studies* 67, 719–742.
- Holmström, Bengt, and Jean Tirole, 1997, Financial intermediation, loanable funds, and the real sector, *Quarterly Journal of Economics* 112, 663–691.
- Hubbard, R. Glenn, Kenneth Kuttner, and Darius Palia, 2002, Are there bank effects in borrowers’ costs of funds? Evidence from a matched sample of borrowers and banks, *Journal of Business* 75, 559–581.
- Krasa, Stefan, and Anne P. Villamil, 1992a, Monitoring the monitor: An incentive structure for a financial intermediary, *Journal of Economic Theory* 57, 197–221.
- Krasa, Stefan, and Anne P. Villamil, 1992b, A theory of optimal bank size, *Oxford Economic Papers* 44, 725–749.
- Lamont, Owen, 1997, Cash flow and investment: Evidence from internal capital markets, *Journal of Finance* 52, 83–109.
- Lamoreaux, Naomi R., 1994, *Insider Lending: Banks, Personal Connections, and Economic Development in Industrial New England* (Cambridge University Press, Cambridge).
- Maksimovic, Vojislav, and Gordon Phillips, 2002, Do conglomerate firms allocate resources inefficiently across industries? Theory and evidence, *Journal of Finance* 57, 721–767.
- Myers, Stewart C., 1977, Determinants of corporate borrowing, *Journal of Financial Economics* 5, 147–175.
- Petersen, Mitchell A., and Raghuram G. Rajan, 1997, Trade credit: Theories and evidence, *Review of Financial Studies* 10, 661–691.
- Petersen, Mitchell A., and Raghuram G. Rajan, 2002, Does distance still matter? The information revolution in small business lending, *Journal of Finance* 57, 2533–2570.
- Shin, Hyun-Han, and René M. Stulz, 1998, Are internal capital markets efficient? *Quarterly Journal of Economics* 113, 531–552.
- Townsend, Robert M., 1979, Optimal contracts and competitive markets with costly state verification, *Journal of Economic Theory* 21, 265–293.
- Williamson, Stephen D., 1986, Costly monitoring, financial intermediation, and equilibrium credit rationing, *Journal of Monetary Economics* 18, 159–179.
- Winton, Andrew, 1995a, Costly state verification and multiple investors: The role of seniority, *Review of Financial Studies* 8, 91–123.
- Winton, Andrew, 1995b, Delegated monitoring and bank structure in a finite economy, *Journal of Financial Intermediation* 4, 158–187.