Moral Hazard and Information Sharing: A Model of Financial Information Gathering Agencies

Marcia H. Millon; Anjan V. Thakor

The Journal of Finance, Vol. 40, No. 5 (Dec., 1985), 1403-1422.

THE JOURNAL OF FINANCE • VOL. XL, NO. 5 • DECEMBER 1985

Moral Hazard and Information Sharing: A Model of Financial Information Gathering Agencies

MARCIA H. MILLON and ANJAN V. THAKOR*

ABSTRACT

We propose a theory of information gathering agencies in a world of informational asymmetries and moral hazard. In a setting in which true firm values are certified by screening agents whose payoffs depend on noisy *ex post* monitors of information quality, the formation of information gathering agencies (groups of screening agents) is justified on two grounds. First, it enables screening agents to diversify their risky payoffs. Second, it allows information sharing. The first effect itself is insufficient despite the risk aversion of screening agents and the stochastic independence of the monitors used to compensate them.

THIS PAPER DEVELOPS A model that explains the formation of financial institutions which acquire and process information for the purpose of certifying asset qualities, but do not get involved in funding. The model makes two basic assumptions. The first is that there is an informational asymmetry in the capital market, involving "insiders" possessing more accurate information about the true economic values of their firms than "outsiders", and the second is that those engaged in the production of information to rectify this asymmetry can gain from sharing their information.

Examples of the *information gathering agencies* (IGA's) we study are credit-rating agencies, financial newsletters, Standard and Poor's Value Line, investment counselors, credit bureaus, etc. IGA's are distinct from *funding financial intermediaries* (FFI's)—like banks and credit unions—that, in addition to acquiring and processing information, also issue claims against themselves, fund their customers' needs, and transform asset characteristics. From a modeling standpoint, a key distinction between IGA's and FFI's is that, in the case of the latter, "outsiders" (depositors, for instance) do *not* necessarily have to assess the quality of the intermediary's information since the intermediary can be made "accountable" for this quality through the positions it takes in the screened assets. Thus, models that explain FFI's focus on how the contracts offered by the intermediary ensure information reliability *without* reliance on explicit monitoring. On the other hand, because of the absence of loan- and deposit-type contracts to dissipate informational problems, a model explaining IGA's must rely on monitoring-related mechanisms.

* Assistant Professor of Finance at Boston College and Associate Professor of Finance at the Indiana University School of Business, respectively. Revisions of this paper were made while Thakor was Visiting Associate Professor of Finance at the J. L. Kellogg Graduate School of Management, Northwestern University. The comments of anonymous referees, and particularly the suggestions of an Associate Editor, have been very helpful in improving this paper. Responsibility for errors or infelicitous remarks, if any, rests solely with the authors.

We consider an economy in which managers of firms wish to sell new shares to investors who do not know the true values of these shares. Managers, however, have priors about their own firms which are superior to the market's. Each firm may have its true value revealed by availing itself of the services of screening agencies (Stiglitz [24] and Viscusi [25]). Screening agents (s.a.'s) are induced to produce reliable information by basing their compensation on a publicly observable, noisy, *ex post* indicator of information quality. Our objective is to establish conditions under which individual s.a.'s choose to form a group and become an IGA (a collection of two or more s.a.'s). Because we assume s.a.'s to be risk-averse and their payoffs to be noisy, a seemingly obvious rationale for group formation is diversification. However, subsequent to developing the model in Section I, we prove in Section II that diversification is insufficient because group formation worsens the severity of moral hazard. We then show in Section III that an IGA will nonetheless emerge (and will generally be of finite size) if group formation permits information sharing by members, and the benefits of such sharing outweigh the cost of heightened moral hazard. Section IV concludes. An Appendix contains all the proofs.

Our paper is part of the literature that views informational imperfections as central to intermediation.[1] A preliminary analysis was presented by Leland and Pyle [19], who conjectured that financial intermediaries emerge to resolve informational asymmetries. Although this reasoning was challenged by Campbell and Kracaw [7], Diamond's [10] formalization of it indicates that intermediation could lower signaling costs.[2] Recently, this line of enquiry has been further pursued by Boyd and Prescott (B-P) [5], whose work differs from ours in that they explain FFI's rather than IGA's.[3] Perhaps the paper most closely related to ours is Ramakrishnan and Thakor (R-T) [20] which also explains IGA's. There are three major differences between their paper and ours. (i) Unlike ours, R-T's intermediary is infinitely large. Thus, our analysis implies that competition-depleting mergers cannot be justified on the grounds that natural monopolies screen most efficiently. (ii) There is *no* information sharing in R-T. (iii) In *addition* to the noisy external monitoring that we assume, R-T allow for *perfect ex post* monitoring of its own members by the IGA. We avoid R-T's strong assumption by explicitly ruling out internal monitoring.

[1] See, e.g., Bhattacharya [4], Chan [8], and Diamond and Dybvig [11]. This literature avoids reliance on traditional arguments such as those based on transactions costs (see Benston [3]). For criticisms of these traditional arguments, see Fama [12].

[2] The major differences between Diamond's [10] analysis and ours are as follows. (i) Diamond's main point is that a diversified intermediary is useful even with risk neutrality because diversification lowers risk and thus signaling costs. Our point—that diversification augments screening cost—is the exact opposite. (ii) Diamond's intermediary is infinitely large whereas ours is finite. (iii) Diamond's model relies on the imposition of a nonpecuniary managerial penalty in the event of failure. We have no such penalty. (iv) There is *no* information sharing in Diamond. (v) Diamond assumes monitoring is costly but *perfect*. Thus, the moral hazard that we focus on—which arises solely because monitoring is noisy—is absent in Diamond.

[3] Other differences are as follows. (i) Because each depositor's payoff must be certain, the B-P [5] intermediary must be infinitely large or at least, quite large and perfectly diversified across projects. Our intermediary can be large or small. (ii) Because of their assumption that project, or agent, types are identical and independent draws, there is *no* information sharing in B-P.

I. The Basic Model of Capital Market Screening under Moral Hazard

Suppose there are firms that wish to sell new shares to raise investment capital. Let Γ denote the true aggregate value of a firm's new shares. In the spirit of asset pricing models, we assume that Γ depends on a variable, δ, that is specific to the firm, and a systematic (market) variable, ω. The link between Γ, δ, and ω is provided by some function, $g(\cdot, \cdot)$. That is,[4]

$$\Gamma = g(\delta, \omega).$$

Although the functional form of g is known, the variables δ and ω are *a priori* unknown to all. Adopting the Bayesian framework, we let the firm's management have prior probability density functions, D and B over δ and ω, respectively. That is, management treats $\tilde{\delta}$ and $\tilde{\omega}$ as random variables (tildes denote random variables) and computes the economic value of its new shares as

$$\Gamma_f = \int\int g(\tilde{\delta}, \tilde{\omega}) D(\tilde{\delta}) B(\tilde{\omega})\, d\tilde{\delta}\, d\tilde{\omega}$$

"Outsiders" are also unaware of the values of δ and ω, and their prior beliefs are condensed in the density functions, Δ and Ω, respectively. Thus, following Akerlof [1], the current market value of the firm's new shares, *given* the market's existing information structure, is

$$\Gamma_m = \int\int g(\tilde{\delta}, \tilde{\omega}) \Delta(\tilde{\delta}) \Omega(\tilde{\omega})\, d\tilde{\delta}\, d\tilde{\omega}.$$

The density functions, D and Δ, are assumed to be different to capture the idea that management's information about factors specific to the firm is usually superior to the market's.[5] If the firm and the market have the same priors about ω, then B and Ω will coincide. We keep them distinct for generality.

Precise information about δ and ω can be acquired by investing in information production. We assume the availability of an elastic supply of agents who can perform this task in exchange for compensation. Following Stiglitz [24], we label this activity "screening," and the agents who perform it "screening agents" (s.a.'s). If management is assumed to maximize current shareholders' wealth, then the

[4] In the context of the single factor asset pricing model of Ross [22], ω may be viewed as the expected return on a well-diversified (zero residual risk) portfolio and δ as the covariance of the firm's own return with the return on this diversified portfolio divided by the variance of the return on the diversified portfolio. In this case $g(\cdot, \cdot)$ will be linear in its argument. We do not impose any functional restrictions on $g(\cdot, \cdot)$, however.

[5] One may assume that management's information structure is *more informative* than the market's, in the sense of Blackwell [6] or more generally, Green and Stokey [13]. Alternatively, suppose $D_i(\delta)$ is the density function assigned to the i[th] firm's δ by its management. Let there be I firms in the market and assume that the market is aware of the set $\{D_i \mid i = 1, \cdots, I\}$, but cannot associate a given D_i with a particular firm. Let $f_1, \cdots, f_I$ be a sequence of real valued parameters satisfying

$$f_i \geq 0\ \forall i,\ \textstyle\sum_{i=1}^{I} f_i = 1.$$

Then, one can view the market's priors as being determined by the relation

$$\Delta(\delta) = \textstyle\sum_{i=1}^{I} f_i D_i(\delta).$$

That is, $\Delta(\delta)$ can be viewed as the density function associated with a *contagious* distribution.

management of a firm for which $\Gamma_f > \Gamma_m$ will negotiate a contract with an s.a. to have the firm screened. The s.a. can discover the firm's δ by choosing an appropriate action, e, from a feasible compact action space, E. And he or she can discover ω by choosing an additional action, α, from a feasible compact action space, A. The s.a.'s compensation for screening the firm is determined by a schedule, $\langle \phi \rangle$. This compensation is provided by the firm. Because we wish to focus on the incentives for the truthful production of information, rather than its accurate dissemination, ϕ is assumed to be *publicly* observable, and any secret side payments are ruled out.

All s.a.'s are assumed to be identical. Each has a von Neumann-Morgenstern utility function,

$$U: \mathbb{R} \rightarrow \mathbb{R}$$

over monetary wealth, where $\mathbb{R}$ is the real line. $U(\cdot)$ is bounded, $U'(\cdot) > 0$, and $U''(\cdot) < 0$, with primes denoting derivatives. Because $U(\cdot)$ is continuous, differentiable, and strictly increasing on $\mathbb{R}$, there exists an inverse function, $h(\cdot) \equiv U^{-1}(\cdot)$, that is itself differentiable everywhere with $h'(\cdot) > 0$, $h''(\cdot) > 0$. Moreover, each s.a. has a disutility for effort, represented by the function,

$$W: E \times A \rightarrow \mathbb{R}.$$

Thus, the s.a.'s net utility is given by the function,

$$Z(e, \alpha, \phi) = U(\phi) - W(e, \alpha), \quad \text{with} \quad W_e > 0,\ W_\alpha > 0.$$

Since E and A are compact, without loss of generality we can take $E \equiv [0, 1]$ and $A \equiv [0, 1]$. We assume that a choice of $e = 1$ enables a *precise* identification of δ and that a choice of $\alpha = 1$ leads to a *precise* identification of ω. But if $e \in [0, 1)$, the s.a.'s posterior density function over δ coincides with his prior density function, $\Delta(\cdot)$. Likewise, $\alpha \in [0, 1)$ implies that, unless an exogenous signal (to be discussed shortly) is received, the s.a.'s posterior density over ω is $\Omega(\cdot)$. Thus, choices of e and α other than 1 generate *no* new information.

The sequence of events leading to the discovery of a firm's true value is as follows. The firm's management designs an incentive contract, $\langle \phi \rangle$, and awards it to an s.a. The s.a. responds to the contract by first choosing either $e = 1$ or $e = 0$. (Note that since $e < 1$ does just as poorly as $e = 0$ and $W_e > 0$, if the s.a. does not choose $e = 1$, it will be optimal for him to choose $e = 0$). A choice of $e = 1$ permits the s.a. to learn the correct value of δ. After having chosen e, the s.a. receives, with probability λ, an exogenous signal that reveals to him the true value of ω. Receipt of this signal implies that the s.a. can choose $\alpha = 0$ and still be accurately informed about ω. With probability $1 - \lambda$, however, the signal (good news) will not be received and the s.a. will have to expend $\alpha = 1$ to learn ω.[6] This appears to be a reasonable way of modeling the cost of searching for

[6] An alternative interpretation of this favorable signal may be to view the s.a. as "lucky" if he receives the signal and "unlucky" if he does not. That is, the s.a. is entrusted with the task of discovering ω. If he is lucky and the pieces of the puzzle fall together right away, he can make this discovery at a very low personal cost to himself. If he is unlucky, it takes more effort to obtain the same information.

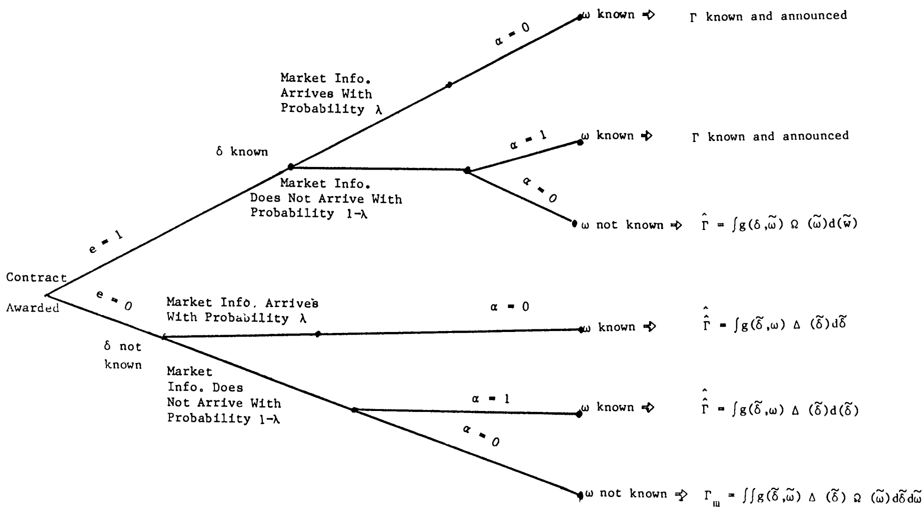

Figure 1. Sequence of Events in Screening

information, for it captures the notion that information gathering is subject to random shocks; an s.a. may be *ex ante* unaware of the exact expenditure of effort required to obtain a piece of information. Of course, one can also capture the uncertain cost of information acquisition by simply specifying a probability distribution around the cost component associated with α. That is, an s.a. who chooses $\alpha = 1$ (after choosing $e = 1$, for instance) bears a cost of $W(1, 0)$ with probability (w.p.) λ and a cost of $W(1, 1)$ w.p. $1 - \lambda$, where $W(1, 0) < W(1, 1)$. If one views the arrival (or nonarrival) of the signal as a state of nature, then in the context of agency models, the difference between this specification and ours is that in one case the agent chooses his action, α, before the state of nature is realized, and in the other (our specification), the agent's action, α, is chosen after the state is realized. Both types of models have been studied (e.g., Harris and Raviv [16] refer to our approach as Model 2 and the other specification as Model 1).[7] Although in the single s.a. case, the s.a.'s *expected* effort disutility (*prior* to the state realization) is identical with both approaches, in the case of IGA's, the two approaches have differing implications, as we discuss later. In either case, once δ and ω are known, the s.a. will employ the function $g(\cdot, \cdot)$ to compute Γ and announce it to the market. A graphic portrayal of these events is given in Figure 1.

It is obvious that $\langle \phi \rangle$ cannot be based on the firm value announced by the s.a., or else all s.a.'s will have an incentive to misrepresent firm values. The only

[7] Many accounting studies have used models similar to ours, the common feature being that the agent receives a signal subsequent to the contract agreement but prior to selecting his action. See, e.g., Baiman and Evans [2]. We should note, though, that our model is a little more complicated in that there are *two* state realizations and two sequential action choices; the first action choice (e) occurs before both states are realized, and the second action choice (α) occurs after the first state (good or bad signal about ω) is realized but before the second state (the result of a noisy monitor to be defined shortly) is realized.

economically sensible mechanism involves making $\langle \phi \rangle$ contingent *only* on e and α. This will work if e and α are observable *ex post*. Realistically, however, neither e nor α will be observable to anyone except the s.a. himself. Because s.a.'s are averse to choosing $e = \alpha = 1$, this lack of observability creates a moral hazard problem (Grossman and Hart [14], Harris and Raviv [16], Holmstrom [17, 18], Ross [21], and Shavell [23]). Every s.a. will find it privately optimal to choose $e = \alpha = 0$, pick any arbitrary numbers from the supports of $\Delta(\cdot)$ and $\Omega(\cdot)$ to use as δ and ω, respectively, and claim $e = \alpha = 1$.

To combat this moral hazard, $\langle \phi \rangle$ can be based on some informative, *ex post* monitor of e and α. This monitor is a function,

$$\Theta: E \times A \times \Xi \rightarrow \{0, 1\},$$

which "measures" the s.a.'s effort input. The set Ξ is the state space of the random variable, ξ, an exogenous source of noise in the *joint* measurement of e and α. Let the random variable, $t \in \{g, b\}$, identify the s.a.'s receipt of the exogenous signal about ω. The t for any s.a. is assumed to be stochastically independent of the t for any other s.a. That is, $t = g$ occurs w.p. λ and represents the state in which the signal is received. If $t = b$ (w.p. $1 - \lambda$), the signal is not received. The symbols g and b can be interpreted as representing "good" and "bad" states respectively. This specification is meant to reflect the notion that the intermediary can be "lucky" and hence be capable of procuring the necessary information at a low private cost. The probability mass function of the monitor Θ is given by

$$\begin{aligned}
\text{Prob}(\Theta = 1 \mid e = 1, \alpha = 1, t = b) &= \text{Prob}(\Theta = 1 \mid e = 1, \alpha = 1, t = g) \\
&= \text{Prob}(\Theta = 1 \mid e = 1, \alpha < 1, t = g) = p, \\
\text{Prob}(\Theta = 0 \mid e = 1, \alpha = 1, t = b) &= \text{Prob}(\Theta = 0 \mid e = 1, \alpha = 1, t = g) \\
&= \text{Prob}(\Theta = 0 \mid e = 1, \alpha < 1, t = g) = 1 - p, \\
\text{Prob}(\Theta = 1 \mid e = 1, \alpha < 1, t = b) &= \text{Prob}(\Theta = 1 \mid e < 1, \alpha = 1, t = b) \\
= \text{Prob}(\Theta = 1 \mid e < 1, \alpha < 1, t = b) &= \text{Prob}(\Theta = 1 \mid e < 1, \alpha = 1, t = g) \\
&= \text{Prob}(\Theta = 1 \mid e < 1, \alpha < 1, t = b) = q, \quad \text{and} \\
\text{Prob}(\Theta = 0 \mid e = 1, \alpha < 1, t = b) &= \text{Prob}(\Theta = 0 \mid e < 1, \alpha = 1, t = b) \\
= \text{Prob}(\Theta = 0 \mid e < 1, \alpha < 1, t = g) &= \text{Prob}(\Theta = 0 \mid e < 1, \alpha = 1, t = g) \\
&= \text{Prob}(\Theta = 0 \mid e < 1, \alpha < 1, t = b) = 1 - q,
\end{aligned}$$

with $p \in (0.5, 1)$ and $q \in (0.0.5]$.

Note that this monitor measures e and α *jointly* rather than separately. Thus, it is coarser than a monitor that provides individual assessments of e and α, but it is just as satisfactory if the object is to discover the overall firm value rather than its components. Also, since $p > 0.5$ and $q \leq 0.5$, we are assuming that the *ex post* verification is not just pure randomization. But since $q > 0$ and $p < 1$, the verification is also assumed to be noisy. Let S represent an s.a.'s reservation utility, i.e., that to be enjoyed in an alternative occupation.

Given the monitoring function, Θ, an s.a.'s compensation can be described by the following schedule,

$$\phi(\Theta) = \begin{cases} M & \text{if } \Theta = 1 \\ N & \text{if } \Theta = 0 \end{cases} \tag{1}$$

where M and N are dollar payoffs. Throughout, we shall let capitals denote monetary payoffs and lower case letters the corresponding utilities of the s.a. Thus, $U(M) = m$ and $U(N) = n$. A firm that desires to have itself screened must design $\phi(\Theta)$ to minimize the expected cost of screening subject to the individual rationality (I.R.) constraint that the s.a. earns at least his reservation utility, and the incentive compatibility (I.C.) constraint that the s.a.'s expected utility from choosing $e = \alpha = 1$ must be at least as great as his expected utility from choosing any other combination of e and α.

Because the marginal effect of an increase in either e or α on the probability of $\Theta = 1$ is zero if the increase is to *any* value in [0, 1), the s.a. will always pick an effort of zero *if* he or she does not find it optimal to pick an effort of one. This simplifies the analysis of the I.C. constraints. The s.a.'s expected utility from choosing $e = 1$ and $\alpha = 1$ (if $t = b$) is

$$pm + [1 - p]n - \lambda W(1, 0) - [1 - \lambda]W(1, 1). \tag{2}$$

The first I.C. constraint is that (2) should be at least as great as

$$qm + [1 - q]n - \lambda W(0, 0) - [1 - \lambda]W(0, 1), \tag{3}$$

which is the s.a.'s expected utility from choosing $e = 0$, and then $\alpha = 1$ if $t = b$ and $\alpha = 0$ if $t = g$. The second I.C. constraint is that (2) be at least as great as

$$qm + [1 - q]n - W(1, 0), \tag{4}$$

the s.a.'s expected utility from choosing $e = 1$, and then $\alpha = 0$, regardless of the realization of t. The third I.C. constraint is that (2) be at least as great as

$$qm + [1 - q]n - W(0, 0), \tag{5}$$

which is the s.a.'s welfare from choosing $e = \alpha = 0$. The final I.C. constraint is that

$$pm + [1 - p]n - W(1, 1) \geq qm + [1 - q]n - W(1, 0). \tag{6}$$

This constraint simply says that *after* the s.a. has chosen $e = 1$, he/she should not find it optimal to choose $\alpha = 0$ if $t = b$.

All of the above I.C. constraints can be condensed into a single constraint,

$$[p - q][m - n] \geq \{W(1, 1) - W(1, 0)\} \vee \{\lambda W(1, 0) + [1 - \lambda]W(1, 1)\}, \tag{6'}$$

upon employing the simplifying assumption that $W(1, 0) = W(0, 1)$ and $W(0, 0) = 0$. Notationally, "$\vee$" is the max operator. Throughout the subsequent analysis, we will assume that

$$\lambda[1 + \lambda]^{-1} \geq W(1, 0)W(1, 1)]^{-1}, \tag{6''}$$

so that (6′) becomes[8]

$$[p - q][m - n] \geq W(1, 1) - W(1, 0). \tag{7}$$

Note that (6″) is an assumption about the slope of the effort disutility function, $W(\cdot, \cdot)$. For a fixed λ, it says that $W(1, 1)$ is much greater than $W(1, 0)$. Thus, the loss in utility in going from the action pair (1, 0) to the action pair (1, 1) is relatively high, i.e., $W(\cdot, \cdot)$ is steeply convex in α. This condition will be useful later in establishing the value of an IGA relative to a market with individual s.a.'s. The intuition is as follows. The principal advantage of an IGA is that it permits member s.a.'s to share their information about ω, thereby avoiding duplication of effort. And the steeper the convexity of $W(\cdot, \cdot)$ in α, the greater is the benefit to an s.a. of choosing $\alpha = 0$ rather than $\alpha = 1$. Thus, the greater the difference between $W(1, 1)$ and $W(1, 0)$, the larger is the gain from avoiding $\alpha = 1$ (for a given s.a.) that is afforded by an IGA.

The I.R. constraint can be written as

$$pm + [1 - p]n - \lambda W(1, 0) - [1 - \lambda]W(1, 1) \geq S, \tag{8}$$

and thus, the firm's objective is to

$$\begin{aligned} &\text{minimize } ph(m) + [1 - p]h(n). \\ &(m, n) \in \mathbb{R}^2 \\ &\text{subject to (7) and (8).} \end{aligned} \tag{9}$$

This formulation assumes that one s.a. can screen only one firm.

We have not dealt with an individual s.a.'s potential inclination to *directly* appropriate the informational rents—by selling the information to other investors or by formulating a *personal* investment strategy—*subsequent* to discovering a firm's value and *prior* to publicly announcing it. The assumption is that such appropriation is impossible because the superior knowledge of informed traders is *instantaneously* reflected in equilibrium prices which are determined through a tatonnement process (see Grossman and Stiglitz [15]). Thus, information takes on the attribute of a public good that will not be purchased at a cost by investors who cannot protect it to extract the associated surplus. This implies that an s.a. must produce costly information *on contract* for a firm, rather than using it for personal portfolio restructuring. In this setting, we will show that an IGA is valuable because its contractual information production occurs at a lower expected cost.

In the remaining sections, our plan is as follows. We first assume that ω is known to all *ex ante.* This can be viewed as a situation in which the probability distribution of the "market" return is intertemporally unchanging, so that historical data allows a correct inference of the expected "market" return. We then solve the optimization problem above when s.a.'s work independently, and when two s.a.'s collaborate to pool their risky payoffs. A comparison of the two solutions reveals that the expected screening cost per firm (given by (9)) is lower with an independent s.a. than with an s.a. who is a member of an IGA. Thus, even though

[8] It is possible to characterize conditions under which intermediation is worthwhile even when (6″) is violated. Details of these conditions are available upon request.

s.a.'s are risk-averse and their individual payoffs are (conditionally) uncorrelated, diversification accompanying IGA formation is inadequate as a sole justification for IGA formation.

Next, we assume that ω is *a priori* unknown and that it will be revealed either through the receipt of an exogenous signal or through the choice of $\alpha = 1$ by the s.a. Under plausible conditions, the expected screening cost per firm is now lower with an IGA than with an independent s.a. The intuition is twofold. First, with an IGA, the odds are higher that *at least* one member s.a. will receive the costless signal that benefits the *whole* organization. Secondly, even if no s.a. receives the signal, an N-s.a. IGA can learn ω by choosing $\alpha = 1$ *in the aggregate*; assuming efforts can be pooled, this implies $\alpha = N^{-1}$ for each s.a. The idea is that commonly useful information can be fruitfully shared. Thus, our rationale for IGA formation requires that there be sufficiently large gains from information sharing *as well as* diversification.

Hence, the key to showing that IGA formation may be advantageous comes from the idea that the accompanying sharing of information decreases the s.a.'s expected disutility for effort and spreads the risk each s.a. bears due to the noise in the monitoring technology. The decrease in the expected effort disutility stems from two sources. First, consider the difference between an s.a.'s disutility for effort if the exogenous signal is received and his disutility for effort if the signal is not received. If this difference is large, information sharing within the IGA has a substantive impact on the s.a.'s expected effort disutility (making it significantly smaller). Thus, the net private gain to an s.a. from an increase in effort will be large (relative to the non-IGA case) at the margin. If the difference is small, however, information sharing within the IGA has a less significant impact on the s.a.'s expected effort disutility. Consequently, the net gain to an s.a. for an increase in effort at the margin will be relatively low, and *ceteris paribus*, it may not be large enough to compensate for the decline in the s.a.'s net private gain (from a marginal increase in effort) caused by the higher moral hazard associated with the IGA. Secondly, consider the probability, λ, that an individual s.a. will exogenously receive information about ω. If this probability is low, there is a relatively large impact of IGA formation on the probability that at least one member of the IGA will receive the exogenous signal. As a result, the net private gain to an s.a. from a marginal increase in effort is high and IGA formation is facilitated. Higher λ's make the IGA less attractive and in the limit, as we know from Section II, IGA formation is inimical to welfare if $\lambda = 1$. But λ cannot be too low either, for if $\lambda = 0$, then the entire burden for showing that an IGA is beneficial shifts to the effort disutility function. For example, an N-s.a. IGA faced with $\lambda = 0$ would presumably decide to have each s.a. choose $\alpha = 1/N$. Such an IGA would dominate N individual s.a.'s only if the IGA's total effort disutility, $N\ W(1, 1/N)$, was substantially lower than the sum of the effort disutilities of the N individual s.a.'s, $N\ W(1, 1)$. Assuming a positive λ adds to the benefit of an IGA and weakens the restrictions one must place on $W(\cdot, \cdot)$. In fact, for a two-s.a. IGA, the probability of the IGA receiving the favorable signal is higher by an amount $\lambda - \lambda^2$ than the probability that an individual s.a. operating alone will get the signal, and this difference is maximized at $\lambda = 1/2$.

The above discussion should clarify why we have modeled the uncertain cost of acquiring information about ω as a sequential "signal arrival-action choice" process. If we just specify a distribution around the cost component associated with α (the s.a. does not know t before choosing α), then every s.a. in the IGA must choose $\alpha = 1$ (in a symmetric Nash equilibrium). In this case, despite an uncertain information acquisition cost, there is no *probabilistic* information sharing gain from IGA formation, and we are back to the case in which $\lambda = 0$. The sequential process allows the IGA to observe the signal *before* deciding on its choice of α, and this introduces an added benefit of group formation, namely an increase in the probability of receiving the signal.

The reader may question why both types of uncertainty—δ and ω—are needed. The reason is that in the absence of either uncertainty, the problem at hand is trivialized. If only δ is present, then there is no commonality in the factors determining firm values. This case is analytically identical to the one in which ω is common knowledge, and our analysis in Section II indicates that IGA formation in this case is counterproductive. On the other hand, if only ω is present, then there is either no idiosyncratic aspect to any firm's value (every firm has the same value) or every firm's idiosyncratic component is common knowledge. In this case an IGA is unnecessary; one s.a. can be employed to discover ω and thus identify the true values of all firms in one screening. [The analogy of δ in the single factor CAPM context would be a firm's "beta." Thus, to say that δ is absent is to assert that all "betas" are one, and to say that δ is known is to assert that every "beta" is known. In either case, just one s.a. would be needed to discover ω (the expected market return) and the process would be over.]

This notion of information sharing has an appealing institutional counterpart. Real world IGA's devote significant resources to gathering economy-wide information that is commonly relevant. For instance, a credit-rating agency assessing numerous municipal bonds will acquire the information needed for a reliable interest rate forecast, and this information will be useful in rating all the bonds. In addition, special information about demographic trends within the municipal unit will be needed to identify factors idiosyncratic to the issuer.

II. IGA Formation without Information Sharing

In what follows, each s.a.'s ϕ is uncertain because it is based on a noisy monitor. Of course, institutional reality is that there is little uncertainty associated with an IGA's compensation from its clients. Explicit monitoring is unnecessary in practice, perhaps because of the effectiveness of reputation effects in ensuring information reliability. We rely on explicit monitoring because our single period framework precludes formal modeling of reputation. The first step in examining the impact of IGA formation without information sharing is to derive the optimal incentive contract for a single s.a. when every s.a. operates independently.

LEMMA 1. *Suppose that ω is a priori known to all and that s.a.'s operate independently. Then, the optimal incentive contract which a firm must offer an s.a. is*

$$\phi^0(\Theta) = \begin{cases} h(S + [1 - q][p - q]^{-1}W(1, 0)) & \text{if } \Theta = 1 \\ h(S - q[p - q]^{-1}W(1, 0)) & \text{if } \Theta = 0. \end{cases}$$

Now suppose two s.a.'s form an IGA. (This is not meant to imply we consider only two-s.a. IGA's. The implications of larger groupings are considered later.) Each will screen one distinct firm and receive a stochastic payoff from *that* firm. Assume that the Θ's for the two s.a.'s are stochastically independent. The s.a.'s will pool their risky payoffs together. And, because they are identical, they will share the pool fifty-fifty (see Wilson [26]). We now examine the potential advantages of IGA formation. Note that our initial size assumption—that the formation of an IGA involves going from one s.a. to two s.a.'s—accentuates the benefits of diversification of ϕ risk; this pooling benefit of further merger will not be particularly large for an IGA that already has many clients. Thus, if we show that even the maximum diversification benefit is insufficient to justify two-s.a. IGA's, then larger IGA's can be ruled out. In what follows, we assume that s.a.'s can pool their efforts to produce information and that individual effort inputs are additive.

PROPOSITION 1. *Suppose that ω is known to all ex ante and that two s.a.'s form an IGA. Then, the expected screening cost for each firm contracting with an s.a. from this IGA is always higher than the expected screening cost it would incur if the s.a.'s functioned independently.*

The problem with payoff pooling here is another example of the basic difficulty in enforcing desired behavior from individual agents who are members of groups (see Holmstrom [18]). As long as "budget balancing"—the sum of the individual s.a.'s payoffs must equal the total payoff of the IGA—is necessary, the imperfection in verifying actions *ex post* creates an externality. Both s.a.'s can conceal improper actions behind the uncertainty concerning who was at fault, since *both* cannot be penalized sufficiently in the final outcome, in the absence of *additional* internal monitoring by the IGA itself. We wish to emphasize that this result is *not* merely an artifact of the *specific* signal distribution employed here. The point is that since group formation permits the s.a.'s to share payoff risk, it also allows each s.a. to share the cost of choosing an *ex post* unverifiable low effort, and this creates a free rider problem. This problem will arise in our context whenever the *ex post* signal of the s.a.'s effort choice is noisy. Holmstrom [18] establishes a similar result when the only available *ex post* noisy signal is the *aggregate* output attributable to the group of agents. There is, however, an important distinction between his finding and ours. In Holmstrom's [18] model, the problem is that there is only a single summary statistic, namely the joint output, that is available to compensate *all* agents. In our model, individual output statistics *are* available, but in equilibrium are discarded by the group in favor of the single statistic, namely the total output, because not doing so effectively "decouples" the intermediary and is not in the group's private interest.

III. IGA's with Information Sharing

A. The Main Results

The restrictive assumption that ω is known *ex ante* is now dropped. In exploring the potential benefits of IGA formation, we first assume that s.a.'s operate independently. The optimal incentive contract which a firm should offer an s.a. is a solution to (9) subject to (7) and (8), and is characterized in the following lemma.

LEMMA 2. *Suppose that* ω *is a priori unknown to all and that s.a.'s operate independently. Then, the optimal incentive contract which an s.a. must be offered is*

$$\phi(\Theta) = \begin{cases} h(S + \zeta_1 + \zeta_2[1 - p][p - q]^{-1}) & \text{if } \Theta = 1 \\ h(S + \zeta_1 - \zeta_2 p[p - q]^{-1}) & \text{if } \Theta = 0 \end{cases}$$

where

$$\zeta_1 \equiv \lambda W(1, 0) + [1 - \lambda]W(1, 1)$$

$$\zeta_2 \equiv W(1, 1) - W(1, 0).$$

Under these conditions, an IGA formed by two s.a.'s must determine an effort allocation strategy to cope with the situation in which *neither* s.a. receives a signal that reveals ω. Since both s.a.'s are identical, and since individual effort inputs can be additively pooled, the simplest arrangement would be to choose $\alpha = 1/2$ for each s.a. But since effort inputs are unobservable, neither s.a. can verify *ex post* whether the other has kept his promise. This implies that firms must design incentive contracts in such a way that, in a Nash equilibrium, it is *privately* optimal for each s.a. to choose $\alpha = 1/2$.

The probability that at least one of the s.a.'s in the IGA will receive the exogenous signal about ω is $2\lambda - \lambda^2$. Hence, each firm designs $\langle\phi\rangle$ to

$$\underset{(m,\,k,\,n) \in \mathbb{R}^3}{\text{minimize}}\; p^2h(m) + 2p[1 - p]h(k) + [1 - p]^2h(n). \tag{10}$$

Subject to

$$\text{I.R.: } p^2m + 2p[1 - p]k + [1 - p]^2n \geq S + \zeta_3 \tag{11}$$

$$\text{I.C.: } p[m - k] + [1 - p][k - n] \geq \zeta_4 \tag{12}$$

$$\text{I.S.: } h(m) + h(n) \equiv 2h(k) \tag{13}$$

where

$$\zeta_3 \equiv [2\lambda - \lambda^2]W(1, 0) + [1 - 2\lambda + \lambda^2]W(1, 1/2)$$

and

$$\zeta_4 \equiv [p - q]^{-1}[W(1, 1/2) - W(1, 0)].$$

Again, in this formulation there are numerous I.C. constraints besides (12). But if one assumes, as we have implicitly done above, that

$$[2\lambda - \lambda^2][1 + 2\lambda - \lambda^2]^{-1} \geq W(1, 0)[W(1, 1/2)]^{-1}, \tag{14}$$

then all the other constraints are automatically satisfied when (12) holds.[9] The optimal solution to this problem is given in the lemma below.

LEMMA 3. *Suppose that ω is a priori unknown to all and that s.a.'s operate independently. Then, the optimal incentive contract which an s.a. must be offered is*

$$m^{**} = p^{-1}\{S + \zeta_3 + \zeta_4[1 - p] - [1 - p]k^{**}\} \tag{15}$$

$$n^{**} = [1 - p]^{-1}\{S + \zeta_3 - \zeta_4 p - pk^{**}\} \tag{16}$$

with k^{**} chosen to satisfy (13).

To emphasize the diversification benefits of an IGA, we assume

$$pW(1, 0) + [1 - p]W(1, 1) \geq W(1, \tfrac{1}{2}). \tag{17}$$

This condition puts an upper bound on p for fixed values of $W(\cdot, \cdot)$. Let $\bar{p} = \max\{p \mid (17) \text{ is satisfied}\}$. Then, this requirement calls for $p \in (0.5, \bar{p}]$. This condition plays an important role in establishing the value of an IGA in the next proposition, and it is easy to see why. The smaller the value of p, the greater is the diversification benefit yielded by an IGA. Thus, a rationale for IGA formation can be expected to rely on p not being very high.

PROPOSITION 2. *Suppose that ω is unknown to all ex ante and that two s.a.'s form an IGA. Assume that* (6″), (14), *and* (17) *hold. Then, the formation of an IGA results in a strict Pareto improvement in the sense that the expected screening cost for each firm contracting with the IGA is lower than the expected cost it would incur by transacting with an independent s.a.*

The role of the probability, $1 - p$, that an s.a.'s action choice will be misidentified, in the establishment of this result is worth noting. The higher this probability, the greater the chance that an s.a. who expends the necessary effort will not be adequately rewarded. IGA formation permits the s.a. to lay off a portion of this risk. Thus, the net private gain to an s.a. from a marginal increase in effort is improved by the IGA.

The preceding analysis seems to contradict B-P's [5] finding that FFI's can emerge *without* information sharing. In our model, the problem of increased moral hazard accompanying group formation arises because outsiders are unable to directly observe whether an individual s.a. has chosen $\alpha = 1$ or $\alpha = 0$. Hence, the offsetting benefit of information sharing is necessary. In B-P [5], the

[9] Like (6″), (14) is also an assumption about the convexity of $W(\cdot, \cdot)$. Intermediation can again be shown to be of value even if this condition is not met. When combined with the earlier condition, the assumption is that

$$W(1, 0) \leq f'W(1, \tfrac{1}{2}) \wedge fW(1, 1)$$

where "$\wedge$" is the min operator and

$$f' \equiv [2\lambda - \lambda^2][1 + 2\lambda - \lambda^2]^{-1},$$

$$f \equiv \lambda[1 + \lambda]^{-1}.$$

equilibrium is sustainable—even without information sharing—because the *act of evaluation* (the equivalent of our $\alpha = 1$) is *costlessly observable by all.* Thus, moral hazard of the type we focus on is ruled out by assumption in B-P [5].

B. Optimal Intermediary Size

We have only considered two-member IGA's. The connotations of larger IGA's are now examined in the context of the endogenous determination of optimal IGA size. The addition of more s.a.'s to the IGA has two diametrically opposite effects. On the one hand, it enhances the likelihood of the IGA receiving the exogenous signal about ω, and on the other hand, it worsens the moral hazard problem. If λ is very close to one or zero *and* $W(\cdot, \cdot)$ has insufficient convexity, IGA's may never form. But if one assumes an "appropriate" λ and a steep and convex $W(\cdot, \cdot)$, competitive market forces will bring two or more s.a.'s together. Optimal IGA size will then be determined endogenously by the relative magnitudes of the parameters in the problem. It will be *finite*, however. The reason is simple. An intermediary consisting of J members faces a probability of $\sum_{j=1}^{J} \binom{J}{j} \lambda^j (1 - \lambda)^{J-j}$ of receiving information about ω exogenously. This probability is an increasing function of J, but is bounded from above by (and asymptotically approaches) 1. Thus, as an IGA grows, the value of joint information increases, but this value has an upper bound, and increases in the group size beyond some point will have diminishing marginal value. But the increased screening cost linked to the moral hazard created by IGA formation rises monotonically with group size and has no upper bound. Hence, a point will be reached at which the marginal increase in the value of joint information exactly equals the marginal increase in the cost of moral hazard, and at this point further growth in size is unproductive.

IV. Concluding Remarks

We have provided a rationale for the existence of financial information gathering agencies that acquire and process information (but do not fund), such as credit rating agencies (like Moody's, Standard and Poor's, and Barron's). These institutions function primarily to certify the values of economic entities that approach them.[10] Our theory does not deal with institutions like commercial banks which intermediate more directly by holding assets, the values of which are not always publicly known.

Existing theories of financial intermediation based on the assumption of informational asymmetry in the capital market have overlooked the importance of information sharing. Asset pricing models have taught us that the values of firms are driven by *common* uncertainties that affect the economy as a whole. Hence, it should be possible for those who produce information about firm values

[10] In this paper, we have ruled out the possibility of side payments that Campbell and Kracaw [7] discuss in detail. We agree with B-P [5] that the role of potential side payments in a theory of financial intermediation may not be a very important one. See their paper for an extended discussion of this issue.

to gain by joining together and sharing information about common uncertainties like the return on the market portfolio, for example. A feature of our analysis that makes it distinct from previous efforts is that the notion of information sharing is assigned a pivotal role in our model.

The reader must have noted an "asymmetry" in our modeling. We assume that a random shock affects the s.a.'s private cost of learning the value of ω, but there is no randomness in the cost attached to discovering δ. Realistically, of course, randomness should be introduced in the private costs associated with *both* ω and δ. We have not done this, however, because it has an unimportant effect on the results. Since δ is purely idiosyncratic, there is no sharing of information possible. Hence, intermediation will not contribute to a reduction in the expected cost of learning δ.

Exploring the existence of financial institutions that produce costly information but do not fund may be useful in understanding more than conventional IGA's like rating agencies. Recent changes in the regulatory environment appear to have spurred a nonfunding trend among many commercial banks which seem to find it profitable to avoid funding loans with deposits and act instead as middlemen underwriting and "certifying" loan qualities for a fee. The following excerpt from the Wall Street Journal [Wednesday, February 1, 1984, p. 25] is illuminating:

> A fundamental change is taking place in the way some of the biggest banks operate in the U.S. and overseas.
>
> For years, banking giants have rushed to build huge portfolios of loans and other assets, bragging as they grow bigger and bigger. But now profits in many traditional loans are falling, U.S. regulators are pressing banks to increase their capital and the international debt crisis has made banks concerned about keeping big chunks of foreign debt.
>
> So some U.S. banks are switching gears and acting more as middlemen. They write big loans for domestic and overseas customers, then quickly parcel them out to smaller banks, thrifts and international investors.

Thus, transitions of intermediaries from one class to the other cannot be ruled out; many of today's banks may become pure "information processors" and insurance agents, with no conventional deposit liabilities or loan portfolios, and may thus look a lot like the intermediary we have portrayed.

Appendix

PROOF OF LEMMA 1. The assumption that ω is known to all *ex ante* is equivalent to assuming that $\lambda = 1$. With this, the problem is to minimize (9) subject to the I.R. constraint,

$$pm + [1 - p]n \geq S + W(1, 0), \tag{A1}$$

and the I.C. constraint,

$$[p - q][m - n] \geq W(1, 0). \tag{A2}$$

The solution to this problem must satisfy the first order necessary conditions for optimality, given by

$$h'(m^0) = \mu + \psi[p - q]p^{-1} \tag{A3}$$

and

$$h'(n^0) = \mu - \psi[p - q][1 - p]^{-1} \tag{A4}$$

where μ and ψ are the Lagrange multipliers attached to the constraints (A1) and (A2), respectively. Suppose $\mu \leq 0$. Then, if $\psi \geq 0$, we have $h'(n) < 0$, which is impossible. And, if $\psi < 0$, then $h'(m) < 0$, which is also impossible. Thus, $\mu > 0$.

Next, combining (A3) and (A4) yields

$$h'(m^0) - h'(n^0) = \psi[p - q]\{p[1 - p]\}^{-1}. \tag{A5}$$

Constraint (A2) requires that $m^0 > n^0$. Because $h(\cdot)$ is strictly increasing, we can conclude from (A5), therefore, that $\psi > 0$. Hence, both multipliers are strictly positive, implying that (A1) and (A2) hold as equalities. Solving them as simultaneous equations provides the optimal solutions,

$$m^0 = S + [1 - q][p - q]^{-1}W(1, 0) \tag{A6}$$

and

$$n^0 = S - q[p - q]^{-1}W(1, 0). \tag{A7}$$

Q.E.D.

PROOF OF PROPOSITION 1. Each s.a. within the IGA will receive a payoff described by (1), and the Θ's for the two s.a.'s are stochastically independent. With payoff pooling and equal sharing, each s.a.'s *effective* compensation will be

$$\begin{array}{lll} M & \text{if} & \Theta = 1 \text{ for both s.a.'s} \\ [M + N]/2 & \text{if} & \Theta = 1 \text{ for only one s.a.} \\ N & \text{if} & \Theta = 0 \text{ for both s.a.'s.} \end{array}$$

Define $K \equiv [M + N]/2$. If both s.a.'s choose $e = 1$, the probability of each receiving M is p^2, the probability of each receiving K is $2p(1 - p)$, and the probability of each receiving N is $(1 - p)^2$. Each contracting firm will recognize that payoffs are being pooled and thus design ψ to

$$\begin{aligned} &\text{minimize } p^2h(m) + 2p[1 - p]h(k) + [1 - p]^2h(n) \\ &(m, k, n) \in \mathbb{R}^3 && \text{(A8)} \\ &\text{subject to} \\ &\text{I.R.: } p^2n + 2p[1 - p]k + [1 - p]n^2 \geq S + W(1, 0) && \text{(A9)} \\ &\text{I.C.: } p[m - k] + [1 - p][k - n] > W(1, 0)[p - q]^{-1} && \text{(A10)} \\ &\text{I.S.: } h(m) + h(n) \equiv 2h(k) && \text{(A11)} \end{aligned}$$

where $k = U(K)$ and I.S. stands for "identity satisfaction." Note that (A10) is a Nash equilibrium constraint. It states that, contingent on the assumption that one s.a. is choosing $e = 1$, the other s.a. must find it privately optimal to also choose $e = 1$.

Let asterisks denote the optimal solution in this case. This solution must satisfy the first-order optimality conditions. Thus, we have

$$h'(m^*) = \mu + \psi p^{-1} \tag{A12}$$

$$h'(k^*) = \mu - \psi[2p - 1][2p\{1 - p\}]^{-1} \tag{A13}$$

and

$$h'(n^*) = \mu - \psi[1 - p]^{-1} \tag{A14}$$

where we have substituted (A11) into (A8) and defined μ and ψ as the Lagrange multipliers for (A9) and (A10), respectively. Suppose $\mu < 0$. Then, if $\psi \leq 0$, we get $h'(m^*) < 0$, and if $\psi > 0$, we get $h'(k^*) < 0$. Neither is possible. Hence, $\mu > 0$. Now suppose $\psi \leq 0$. Then, since $p > 0.5$, we obtain $h'(n^*) \geq h'(k^*) \geq h'(m^*)$. But this violates (A10). We can conclude, therefore, that $\psi > 0$. Since both multipliers are strictly positive, (A9) and (A10) must hold as equalities. Treating them as such provides the following optimal solution:

$$m^* = p^{-1}\{S + W(1, 0)[1 - q][p - q]^{-1} - k^*[1 - p]\}, \tag{A15}$$

and

$$n^* = [1 - p]^{-1}\{S - W(1, 0)q[p - q]^{-1} - k^*p\}, \tag{A16}$$

with k^* chosen to satisfy (A11). To complete the proof, we need to show that the expected screening cost resulting from using the contract in (A15) and (A16) exceeds that resulting from using the contract in (A6) and (A7). To see this, note that

$$\begin{aligned} &p^2h(m^*) + 2p[1 - p]h(k^*) + [1 - p]^2h(n^*) \\ &= p[h(m^*) + [1 - p]h(k^*)] + [1- p][\{1 - p\}h(n^*) + ph(k^*)] \\ &> p[h(pm^* + [1 - p]k^*)] + [1 - p][h([1 - p]n^* + pk^*)] \\ &= ph(m^0) + [1- p]h(n^0), \end{aligned}$$

where the penultimate step follows from Jensen's inequality (see de Groot [9]). Q.E.D.

PROOF OF LEMMA 2. The proof is similar to that of Lemma 1. The Lagrange multipliers attached to (7) and (8) can be shown to be strictly positive. Solving (7) and (8) as simultaneous equations yields

$$\hat{m} = S + \zeta_1 + \zeta_2[1 - p][p - q]^{-1}, \tag{A17}$$

and

$$\hat{n} = S + \zeta_1 - \zeta_2 p[p - q]^{-1}, \tag{A18}$$

completing the proof. Q.E.D.

PROOF OF LEMMA 3. Familiar arguments can be repeated to prove that the

Lagrange multipliers associated with (11) and (12) are strictly positive. Viewing (11) and (12) as simultaneous equations and solving them gives the desired result. Q.E.D.

PROOF OF PROPOSITION 2. The optimal solution for the non-IGA case is $\{\hat{m}, \hat{n}\}$, with these variables defined in (A17) and (A18), respectively. Let $\{\hat{m} - \varepsilon, \hat{k}, \hat{n}\}$ be a candidate solution for the two-s.a. IGA, where $\varepsilon > 0$ is a very small positive scalar. Since this solution must satisfy (13), we take $2h(\hat{k}) = h(\hat{m} - \varepsilon) + h(\hat{n})$. Thus, by the convexity of $h(\cdot)$, we must have

$$\hat{k} > [\hat{m} - \varepsilon + \hat{n}]/2. \tag{A19}$$

We will first show that the proposed solution is feasible. Define $\hat{\hat{m}} = \hat{m} - \varepsilon$ and note that

$$\begin{aligned} p^2\hat{\hat{m}} &+ 2p[1 - p]\hat{k} + [1 - p]^2\hat{n} \\ &> p^2\hat{\hat{m}} + 2p[1 - p][\hat{\hat{m}} + \hat{n}]/2 + [1 - p]^2\hat{n}. \\ &= p\hat{\hat{m}} + [1 - p]\hat{n}. \end{aligned} \tag{A20}$$

Also

$$\begin{aligned} S + \zeta_3 &< S + \lambda W(1, 0) + [1 - \lambda]W(1, 1) \\ &= p\hat{m} + [1 - p]\hat{n}, \end{aligned} \tag{A21}$$

since the I.R. constraint, (8), is binding at the optimum. Combining (A20) and (A21), we see that for ε close enough to zero,

$$p^2\hat{\hat{m}} + 2p[1 - p]\hat{k} + [1 - p]^2\hat{n} > S + \zeta_3.$$

Thus, the candidate solution satisfies the I.R. constraint (11).

For the non-IGA solution we know that, at the optimum, the I.C. constraint (7) holds as an equality. That is,

$$[\hat{m} - \hat{h}][p - q] = W(1, 1) - W(1, 0). \tag{A22}$$

Now, since $k \in (n, m)$ for any convex $h(\cdot)$, we know that

$$\inf_{k \in (n,m)} \{p[m - k] + [1 - p][k - n]\} = [1 - p][m - n]. \tag{A23}$$

It is easy to see that if (17) holds, then (A22) implies that

$$[1 - p][\hat{m} - \hat{n}][p - q] \geq W(1, ½) - W(1, 0),$$

which means

$$\{p[\hat{m} - \hat{k}] + [1 - p][\hat{k} - \hat{n}]\}[p - q] > W(1, ½) - W(1, 0).$$

Thus, for ε small enough, we have

$$\{p[\hat{\hat{m}} - \hat{k}] + [1 - p][\hat{k} - \hat{n}]\}[p - q] \geq W(1, ½) - W(1, 0).$$

Hence, (12) holds. Because $\hat{k}$ was explicitly chosen to meet (13), the candidate solution is feasible.

The expected screening cost induced by the candidate contract is (upon substituting (13) in (10))

$$ph(\hat{\hat{m}}) + [1 - p]h(\hat{n}) < ph(\hat{m}) + [1 - p]h(\hat{n})$$

$$= \text{expected screening cost for a single agent.}$$

Since the optimal contract with an IGA, $\{m^{**}, k^{**}, n^{**}\}$, must do at least as well as the arbitrary contract, $\{\hat{\hat{m}}, \hat{k}, \hat{n}\}$, we are done. Q.E.D.

REFERENCES

1. G. A. Akerlof. "The Market for 'Lemons': Qualitative Uncertainty and the Market Mechanism." *Quarterly Journal of Economics* 84 (August 1970), 488–500.
2. S. Baiman and J. H. Evans. "Pre-Decision Information and Participative Management Control Systems." *Journal of Accounting Research* 21 (Autumn 1983), 371–95.
3. G. Benston. "A Transactions Cost Approach to the Theory of Financial Intermediation." *The Journal of Finance* 31 (May 1976), 215–31.
4. S. Bhattacharya. "Aspects of Monetary and Banking Theory and Moral Hazard." *The Journal of Finance* 37 (May 1982), 371–84.
5. J. Boyd and E. Prescott. "Financial Intermediaries." Working Paper, 272, Federal Reserve Bank of Minneapolis, March 1983.
6. D. Blackwell. "Comparison of Experiments." In J. Neyman (ed.), *Second Berkeley Symposium on Mathematical Statistics and Probability.* CA: University of California Press, 1970.
7. T. Campbell and W. Kracaw. "Information Production, Market Signalling and The Theory of Financial Intermediation." *The Journal of Finance* 35 (September 1980), 863–82.
8. Y. Chan. "On the Positive Role of Financial Intermediation in Allocation of Venture Capital in a Market with Imperfect Information." *Journal of Finance* (December 1983), 1543–68.
9. M. De Groot. *Optimal Statistical Decisions.* NY: McGraw-Hill, 1970.
10. D. Diamond. "Delegated Monitoring and Financial Intermediation." *Review of Economic Studies* LI (July 1984), 393–415.
11. ——— and P. Dybvig. "Deposit Insurance, Liquidity and Bank Runs." *Journal of Political Economy* 91 (June 1983), 401–19.
12. E. F. Fama. "Banking in the Theory of Finance." *The Journal of Monetary Economics* 6 (January 1980), 39–51.
13. J. Green and N. Stokey. "Two Representations of Information Structures and Their Comparisons." Unpublished manuscript, Harvard University and Northwestern University, 1979.
14. S. J. Grossman and O. D. Hart. "An Analysis of The Principal-Agent Problem." *Econometrica* 51 (January 1983), 7–46.
15. S. J. Grossman and J. E. Stiglitz. "Information and Competitive Price Systems." *American Economic Review* 66 (May 1976), 246–53.
16. M. Harris and A. Raviv. "Optimal Incentive Contracts with Imperfect Information." *Journal of Economic Theory* 20 (April 1979), 231–59.
17. B. Holmstrom. "Moral Hazard and Observability." *The Bell Journal of Economics and Management Science* 10 (Spring 1979), 74–91.
18. ———. "Moral Hazard in Teams." *The Bell Journal of Economics and Management Science* 13 (Autumn 1982), 324–40.
19. H. Leland and D. Pyle. "Informational Asymmetries, Financial Structure and Financial Intermediation." *The Journal of Finance* 32 (May 1977), 371–87.
20. R. T. S. Ramakrishnan and A. V. Thakor. "Information Reliability and a Theory of Financial Intermediation." *Review of Economic Studies* LI (July 1984), 415–32.
21. S. A. Ross. "The Economic Theory of Agency: The Principal's Problem." *American Economic Review* 63 (May 1973), 134–39.

22. ——— "The Arbitrage Theory of Capital Asset Pricing." *Journal of Economic Theory* 13 (December 1976), 341–60.
23. S. Shavell. "Risk Sharing and Incentives in the Principal and Agent Relationship." *The Bell Journal of Economics and Management Science* 10 (Spring 1979), 55–73.
24. J. E. Stiglitz. "The Theory of 'Screening,' Education and the Distribution of Income." *The American Economic Review* 71 (June 1975), 283–300.
25. W. K. Viscusi. "A Note on 'Lemons' Markets with Quality Certification." *The Bell Journal of Economics and Management Science* 9 (Spring 1978), 277–80.
26. R. Wilson. "The Theory of Syndicates." *Econometrica* 36 (January 1968), 119–132.