CORPORATE DIVERSIFICATION AND AGENCY

Benjamin E. Hermalin
University of California at Berkeley

Michael L. Katz
University of California at Berkeley

10 November 1996

Abstract

Firms undertake a variety of actions to reduce risk through diversification, including entering diverse lines of business, taking on project partners, and maintaining portfolios of risky projects such as R&D or natural resource exploration. By a well-known argument, securities holders do not directly benefit from risk-reducing corporate diversification when they can replicate this diversification on their own. Moreover, shareholders should be risk neutral with respect to the unsystematic risk that is associated with many research projects. Some have argued that corporate risk reduction may be of value, or can otherwise be explained by, the agency relationship between securities holders and managers. We argue that the value of diversification strategies in an agency relationship derives not from its effects on risk, but rather from its effects on the principal's information about the agent's actions. We demonstrate by example that diversification activities may increase or decrease the principal's information, depending on the particular structure of the activity.

Keywords: diversification; principal-agent relationship.

This material is based on work supported by the Berkeley Program in Finance, the Olin Foundation, and the National Science Foundation (the latter under Award No. SES-9112076).

1. INTRODUCTION

A frequently stated motive for joint ventures, conglomerate mergers, and investments in new lines of business is a desire to diversify a firm and reduce the riskiness of its returns. According to this view, joint ventures spread the risk of major projects, while conglomerate diversification creates portfolio benefits by pooling uncorrelated returns. A similar logic may drive a corporate decision to pursue simultaneously a variety of different R&D strategies aimed at a common goal.[1] A standard argument, however, suggests that with perfect capital markets there is no financial value to within-firm diversification because investors could instead diversify their own portfolios (Alberts [1966] and Levy and Sarnat [1970]). Indeed, it may be worse to have the firms themselves diversify because this reduces the number of pure securities that are traded.[2]

One can construct a model in which diversification generates financial benefits by reducing the costs of bankruptcy, but (as we discuss further in the conclusion to this paper) doing so appears to be surprisingly delicate. One might also argue that economies of scale in making investment decisions generate a reason to create diversified firms. But—even if such economies exist—they explain only the existence of organizations like mutual funds, not companies that exercise direct control over the business decisions of several different operational divisions.[3]

Why, then, do firms undertake activities to diversify risk? Many analysts have suggested that the answer lies with the separation of ownership and control and the resulting agency

[1] Of course, the firm also could be driven to pursue a variety of positive net present value projects where the extensive margin is greater than the intensive margin.

[2] Recently, for example, both the Pacific Telesis Group and the Litton Corporation decided to split themselves into separately traded divisions in part for this reason. Other firms have created separate classes of stock whose returns are tied to the performance of specific divisions within the firm. Porter [1987] also makes this point from the perspective of corporate strategy.

[3] Errunza and Senbet [1984] argue that, in some developing countries, foreign direct investment may be needed to fill in incomplete capital markets.

relationship that exists between managers and outside investors. The building blocks for an agency-based explanation of diversification are these. Risk-averse owners will be motivated to hold diversified portfolios and thus (to a first-order approximation) behave in a risk-neutral fashion with respect to the firm's investment decisions.[4] Risk-averse managers also would like to diversify away the risks associated with their firms, but for a variety of reasons typically are unable to do so. One, to ensure that a manager has incentives to serve the owners' interests, her current income stream may be tied in an undiversified way to the performance of the firm that employs her. Two, through the effects on her reputation, a manager's future income may depend on the returns of the one firm that she currently manages. It is difficult to imagine managers taking multiple part-time jobs in order to diversify their human capital. Lacking the opportunity to diversify in other ways, managers may value within-firm diversification.

Viewed in this light, there are two possible versions of the claim that firm diversification is driven by agency considerations. The first is that, *given the managerial compensation scheme in place,* managers may pursue their own self-interest in risk reduction, even if doing so comes at security holders' expense. There is a sizable empirical literature that examines whether managers pursue acquisitions, conglomerate and otherwise, at shareholders' expense (e.g., Amihud and Lev [1981], Amihud *et al.* [1986], and Lewellen *et al.* [1989] focus on risk-reduction motives, while Morck *et al.* [1990] and the references cited therein consider acquisitions driven by managerial interests more generally). The empirical findings with respect to whether there is a conflict between managerial desires to reduce risk and shareholder interests is mixed. Moreover, the logic underlying this debate is seriously incomplete. The relation between the riskiness of a

[4] We do not explore the possibility that a "large shareholder" is needed to exercise control of the firm and this

firm's returns and its manager's well-being is largely endogenous: It is driven by the manager's compensation scheme.[5] Consequently, cross-sectional studies that take managerial compensation schemes (such as the degree of stock ownership) as exogenous are seriously flawed.

Even if the relation between the riskiness of a firm's returns and it manager's well-being is not entirely endogenous (*e.g.*, reputational concerns arise), it does not necessarily follow the manager will prefer to diversify. Hermalin [1993] has shown that a manager whose future income is tied to the labor market's posterior estimate of her abilities could rationally prefer that her firm's returns be as risky as possible in order to reduce the informativeness of these returns as signals of her underlying ability. The risk to the manager's reputation is diminished, even though current returns are riskier.

The second version of the claim that firm diversification is driven by agency considerations asserts that risk reduction is a second-best efficient response to agency and managerial risk aversion. Intuitively, even diversified shareholders would benefit from a reduction in the riskiness of the firm's profit stream if the risk reduction lowered the compensation level needed to attract and retain (risk-averse) managers, or if it reduced the distortions in managers' investment decisions. Diamond and Verrecchia [1982] present a model in which they fully characterize the optimal managerial compensation scheme and examine the effects of risk reduction on shareholder welfare. They interpret their results as showing that—through the effects on agency costs—risk reduction does benefit shareholders. Unfortunately, Diamond and Verrecchia considered a very specific formulation of the agency problem. The

---

shareholder is, by necessity, undiversified.

[5] Indeed, with limited liability, golden parachutes, and stock options, management compensation schemes may be convex functions of firm performance and, thus, induce risk loving.

parameter in their model that measures riskiness also measures the informativeness of the firm's returns as a signal of the agent's effort. We will argue that changes in informativeness, not risk, are what drive Diamond and Verrecchia's findings.

Marshall *et al.* [1984] argued that diversification can ameliorate agency problems both by improving the informativeness of the firm's returns as a signal of the agent's effort and by reducing the risk borne by the agent for any given contract. It is well-known that in situations in which the manager's actions cannot be observed directly by the owners, the agency problem can be viewed as a statistical inference problem and the cost of inducing the manager to take a certain action depends on the statistical relationship between that action and the possible outcomes that the owners might observe.[6] To the extent diversification alters this inference problem, it can affect the cost of agency. Marshall *et al.*, however, characterize informativeness in an informal and, to some extent, inaccurate manner. Moreover, their discussion of risk reduction is incomplete and potentially misleading in that it fails to account for changes in the optimal compensation scheme in response to changes in risk.

Aron [1988] also argues that diversification is valuable because of the potential benefits of improved information that reduces the cost of agency contracting. In her model, diversification into an additional line of business generates a new signal whose noise is independent of the other signals' and which has no effect on the distribution of the other signals. For this independent returns case, she shows that diversification reduces the cost incurred by the principal to induce the agent to undertake effort. The optimal extent of diversification is limited in her model by economies of scale in production—in order to diversify into additional product

[6] See Holmstrom [1979] and Shavell [1979].

lines, the firm would have to operate each at a smaller scale, which raises average production costs. Aron does not formalize the concept of increased informativeness. Indeed, in her comparison of diversification and the use of relative performance schemes, she focuses her answer on the correlation of returns across projects. Unfortunately, this is not directly related to informativeness, which as Holmstrom [1979] has shown is related to statistical sufficiency.[7]

The effects of corporate diversification on the agency relationship are an example of a more general problem. Changes in the structure of the corporation or the agent's tasks more generally can change the statistical relationship between the agent's ability and effort and the resulting returns. When executive compensation contracts depend on measures of performance, such as the level of sales or profits, the value of the measure is of interest not just as a signal of the agent's effort or ability, but also because it represents the financial returns to ownership. The statistical distribution of these variables may be affected by the organization of real activities (as opposed to accounting systems) within the firm. In evaluating alternative forms of organization, the owners of the firm may care about: (1) the expected level of returns; (2) the riskiness of the returns; and (3) the informativeness of the returns as a signal of managerial effort or ability.

An important question is how to disentangle these different effects. It is well known that additional information is valuable (for instance, in the standard agency problem the principal benefits from getting an informative signal that has no effect on the actual distribution of returns). Thus, in choosing how to structure its activities, a firm may be willing to trade off expected returns for improved infomation. A more difficult question is how a firm is willing to trade improvements in informativeness against increases in risk, or whether risk even matters.

[7] Consider, for example, two random variables s_1 and s_2. As long as there exists a function, $t(\cdot)$, such that $t(s_1) = s_2$,

As noted above, earlier authors were unable to disentangle the effects of risk and information.

In the present paper, we examine the effects of diversification in a model of optimal agency contracting that is general enough to distinguish between changes in riskiness and changes in informativeness. The next section presents a standard agency model in which we show that using the conventional Blackwell conception of informativeness, an improvement in information (holding the mean return constant) implies a reduction in risk. Thus, this approach is not well suited to examining these issues. Moreover, we join Gjesdal [1981] and Kim [1995] in arguing that Blackwell's definition of informativeness is unduly restrictive for agency purposes. We develop alternative measures of informativeness to demonstrate that, when these concepts do not coincide, information—not risk—is what matters in the agency setting. We show, by example, that shareholders may prefer a returns structure that entails a high degree of risk, but is highly informative, to one that is low-risk, but uninformative.

In Section 3, we examine the effects of corporate diversification on informativeness. While in Aron's model diversification always improves the information structure and the costs of diversification derive entirely from the effects of reassigning capital, we allow for the possibility that spreading effort itself affects the pattern of returns. In contrast to Aron, we argue that there are plausible cases in which diversification or risk-spreading activities actually reduce informativeness.

In Section 4, we examine what happens when the diversification choice is not verifiable in court and thus the managerial compensation scheme cannot be made dependent on it. While in many situations (*e.g.,* mergers and acquisitions) the reasonable assumption is that the

---

s_1 is a sufficient statistic for s_2 regardless of their correlation.

information structure is verifiable, there are diversification decisions (*e.g.*, the number of paths to pursue in new product development or marketing) that the manager makes with little direct oversight. While in principle the principal might be able to obtain such information and verify it in court, it strikes us as very unlikely that shareholders have anything like this level of detailed knowledge in practice. Thus, it is worth investigating what happens when the manager controls the diversification decision after the principal has set the compensation scheme.[8] We derive conditions under which the lack of verifiability does not affect the principal because the agent will choose the principal's preferred information structure. Thus, it is rational for principals to forgo trying to verify the information structure in these cases. We show by example, however, that there are other cases in which the lack of verifiability with respect to the agent's diversification decision is costly to the principal.

[8] We do not examine the (to us) less plausible case where the structure is not verifiable and is controlled by the principal.

2. THE VALUE OF INFORMATION AND RISK REDUCTION IN AN AGENCY MODEL

Consider a standard principal-agent model similar to Grossman and Hart's [1983]. A set of identical risk-neutral owners hire a risk-averse manager by offering her an incentive contract on a take-it-or-leave-it basis. The manager accepts if her expected utility under the contract is at least as great as her reservation utility, which we normalize to zero. The manager's utility is $V(y)$ - $K(a)$, where y is her monetary compensation and a is her action chosen from the finite set $\mathbf{A}$. $K(\bullet)$, the disutility-of-action function, is increasing; $V(\bullet)$, the utility of money function, is strictly increasing, strictly concave, and has an unbounded range. The owners' utility is $x - y$, where x is total revenue generated.

A *returns structure* is a set of densities over revenue levels, denoted by Π^{i}, where i indexes different returns structures. The vector $\mathbf{x} \equiv (x_1, x_2, \ldots , x_N)'$ represents the set of possible revenue levels.[9] Returns are indexed so that $m < n$ implies $x_m < x_n$. An element of Π^{i} is $\pi_n^i(a)$, the probability that revenues are x_n under the ith returns structure conditional on action a having been taken. The density over revenue levels conditional on action a is

$\pi^i(a) = \left(\pi_1^i(a), \ldots, \pi_N^i(a)\right)'$. For convenience, we define the index set $\mathbf{N} \equiv \{1, 2, \ldots , N\}$.

As is well known, the returns in this agency problem play two roles. One, they are the source of income for the principal and agent. In this capacity, we are interested in the *riskiness* of the returns, and second-degree stochastic dominance is the standard measure of riskiness. For the discrete setting under examination here, $\pi^2(a)$ is riskier than $\pi^1(a)$ if and only if

9 The prime (′) denotes vector or matrix transpose.

$$\sum_{i=1}^{n-1}\left(\sum_{j=1}^{i}\pi_j^1(a) - \pi_j^2(a)\right)(x_{i+1} - x_i) \ \leq \ 0 \quad \forall n \in \{2,\ldots,N\}. \tag{1}$$

and

$$\pi^2(a)'\mathbf{x} = \pi^1(a)'\mathbf{x}.$$

Returns also serve as a source of information—the level of returns can serve as a signal to the principal of the action taken by the agent. In this role, we are interested in the *informativeness* of the returns. The notion of Blackwell informativeness provides a starting point for our analysis. Consider two standard agency problems that are identical except for the stochastic relation between revenues and actions. If, for *all* possible a, $\pi^2(a) = \mathbf{Q}\pi^1(a)$, where $\mathbf{Q}$ is a constant stochastic transformation matrix (a matrix with non-negative elements in which each column sums to one and at least one column has two positive elements), then returns are a more (Blackwell) informative signal of the action under structure 1 than under structure 2. One can think of $\mathbf{Q}$ as "garbling" the signal that would have obtained in the first agency problem.

We want to examine the risk properties of diversification, and thus want to hold the mean level of returns constant to facilitate comparison. Hence, we will examine informativeness in settings in which one returns structure is more informative than another but the expected revenues depend only on the action taken and not on the returns structure. This leads us to restrict attention to *mean-preserving garblings*: stochastic transformation matrices, $\mathbf{Q}$, such that $\mathbf{Q}'\mathbf{x} = \mathbf{x}$. Hence, if $\pi^2(a) = \mathbf{Q}\pi^1(a)$ for all $a \in \mathbf{A}$, then $\pi^2(a)'\mathbf{x} = \pi^1(a)'\mathbf{x}$ for all $a \in \mathbf{A}$.[10]

---

[10] If Π^1 has less than full rank, then there may exist stochastic transformation matrices such that $\pi^2(a) = \mathbf{Q}\pi^1(a)$ for all $a \in \mathbf{A}$ and $\pi^2(a)'\mathbf{x} = \pi^1(a)'\mathbf{x}$ for all $a \in \mathbf{A}$, but $\mathbf{Q}'\mathbf{x} \neq \mathbf{x}$. That is, $\mathbf{Q}'\mathbf{x} = \mathbf{x}$ is a necessary condition only if Π^1 has full rank. Since little use can be made of this additional flexibility that arises if Π^1 has less than full rank, we have chosen not to divide the analysis according to the rank of Π^1.

Economically, one can interpret a mean-preserving garbling as follows. Under returns structure Π^1, the payoff if the nth state occurs is x_n. Under returns structure Π^2, a lottery over revenues is held if the nth state occurs, where the lottery's probabilities are given by the nth column of $\mathbf{Q}$, $\mathbf{q}_{\bullet n}$. Each of these lotteries is mean-preserving (*i.e.*, the nth lottery has mean x_n for all n) since $\mathbf{q}_{\bullet n}'\mathbf{x} = x_n$ by construction.

Two Results that Make Distinctions Difficult

While distinct concepts, informativeness and riskiness are related, and this relationship can make it difficult to determine which property drives agency costs.

It is no surprise that, when $\pi^2(a) = \mathbf{Q}\pi^1(a)$ for all $a \in \mathbf{A}$ and $\mathbf{Q}$ is a mean-preserving garbling, the densities in Π^1 are less risky than the densities in Π^2 in the sense of second-degree stochastic dominance—by construction $\pi^2(a)$ is a mean-preserving spread of $\pi^1(a)$.

Proposition 1: *Consider two returns structures,* Π^1 *and* Π^2, *for which the stochastic transformation matrix between the first and second returns structures is a mean-preserving garbling. Then:*

i) Π^1 *is more informative than* $\Pi^{2.}$

ii) *For all* $a \in \mathbf{A}$, $\pi^1(a)$ *is less risky than* $\pi^2(a)$ *in the sense of second-degree stochastic dominance.*

Proposition 1(ii) is a straightforward corollary of a theorem of Blackwell (Marshall and Olkin [1979, p.417]):

Theorem (Blackwell): *Consider two densities,* π^1 *and* π^2, *with support* $\mathbf{x}$. *The first density is* *less risky than the second, in the sense of second-degree stochastic dominance, if and only if there exists a mean-preserving garbling,* **Q***, such that* $\pi^2 = \mathbf{Q}\pi^1$.

One might suspect that the converse of Proposition 1 is also a corollary of this Theorem. It is not. The theorem holds for a pair of *densities* and not a pair of returns structures. Although the Theorem does not imply the converse of Proposition 1, it does imply the following:[11]

Proposition 2: *Let* Π^1 *and* Π^2 *be two returns structures, where: (i)* Π^2 *is of full rank; (ii) the expected returns conditional on the action taken are the same under the two returns structures; and (iii) for some action* $\tilde{a} \in \mathbf{A}$, $\pi^1(\tilde{a})$ *is strictly less risky than* $\pi^2(\tilde{a})$ *in the sense of second-degree stochastic dominance. Then* Π^2 *is* not *more informative than* Π^1 *in the Blackwell sense.*

Reasons to Believe that Informativeness is What Matters

Although linked, informativeness and riskiness remain different concepts. Indeed, it is possible to rank individual pairs of densities by second-degree stochastic dominance without being able to rank the returns structures by informativeness and vice versa. In this section, we will argue that informativeness is what matters from an agency perspective when the principal is risk neutral.

The basic intuition is as follows. The principal is risk neutral and thus has incentives to insure the risk-averse agent. But in the presence of a moral hazard problem, the principal must trade off the provision of insurance against the provision of incentives. The less the principal insures the agent, the greater the principal's expected cost of implementing a given action because he has to pay the agent more money to compensate for the risk. The terms of this trade-off depend on the extent to which the principal is informed about the agent's actions. If the principal had perfect information about the agent's action, for example, then the principal would fully insure the agent and the riskiness of the firm's returns would be irrelevant.

Grossman and Hart [1983] show how the trade-off is affected by an improvement in information in the Blackwell sense. Their Proposition 13 establishes that, if there exists a stochastic transformation matrix **Q** such that $\pi^2(a) = \mathbf{Q}\pi^1(a)$ for all actions a, then the principal's expected cost of implementing a given action in the first agency problem is no greater than the expected cost of implementing that action in the second agency problem.

Like Gjesdal [1981] and Kim [1995], we believe that Blackwell's definition of informativeness is unduly restrictive for agency purposes. It is well known, for instance, that, if the support of the returns varies with the agent's action in the right way, then the principal can completely solve the incentive problem while providing full insurance to an agent who takes the action desired by the principal.[12] The following example exploits this fact to show that the principal may well prefer a returns structure that gives rise to a high level of risk (and information) to one that does not.

Example 1: Suppose $\mathbf{x} = (1,2,3)'$; $\mathbf{A} = \{0,1\}$; and

$$\pi^1(0) = \begin{pmatrix} \frac{1}{6} \\ \frac{2}{3} \\ \frac{1}{6} \end{pmatrix}, \ \pi^2(0) = \begin{pmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{pmatrix}, \ \pi^1(1) = \begin{pmatrix} \frac{1}{6} \\ \frac{1}{6} \\ \frac{2}{3} \end{pmatrix}, \text{ and } \ \pi^2(1) = \begin{pmatrix} \frac{1}{4} \\ 0 \\ \frac{3}{4} \end{pmatrix}.$$

It is straightforward to verify that, for both a, $\pi^1(a)$ is less risky than $\pi^2(a)$.[13] Under the first

---

[11] Proofs not in the text may be found in the Appendix.

[12] In a related context, Hermalin and Katz [1991] show that renegotiation can mitigate the effects of risk in an agency problem, so that even a very limited amount of information is sufficient to attain the first-best outcome.

[13] Note that there is no *common* stochastic transformation matrix such that $\pi^2(a) = \mathbf{Q}\pi^1(a)$: Since $\pi^2(1) = 0$, the middle row of **Q** would necessarily consist entirely of zeros, which would be inconsistent with $\pi^2(0) = 1/3$. Thus, the

returns structure, the cost of implementing $a = 1$ exceeds the full-information level because the optimal contract that induces the agent to take this action entails her bearing some of the risk. Under the second returns structure, however, obtaining a revenue of 2 is proof that the manager took action $a = 0$, and the owners can thus implement $a = 1$ at the full-information cost. Therefore, the owners prefer the second, riskier, returns structure.

Example 1 establishes that, when riskiness and informativeness do not coincide, the owners can rationally prefer the riskier returns structure. This example also illustrates the fact that Blackwell informativeness is an imperfect measure of informativeness in an agency problem. Our next result shows that this point is not dependent on the existence of a shifting support.

This result builds on the following intuitive notion of informativeness. Suppose that the principal wants to induce the agent to choose action $\hat{a}$. Intuitively, the principal's ability to discern whether the agent has indeed chosen this action will depend on how dissimilar are the densities over profits associated with other actions. Thus, we will interpret returns structure 1 as being more informative than returns structure 2 about whether action a was taken if for all $a \in \mathbf{A} \setminus \{\hat{a}\}$, $\pi^1(a)$ is "farther" from $\pi^1(\hat{a})$ than $\pi^2(a)$ is from $\pi^2(\hat{a})$.

Proposition 3: *If two returns structures,* Π^1 *and* Π^2*, are such that:*

(i) *for each* $n \in \mathbf{N}$ *either*

$$\pi_n^1(a) - \pi_n^1(\hat{a}) > \pi_n^2(a) - \pi_n^2(\hat{a}) \geq 0 \quad \forall a \in \mathbf{A} \setminus \{\hat{a}\}$$

or

$$\pi_n^1(a) - \pi_n^1(\hat{a}) < \pi_n^2(a) - \pi_n^2(\hat{a}) \leq 0 \quad \forall a \in \mathbf{A} \setminus \{\hat{a}\};$$

---

returns structures cannot be ranked by the criterion of Blackwell informativeness.

and

(ii) for all a $\in \mathbf{A}\backslash\{\hat{a}\}$

$$\min_{n\in\mathbf{N}} \frac{\pi_n^2(\hat{a})}{\pi_n^1(\hat{a})} > \max_{m\in\mathbf{N}} \frac{\pi_m^2(a)-\pi_m^2(\hat{a})}{\pi_m^1(a)-\pi_m^1(\hat{a})},$$

then the cost of implementing â *under returns structure* Π^1 *is less than or equal to the cost under* Π^2. *If* â *is not a least-cost action (i.e., if* K(â) > K(a) *for some* a $\in$ **A**), *then the inequality is strict.*

It is well known within agency theory that what is important for solving an agency problem are the likelihood ratios,

$$h_n^i(a,\hat{a}) \equiv \frac{\pi_n^i(a)-\pi_n^i(\hat{a})}{\pi_n^i(\hat{a})},$$

since compensation in state *n* is an increasing function of

$$\mu^i - \sum_{a\in\mathbf{A}\backslash\{\hat{a}\}} \lambda^i(a) h_n^i(a,\hat{a}), \tag{2}$$

where μ^i and $\lambda^i(a)$ are positive constants. Proposition 3's two conditions are conditions on the way the $h^i_n(a,\hat{a})$ are distributed under the two information systems. As such, Proposition 3 is related to Kim [1995]'s work, which shows that if the distribution of the likelihood ratios under information structure 1 is a mean-preserving spread of the distribution under information structure 2, then 1 is more informative (leads to a better solution of the agency problem) than 2. Our conditions, although closely related to Kim's, are different. In particular, it is possible to show by example that the information structures can satisfy our conditions without their likelihood ratios being ordered by second-degree stochastic dominance as required by Kim.[14] Hence, stochastic dominance is not a necessary condition. The reason is that although the

[14] An example can be constructed when there are two states and two actions, *a* and $\hat{a}$, such that $\pi_1^1(a) = .7$,

principal's costs are convex in (2), the weights μ^i and $\lambda^i(a)$ also depend on the information.

The important point to note is that $\pi^1(\hat{a})$ and $\pi^2(\hat{a})$ may satisfy these conditions even though $\pi^1(\hat{a})$ is riskier than $\pi^2(\hat{a})$. This can be seen by comparing these two conditions with the definition of second-degree stochastic dominance, inequality (1) above. The following example further illustrates this point.

Example 2: Suppose that $\mathbf{x} = (0, \ 100, \ \ 200)'$,

$$\pi^1(a_1) = \pi^2(a_1) = (.395,.48,.125)',$$
$$\pi^1(a_2) = \pi^2(a_2) = (.32,.48,.2)',$$
$$\pi^1(a_3) = (.48,.04,.48)',$$

and

$$\pi^2(a_3) = (.46,.08,.46)'.$$

Moreover, suppose that $3K(a_1) = 2K(a_2) = K(a_3) = \theta > 0$.

The expected returns conditional on actions a_1, a_2, and a_3 are 73, 88, and 100, respectively. Thus, for θ sufficiently low, the principal will choose a contract that induces the agent to take action a_3. By Proposition 3, it costs less to implement a_3 under the first information structure than under the second information structure. But $\pi^1(a_3)$ is a mean-preserving spread of $\pi^2(a_3)$—the principal chooses the riskier returns structure.

We gain additional insight into Proposition 3 by holding the probability density associated with $\hat{a}$ invariant across the two returns structures being compared while varying the densities associated with other actions. By doing this, riskiness is held constant across the two

---

$\pi_1^{\,1}(\hat{a}) = .3$, $\pi_1^{\,2}(a) = .7$, and $\pi_1^{\,2}(\hat{a}) = .35$. Details available upon request.

returns structures but informativeness may vary. The corollary establishes that given the choice between two returns structures with equal riskiness, the principal chooses the one that is more informative as measured by the distance between densities.

Corollary: *If two returns structures,* Π^1 *and* Π^2*, are such that*

i) $\pi^1(\hat{a}) = \pi^2(\hat{a}) \equiv \hat{\pi}$,

and

ii) *for each* $n \in \mathbf{N}$ *either* $\pi_n^1(a) > \pi_n^2(a) \geq \hat{\pi}_n$ $\forall a \in \mathbf{A}\backslash\{\hat{a}\}$ *or* $\pi_n^1(a) < \pi_n^2(a) \leq \hat{\pi}_n$ $\forall a \in \mathbf{A}\backslash\{\hat{a}\}$,

then the cost of implementing $\hat{a}$ *is less under returns structure* Π^1 *than under* Π^2. [15]

The corollary underscores again that when risk and informativeness are not related, it is informativeness that matters in an agency context.

3. RISK AND INFORMATION EFFECTS OF MANAGERIAL DIVERSIFICATION

Having looked at returns structures in the abstract, we now ask how they relate to diversification. We do so in the context of a simple illustrative example. The manager must choose a total effort level $a \in \{0,1\}$. The owners choose whether the firm diversifies or not. There are two potential projects, 1 and 2. Let λ denote the fraction of total effort devoted to project 1. For now, we assume that the owners choose λ.

Each project can either succeed or fail. The level of effort devoted to a project can affect both the project's probability of success and the payoff realized in the event of success. Because our focus is on the informational and risk impacts of diversification in a hidden-action problem, we want the expected revenue (conditional on a given level of total effort) to be invariant with respect to the extent of diversification.[16] That is, although effort gets divided across projects, this is done in such a way that the firm has constant returns to both scale and scope, thereby isolating the informational effects.

We assume that a project fails with certainty if no effort is allocated to it and that an unsuccessful project's returns are 0. Hence, the firm's expected returns conditional on $a = 0$ are equal to 0 no matter what the value of λ.

Let $p_{ss}(\lambda)$ denote the probability that both projects succeed when $a = 1$ and λ of the total effort is devoted to project 1. Let $p_{ff}(\lambda)$ denote the probability that both projects fail given λ when $a = 1$, and define $p_{fs}(\lambda)$ and $p_{sf}(\lambda)$ in a similar manner. We put no restriction on the possibility that the projects are correlated (positively or negatively). We assume that, when $a = 1$, a successful project generates returns $\alpha + \beta\lambda$, $\alpha,\beta \in [0,\infty)$.

There are two polar cases to consider. When $\alpha = 0$, a diversified firm's revenue from a successful project is proportional to the percentage of total effort allocated to it. This polar case includes the subcase in which a project's probability of success is unaffected by diversification, but the value of a successful project is lowered. This might loosely be thought of as the effect of entering into a joint venture and sharing the benefits of a successful project with another firm.

When $\beta = 0$, a diversified firm's revenue from each successful project is α. This model captures the possibility that project value is unaffected by diversification, but that the probability

---

[15] The corollary also follows from Kim [1995]'s main proposition suitably modified for a discrete state space.

[16] We are not allowing the firm to expand by simply adding more projects of equal size. As Samuelson [1963] (cited in Diamond [1984]) showed, adding statistically independent projects in this manner might not reduce the manager's aversion to the risk of any one project. Moreover, managers might be unable to control a larger number of projects if they had to undertake them all without either hiring additional managers or taking on joint venture partners.

of a successful project is lowered. Such effects might arise when a fixed managerial team spreads its efforts across an increased number of projects and thus dilutes its efforts.

The following result greatly simplifies the analysis.

Lemma 1: *For the principal-agent model of this section, if the probability of complete failure (i.e.,* $x = 0$*) conditional on* $a = 1$ *is lower under returns structure 1 than under returns structure 2, then the principal's cost of implementing total effort level* $a = 1$ *is lower under returns structure 1.*

Proposition 4: *Suppose that the owners wish to implement* $a = 1$*. When* $\alpha = 0$*, diversification reduces risk, and increases informativeness and expected profits (relative to non-diversification), as long as the returns of the two projects are not perfectly correlated. When* $\beta = 0$*, diversification increases risk, and lowers informativeness and expected profits, as long as there is a positive probability that the two projects will succeed simultaneously.*

Proof: By Lemma 1, our interest is in comparing $p_{ff}(1)$ with $p_{ff}(\lambda)$ for $\lambda \in (0,1)$. By hypothesis, diversification does not affect the mean return. Hence, there is exists a constant τ such that

$$\{2\alpha+\beta\}p_{ss}(\lambda) + \{\alpha+\beta\lambda\}p_{sf}(\lambda) + \{\alpha+\beta(1-\lambda)\}p_{fs}(\lambda) = \tau \quad \forall\lambda \in [0,1].$$

From this expression and the fact that $p_{ss}(\lambda) + p_{sf}(\lambda) + p_{fs}(\lambda) + p_{ff}(\lambda) = 1$, it follows that

$$p_{ff}(\lambda) = p_{ff}(1) + \{\alpha p_{ss}(\lambda) - \beta(1-\lambda)p_{sf}(\lambda) - \beta\lambda p_{fs}(\lambda)\}/(\alpha+\beta), \tag{3}$$

where we have made use of the fact that $p_{ss}(1) = p_{fs}(1) = 0$.

When $\alpha = 0$, equation (3) implies $p_{ff}(\lambda) < p_{ff}(1)$ for all $\lambda \in (0,1)$ as long as the returns of the two projects are not perfectly correlated. Therefore, by Lemma 1, agency costs are lower, and thus the owners' expected profits are higher under diversification. Using (1), it is straightforward to show that, for any λ, the returns density under diversification is less risky than the density without diversification.

When $\beta = 0$, equation (3) implies $p_{ff}(\lambda) > p_{ff}(1)$ for all $\lambda \in (0,1)$ such that $p_{ss}(\lambda) > 0$. Therefore, by the logic used for the case of $\alpha = 0$, one can establish that agency costs are higher, and thus the owners' expected profits are lower under diversification. Moreover, straightforward calculations show that, for any λ, the returns density with diversification is riskier than the density under diversification. **Q.E.D.**

Admittedly, Proposition 4 does not pertain to a particularly robust model. Nonetheless, it is sufficient to point out that diversification can have ambiguous effects.[17] Specifically, diversification can make signals about the manager's actions more or less informative depending on how diversification changes the stochastic relation between signals (revenues) and actions. When $\alpha = 0$, diversification essentially consists of doubling the sample size—roughly, the outcomes of two Bernoulli trials are observed rather the outcome of only one trial. Since a larger sample allows for better inference, it is not surprising that diversification of this form reduces agency costs. When $\beta = 0$, diversification makes the outcome of each trial less informative—in fact, to the point that the loss of information outweighs the informational benefit of a second trial—which raises agency costs.

4. WHO DECIDES WHETHER TO DIVERSIFY OR NOT TO DIVERSIFY?

[17] The case of $\alpha = 0$ corresponds to Aron's [1988] analysis.

So far, we have assumed that the owners decide whether the firm should diversify or not. In many situations, this is a reasonable assumption: Mergers and acquisitions often require shareholder approval, for example. On the other hand, there are diversification decisions (*e.g.*, the number of paths to pursue in new product development) that the manager makes with little direct oversight. For there to be a meaningfully distinction between situations in which the principal chooses the information structure and those in which the agent chooses, it must be in the latter case that the principal cannot verify the agent's choice of information structure (preventing the principal from making the agent's compensation contingent on the choice of information structure). While this is implausible when the choice entails new lines of business, it is much more likely when the choice is with regard to the number of R&D or marketing projects to pursue. While in principle the principal might be able to obtain such information, it strikes us as very unlikely that shareholders have anything like this level of detailed knowledge in practice. Thus, it is worth investigating what happens when the manager controls the diversification.

We can embed the choice of returns structure in a standard agency where the agent chooses an action pair consisting of a productive action, a, and an "information-structure" action, i. We call the principal-agent problem in which the owners can specify the returns structure in the contract (*i.e.*, the diversification choice is verifiable) the *direct-choice problem* because the owners directly choose the returns structure. We call the principal-agent problem in which the agency contract depends solely on the firm's realized returns the *indirect-choice problem* because the owners indirectly choose the returns structure through the incentive scheme they offer to their manager.[18]

[18] The indirect-choice model is an example of a multi-task principal-agent model. For more on multi-task principal-

Obviously, the owners' expected profit in the indirect-choice problem is no greater than their expected profit in the direct-choice problem. The first question we ask is: when is their expected profit no less? It is easy to see that the direct- and indirect-choice problems yield the principal the same payoff in the example of Section 3. The argument is as follows. As shown in the previous section, the optimal contract in the direct-choice problem entails one payment if the firm earns a positive return and another, lower payment if returns equal 0. Hence, if the principal used this contract in the indirect-choice problem, the agent would be induced to choose the returns structure (*i.e.*, the extent of diversification) to minimize the probability of simultaneous failure by all projects. But this is just the choice that the principal makes in the direct-choice problem. Hence, the principal implements the same outcome with the same contract in both the direct- and indirect-choice problems.

Consider now a more general setting. Because of the congruence of Blackwell informativeness and riskiness, we know that the manager will choose the more informative returns structure if there is a mean-preserving garbling between the two returns structures and the manager is risk averse with respect to the firm's profits (*i.e.*, the composition of the agent's compensation scheme and her utility function is concave). Define $v_i \equiv V(y_i)$, where y_i is the agent's monetary compensation contingent on the firm's returns being x_i.

Corollary to Proposition 1: *Consider two returns structures,* Π^1 *and* Π^2*, such that the second returns structure is a mean-preserving garbling of the first, and suppose that the composition of the agent's compensation scheme and utility function is concave:*

---

agent problems see Holmstrom and Milgrom [1991].

$$\frac{v_{n+1} - v_n}{x_{n+1} - x_n} < \frac{v_n - v_{n-1}}{x_n - x_{n-1}} \;\forall n \in \mathbf{N} \setminus \{1, N\}. \tag{4}$$

Then the manager prefers the first returns structure to the second.

Since, for each action, the density in Π^2 is a riskier distribution over revenues than the corresponding density in Π^1, it follows that, for any action she might choose, the manager expects to do better under the first returns structure than under the second.

Condition (4) is unsatisfactory, in that it is an assumption about endogenous variables. The next proposition builds on an earlier result of Grossman and Hart to state its assumptions in terms of exogenous conditions.

Proposition 5: *Consider two returns structures,* Π^1 *and* Π^2*, where the second is a mean-preserving garbling of the first. Suppose, too, that the following conditions are satisfied:*

A1 *(No Shifting Support):* $\pi_n^i(\mathrm{a}) > 0$ *for all* a ∈ **A**, *for all* n ∈ **N** *and* i ∈ *{1,2}.*

A2 *(Monotone Likelihood Ratio Property): For all* a *and* â *in* **A**, K(a) < K(â) *implies that* $\pi_n^i(\mathrm{a}) / \pi_n^i(\hat{\mathrm{a}})$ *is non-increasing and convex in* n *and* i ∈ *{1,2}.*

A3 *(Concavity of Distribution Function Property):* K(a) = γK(â) + (1-γ)K(ã) (γ ∈ [0,1]) *implies* $\mathbf{P}^i(\mathrm{a}) \le \gamma \mathbf{P}^i(\hat{\mathrm{a}}) + (1-\gamma)\mathbf{P}^i(\tilde{\mathrm{a}})$, *where* $\mathbf{P}^i(\mathrm{a})$ *is the cumulative distribution function vector over revenues induced by action* a *(*i.e., $\mathbf{P}^i(\mathrm{a}) =$ $\left(\pi_1^i(\mathrm{a}), \pi_1^i(\mathrm{a}) + \pi_2^i(\mathrm{a}), \ldots, \pi_1^i(\mathrm{a}) + \ldots + \pi_N^i(\mathrm{a})\right)'$ *),* i ∈ *{1,2}.*

A4 *(Income Effects on Attitudes toward Risk):* $1/V'(y)$ *is concave in* y.

A5 *(Convex Revenues):* $x_n - x_{n-1}$ *is non-decreasing in* n.

Then the optimal contract for the direct-choice problem is also an optimal contract for the indirect-choice problem. Moreover, this contract induces the manager to choose the more informative returns structure and yields the owners an expected profit equal to what they would receive in the direct-choice problem.

Assumptions 2-5 strike us as restrictive, and there exist plausible settings in which the optimal contract entails the manager's utility being a non-concave function of returns, so that the manager prefers the less informative returns structure. In such cases, the owners' expected profit is strictly less in the indirect-choice problem, and it is useful to explore why. In particular, profits could be lower because fewer actions are implementable in the indirect-choice problem or because those actions that are implementable cost more to implement.

We first compare the set of implementable actions in the two problems. To this end, we define the following property.

Definition (Convexity of Disutility Property): *Action* a *satisfies* convexity of disutility *if* $\pi(a)'\mathbf{x} = \sum_{j\in \mathbf{J}}\lambda_j\pi(a_j)'\mathbf{x}$ *implies that* $K(a) \leq \sum_{j\in \mathbf{J}}\lambda_j K(a_j)$, *where* $\{\lambda_j\}$ *is a set of non-negative weights summing to one and* **J** *is an index set for actions other than* a.[19]

This property implies that any mixed strategy over actions that yields the same expected revenue as action *a* must yield the manager a greater expected disutility of effort than *a*. The convexity of disutility property is a more stringent property than implementability in the direct-choice

[19] This property is related to the Concavity of Distribution Function Property above. It differs in that it applies to any convex combination of actions (any mixed strategy) and is based on expected values rather than first-degree stochastic

problem.[20]

Proposition 6: *Consider two returns structures,* Π^1 *and* Π^2, *where the second is a mean-preserving garbling of the first. Suppose that action* â *satisfies the convexity of disutility property in the direct-choice problem (*i.e., *under the first returns structure). Then there exists a contract in the indirect-choice problem that implements* â *and induces the manager to choose the more informative returns structure.*

Of course being able to implement an action pair does not imply that the owners would want to implement it. Indeed, it is possible in the indirect-choice problem that the owners would prefer to implement the *less* informative returns structure. We illustrate this with an example.

Example 3: Suppose $V(y) = \ln(y)$, $\mathbf{A} = \{0,1\}$, $K(0) = 0$, $K(1) = 1$, $\pi^1(0) = (.5,.25,.25)'$, $\pi^1(1) = (.4,.05,.55)'$, and $\mathbf{x} = (0,12,24)'$. Intuitively, this example represents the following situation. By working hard ($a = 1$), the manager does not greatly reduce the chance that the project will be a complete failure ($x = 0$). But the manager's effort does affect whether a successful project will do all right or extremely well. Now, consider returns structure $\Pi^2 = \mathbf{Q}\Pi^1$, where $\mathbf{Q}$ is the mean-preserving garbling

$$\mathbf{Q} = \begin{pmatrix} 1 & .05 & 0 \\ 0 & .9 & 0 \\ 0 & .05 & 1 \end{pmatrix}.$$

Under either information structure, the optimal contract for implementing $a = 0$ is to pay the

---

dominance.

[20] Implementability of a in the direct-choice problem requires only that there exists *some* vector $\mathbf{v}$ such that, if $\pi(a)'\mathbf{v} = \Sigma_{j\in J}\lambda_j\pi(a_j)'\mathbf{v}$, then $K(a) \leq \Sigma_{j\in J}\lambda_j K(a_j)$, where $\{\lambda_j\}$ is a set of non-negative weights summing to one.

agent y = 1 for all outcomes. The owners' expected profit is 8. The optimal contract for implementing (a=1,i=1) is $\tilde{y}_1 \approx .15$, $\tilde{y}_2 \approx 1.87$, and $\tilde{y}_3 \approx 22.76$, with an expected cost of 12.67. The optimal contract for implementing (a=1,i=2) is $y_1 \approx 3.09$, $y_2 \approx .02$, and $y_3 \approx 3.75$. This contract's expected cost is 3.32, and the owners' expected profit is 10.48. Therefore, the owners will induce the agent to choose the less informative information structure, i=2.

In this example, because the two returns structures are so close to each other, the optimal contract for implementing a = 1 remains one that punishes the manager for achieving x = 12 (since x = 12 is a relatively rare event if a = 1, but a relatively more common event if a = 0). Consequently, the manager's utility is a convex function of revenue, so she prefers the less informative—and, therefore, riskier—returns structure.

The question remains: would the owners implement the less informative returns structure if the two returns structures were "far apart"? The following proposition shows the answer is no when we interpret the notion of "far apart" as the repeated application of a mean-preserving garbling:

Proposition 7: *Consider returns structure* Π^1 *and mean-preserving garbling* $\mathbf{Q}$. *Suppose that* $(\mathbf{Q})^n$ *converges to* $\mathbf{Q}^\infty$ *as* n *goes to* ∞. *In the indirect-choice problem with returns structures* Π^1 *and* $\Pi^2 \equiv \mathbf{Q}^\infty\Pi^1$, *the owners will choose a contract that induces the agent to choose the more informative returns structure,* Π^1.

5. CONCLUSION

In this paper, we explored agency considerations as a possible explanation for why corporations diversify even when their shareholders could otherwise do so on their own. In evaluating alternative forms of organizing the firm and the agent's activities, the owners of the firm potentially care about the returns' expected level, riskiness, and informativeness.

We first examined the relationship between riskiness and informativeness. We showed that diversification that improves the informativeness of the returns structure in the Blackwell sense also reduces the firm's riskiness (Proposition 1). Thus, even if risk-neutral owners did not desire a reduction in risk, one might observe a positive correlation between diversification and risk reduction. To explore this issue further, we proposed new measures of informativeness appropriate for agency relationships. Using these measures, we showed (Example 1 and Proposition 3) that owners may choose returns structures (*e.g.*, diversification strategies) that are more informative but entail increased risk.

We next turned to examination of the effects of diversification on informativeness and riskiness. In the context of an extended example, we showed that diversification that entails the spreading of managerial effort—rather than purely financial diversification—has ambiguous effects on risk and information.

Lastly, we considered situations in which the manager controls the diversification decision. We derived conditions under which the manager would choose to adopt the owners' preferred diversification strategy. Consequently, leaving the diversification decision to the manager under these conditions would be without cost to the owners. More generally, delegating the diversification decision to the manager can be costly for the owners. This is not so much because delegation makes it impossible to implement certain courses of action—the condition under which it has no impact on the set of implementable actions is relatively innocuous—but rather because of the loss of information. It is ambiguous whether the manager will, in equilibrium, choose the same diversification strategy as would the owners had they controlled the diversification decision. In Example 2, we showed that choices could be different; but, in Propositions 5 and 7, we derived conditions under which the choices would be the same.

In closing, we want to discuss the claim that diversification generates financial benefits by economizing on bankruptcy costs.[21] Without getting into a full analysis of bankruptcy here, we want to raise several issues. First, it is not obvious that diversification economizes on bankruptcy costs. In general, combining projects can raise or lower the expected costs of bankruptcy since a single failing project may be "saved" by other (successful) projects, or it may drag all of them down.[22] Second, it is critical to understand why firms issue securities that give rise to the possibility of bankruptcy. If bankruptcy serves a useful role, then diversification to avoid bankruptcy may be costly. For example, if bankruptcy is a device to discipline managers (*e.g.*, the manager gets in trouble only when there is a bankruptcy), then a joint venture or some other form of diversification may allow a manager to undertake projects that investors would otherwise reject.[23] Finally, one does not want to overstate the costs of bankruptcy. A firm with value as an ongoing concern may renegotiate its debts without formally going bankrupt or may reorganize (and continue to operate) under bankruptcy. In either case, the costs are largely

[21] For an early example of a model in which this occurs, see Lewellen [1971].
[22] For more on this point, see Higgins' [1971] discussion of Lewellen [1971].
[23] Recently, a number of papers have sought to model these issues formally. See, e.g., Aghion and Bolton [1992], Harris and Raviv [1990], and Hart [1991].

administrative ones, not the loss in the value of firm's productive activities.

REFERENCES

Aghion, P. and P. Bolton. 1992. "An Incomplete Contract Approach to Bankruptcy and the Financial Structure of the Firm." *Review of Economics Studies* **59**:473-494.

Alberts, W.W. 1966. "The Profitability of Growth by Merger," in *The Corporate Merger*, W.W. Alberts and J.E. Segall, eds. Chicago: The University of Chicago Press.

Amihud, Yakov, Peter Dodd, and Mark Weinstein. 1986. "Conglomerate Mergers, Managerial Motives, and Stockholder Wealth." *Journal of Banking and Finance* **10**:401-410.

Amihud, Yakov, and Baruch Lev. 1981. "Risk Reduction as a Managerial Motive for Conglomerate Mergers." *Bell Journal of Economics* **12**:605-617.

Aron, Debra J. 1988. "Ability, Moral hazard, Firm Size, and Diversification." *RAND Journal of Economics* **19**:72-87.

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

Diamond, Douglas W., and Robert E. Verrecchia. 1982. "Optimal Managerial Contracts and Equilibrium Security Prices." *The Journal of Finance* **37**:275-288.

Errunza, Vihang R., and Lemma W. Senbet. 1984. "International Corporate Diversification, Market Valuation, and Size-Adjusted Evidence." *The Journal of Finance* **34**:727-743.

Gjesdal, Froystein. 1981. "Accounting for Stewardship." *Journal of Accounting* **19**:208-231.

Grossman, Sanford, and Oliver D. Hart. 1983. "An Analysis of the Principal-Agent Problem." *Econometrica* **51**:7-45.

Harris, Milton, and Artur Raviv. 1990. "Capital Structure and the Informational Role of Debt." *The Journal of Finance* **45**:321-349.

Hart, Oliver. 1991. "Theories of Optimal Capital Structure: A Principal-Agent Perspective." Unpublished draft, MIT.

Hermalin, Benjamin E. 1993. "Managerial Preferences Concerning Risky Projects," *Journal of Law, Economics, & Organization* **9**:127-135.

Hermalin, Benjamin E., and Michael L. Katz. 1991. "Moral Hazard and Verifiability: The Effects of Renegotiation in Agency." *Econometrica* **59**:1735-1753.

Higgins, Robert C. 1971. "Discussion of 'A Pure Financial Rationale for the Conglomerate Merger.'" *The Journal of Finance* **26**:543-545.

Holmstrom, Bengt. 1979. "Moral Hazard and Observability." *Bell Journal of Economics* **10**:74-91.

Holmstrom, Bengt, and Paul Milgrom. 1991. "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design." *Journal of Law, Economics, and Organization* **7**:24-52.

Kim, Son Ku. 1995. "Efficiency of an Information System in an Agency Model," *Econometrica*, **63:**89-102.

Levy, Haim, and Marshall Sarnat. 1970. "Diversification, Portfolio Analysis and the Uneasy Case for Conglomerate Mergers." *The Journal of Finance* **25**:795-802.

Lewellen, Wilbur G. 1971. "A Pure Financial Rationale for the Conglomerate Merger." *The Journal of Finance* **26**:521-537.

Lewellen, Wilbur G., Claudio Loderer, and Ahron Rosenfeld. 1989. "Mergers, Executive Risk Reduction, and Stockholder Wealth." *Journal of Financial and Quantitative Analysis* **24**:459-472.

Marshall, Albert W., and Ingram Olkin. 1979. *Inequalities: Theory of Majorization and Its Applications*. San Diego: Academic Press.

Marshall, William J., Jess B. Yawitz, and Edward Greenberg. 1984. "Incentives for Diversification and the Structure of the Conglomerate Firm." *Southern Economic Journal* **51**:1-23.

Morck, Randall, Andrei Shleifer, and Robert W. Vishny. 1990. "Do Managerial Objectives Drive Bad Acquisitions?" *The Journal of Finance* **45**:31-48.

Porter, Michael. 1987. "From Competitive Advantage to Corporate Strategy." *Harvard Business Review*, May-June.

Samuelson, Paul A. 1963. "Risk and Uncertainty: A Fallacy of Large Numbers." *Scientia*.

Shavell, Steven. 1979. "Risk Sharing and Incentives in the Principal and Agent Relationship." *The Bell Journal of Economics* **10**:55-73.

APPENDIX

Proof of Proposition 2: Suppose, contrary to the statement of the proposition, that Π^2 is more Blackwell informative. Then there exists a stochastic transformation matrix $\mathbf{R}$ such that $\Pi^1 = \mathbf{R}\Pi^2$. Since the two returns structures yield the same expected returns conditional on the same action, $\Pi^{2\prime}(\mathbf{R} - \mathbf{I})'\mathbf{x} = \mathbf{0}$. Because Π^2 has full rank, it follows that $\mathbf{R}'\mathbf{x} = \mathbf{x}$. Thus, $\mathbf{R}$ is a mean-preserving garbling. From Blackwell's Theorem, it follows that $\pi^2(a)$ is less risky than $\pi^1(a)$—but this contradicts the assumption that $\pi^1(a)$ is strictly less risky than $\pi^2(a)$. **Q.E.D.**

In the proofs below, we will make use of the following two facts about the individual rationality (IR) and incentive compatibility (IC) constraints stated below. One, if contract $\mathbf{v}$ satisfies these constraints, then it implements action $\hat{a}$. Two, if the optimal solution to the agency problem entails implementing action $\hat{a}$, then it must satisfy these constraints.

$$\sum_{i\in\mathbf{N}} \pi_i(\hat{a})v_i = K(\hat{a}) \qquad \text{(IR)}$$

and

$$\sum_{i\in\mathbf{N}} \{\pi_i(a)-\pi_i(\hat{a})\}v_i \leq K(a) - K(\hat{a}) \quad \forall a \in \mathbf{A}.^{24} \qquad \text{(IC)}$$

Lemma A1: If $\mathbf{v}^*$ is the solution to an agency problem in which $\pi_m(a)-\pi_m(\hat{a}) > 0 > \pi_n(a)-\pi_n(\hat{a})$ $\forall a \in \mathbf{A}\backslash\{\hat{a}\}$, then $v_m^* \leq v_n^*$.

Proof of Lemma A1: Suppose that, contrary to the hypothesis of the lemma, $\pi_m(a)-\pi_m(\hat{a}) > 0 > \pi_n(a)-\pi_n(\hat{a})$ $\forall a \in \mathbf{A}\backslash\{\hat{a}\}$ and $v_m^* > v_n^*$.

[24] When the agent's utility is additively separable—as it is here—there is no loss in generality from assuming that the (IR) constraint is binding (see Grossman and Hart [1983]).

Define $\mathbf{v}^a$ by

$$v_i^a = \begin{cases} v_i^* \text{ if } i \neq m,n \\ v_0 \equiv \dfrac{\pi_m(\hat{a})v_m^* + \pi_n(\hat{a})v_n^*}{\pi_m(\hat{a}) + \pi_n(\hat{a})} \text{ if } i = m,n \end{cases}.$$

Since $\pi_n(\hat{a})$ cannot be zero, v_0 is well defined. It is trivial to verify that $\mathbf{v}^a$ satisfies the (IR) constraint.

Turning to the (IC) constraints, we have

$$\sum_{i\in \mathbf{N}} \{\pi_i(a)-\pi_i(\hat{a})\}\, v_i^a = \sum_{i\in \mathbf{N}} \{\pi_i(a)-\pi_i(\hat{a})\}v_i^* + \{v_0-v_m^*\}\{\pi_m(a)-\pi_m(\hat{a})\} + \{v_0-v_n^*\}\{\pi_n(a)-\pi_n(\hat{a})\}$$

$$< \sum_{i\in \mathbf{N}} \{\pi_i(a)-\pi_i(\hat{a})\}v_i^* \leq K(a) - K(\hat{a}) \;\; \forall a \in \mathbf{A}.$$

Thus, the contract $\mathbf{v}^a$ implements $\hat{a}$. Since the agent's expected utility is the same, but the contract entails less variability, Jensen's inequality implies that the principal prefers $\mathbf{v}^a$ to $\mathbf{v}^*$, which contradicts the optimality of $\mathbf{v}^*$. **Q.E.D.**

Proof of Proposition 3: Condition (ii) implies $\pi_n^2(\hat{a}) > 0$ for all n. So, if $\hat{a}$ can be implemented at first-best cost under the second returns structure, it follows from Proposition 3 of Grossman and Hart [1983] that it must be a least-cost action. But if $\hat{a}$ is a least-cost action, it can be implemented at first-best cost under either information structure. This completes the proof if $\hat{a}$ is a least-cost action. It also shows that if $\hat{a}$ is not a least-cost action, then it cannot be implemented at first-best cost (*i.e.*, under a full-insurance contract).

Assume, henceforth, that $\hat{a}$ is not a least-cost action. Suppose that $\mathbf{v}^*$ is an optimal contract when the returns structure is Π^2. Pick a constant, $\beta \in (0,1)$, such that

$$\min_{n\in\mathbf{N}} \frac{\pi_n^2(\hat{a})}{\pi_n^1(\hat{a})} > \beta > \max_{m\in\mathbf{N}} \frac{\pi_m^2(a)-\pi_m^2(\hat{a})}{\pi_m^1(a)-\pi_m^1(\hat{a})}.$$

Such a β exists by hypothesis. Define $\mathbf{R} \equiv \beta\mathbf{I} + [\pi^2(\hat{a})] - \beta[\pi^1(\hat{a})]$, where $[\pi]$ is the $N \times N$ matrix in which each column is π. Combining the result that $\pi_n^2(\hat{a}) > 0$ for all n with the definition of β, it follows that every element of $\mathbf{R}$ is positive. Define $\tilde{\mathbf{v}} \equiv \mathbf{R}'\mathbf{v}^*$, so that

$$\tilde{v}_i = \beta v_i^* + \sum_{n\in\mathbf{N}} \left(\pi_n^2(\hat{a}) - \beta\pi_n^1(\hat{a})\right) v_n^*.$$

Define $\mathbf{S} \equiv \left\{i | \pi_i^1(a) - \pi_i^1(\hat{a}) > 0\right\}$. Lastly, define $v_0 \equiv \max_{i\in\mathbf{S}} v_i^*$. Note that, by Lemma A1, $v_0 \le \min_{i\in\mathbf{N}\backslash\mathbf{S}} v_i^*$.

It is trivial to verify that $\tilde{\mathbf{v}}$ satisfies the (IR) constraint when the information structure is Π^1. Now, consider the (IC) constraints. Using the fact that $\sum_{i\in\mathbf{N}} \{\pi_i(a)-\pi_i(\hat{a})\} = 0$, we have

$$\begin{aligned}
&\sum_{i\in\mathbf{N}} \left(\pi_i^1(a) - \pi_i^1(\hat{a})\right)\tilde{v}_i = \sum_{i\in\mathbf{N}} \left(\pi_i^1(a) - \pi_i^1(\hat{a})\right)\beta v_i^* \\
&= \sum_{i\in\mathbf{S}} \left(\pi_i^1(a) - \pi_i^1(\hat{a})\right)\beta\left(v_i^* - v_0\right) + \sum_{i\in\mathbf{N}\backslash\mathbf{S}} \left(\pi_i^1(a) - \pi_i^1(\hat{a})\right)\beta\left(v_i^* - v_0\right) \\
&\le \sum_{i\in\mathbf{S}} \left(\pi_i^2(a) - \pi_i^2(\hat{a})\right)\left(v_i^* - v_0\right) + \sum_{i\in\mathbf{N}\backslash\mathbf{S}} \left(\pi_i^2(a) - \pi_i^2(\hat{a})\right)\left(v_i^* - v_0\right)
\end{aligned}$$

by condition (i) of the Proposition, the construction of β, and the definition of v_0. Simplifying,

$$\begin{aligned}
&\sum_{i\in\mathbf{N}} \left(\pi_i^1(a) - \pi_i^1(\hat{a})\right)\tilde{v}_i \le \sum_{i\in\mathbf{N}} \left(\pi_i^2(a) - \pi_i^2(\hat{a})\right) v_i^* \\
&\le K(a) - K(\hat{a}) \quad \forall a \in \mathbf{A},
\end{aligned}$$

where the last inequality follows from the fact that $\mathbf{v}^*$ is a solution to the agency problem under returns structure Π^2. Hence, the (IC) constraints are satisfied by $\tilde{\mathbf{v}}$ under returns structure Π^1.

It remains to show that the principal prefers contract under returns structure Π^1 to contract $\mathbf{v}^*$ under returns structure Π^2. Define $\mathbf{y} \equiv (y_1, \ldots, y_N)'$, where $y_n = V^{-1}(v_n)$ (adding asterisks or tildes as appropriate). By construction $\tilde{v}_n = \mathbf{r}_{\bullet n}'\mathbf{v}^*$, where $\mathbf{r}_{\bullet n}$ is the nth column of $\mathbf{R}$. Because (i) $V^{-1}(\cdot)$ is a strictly convex function, (ii) $\mathbf{y}^*$ is not a full-insurance contract, and (iii) $\mathbf{r}_{\bullet n}$ has no zero element, Jensen's inequality implies that $\tilde{y}_n < \mathbf{r}_{\bullet n}'\mathbf{y}^*$ for all n. Hence,

$$\pi^1(\hat{a})'\tilde{\mathbf{y}} = \sum_{n\in\mathbf{N}} \pi_n^1(\hat{a})\tilde{y}_n < \sum_{n\in\mathbf{N}} \pi_n^1(\hat{a})\left(\mathbf{r}_{\bullet n}'\mathbf{y}^*\right) = \pi^2(\hat{a})'\mathbf{y}^*,$$

where the final equality follows from the fact that $\mathbf{R}\pi^1(a) = \pi^2(a)$. Therefore, the expected monetary compensation under the first information structure is strictly less than the expected monetary compensation under the second information structure. **Q.E.D.**

Proof of Corollary to Proposition 3: Assumptions (i) and (ii) of the corollary imply that conditions (i) and (ii) of Proposition 3 are satisfied, since $\pi^1(\hat{a}) = \pi^2(\hat{a}) \equiv \hat{\pi}$ implies that $\min_{n\in N} \pi_n^2(\hat{a})/\pi_n^1(\hat{a}) = 1$. **Q.E.D.**

Proof of Lemma 1: The problem of implementing $a = 1$ can be expressed as

$$\text{minimize } \textstyle\sum_{i\in\mathbf{N}} \pi_i(1)\, V^{-1}(v_i)$$

$$\text{subject to} \quad \textstyle\sum_{i\in\mathbf{N}} \pi_i(1) v_i = K(1) \quad \text{and} \quad v_1 \le K(0).$$

By the convexity of $V^{-1}(\cdot)$, the solution is $v_1 = K(0)$ and $v_j = \{K(1) - \pi_1(1)K(0)\}/\{1-\pi_1(1)\} \equiv \bar{v}$ for all $j\neq 1$. The principal's expected cost of implementing $a = 1$ is

$$C \equiv \pi_1(1)\, V^{-1}(v_1) + \{1-\pi_1(1)\}\, V^{-1}(\bar{v}).$$

Total differentiation with respect $\pi_1(1)$ yields

$$\frac{\mathrm{d}C}{\mathrm{d}\pi_1(1)} = y_1 - \left[\bar{y} - \frac{\mathrm{d}V^{-1}}{\mathrm{d}v} \times \left(\frac{K(1) - K(0)}{1 - \pi_1(1)}\right)\right],$$

where $y_1 = V^{-1}(v_1)$ and $\bar{y} = V^{-1}(\bar{v})$. The expression in brackets is the first-order Taylor series approximation of y_1 starting from $\bar{y}$. Since $V^{-1}(\cdot)$ is a convex function, it follows that this Taylor series approximation is less than y_1. Therefore, $\mathrm{d}C/\mathrm{d}\pi_1(1) > 0$. **Q.E.D.**

Proof of Proposition 5: Proposition 9 of Grossman and Hart [1983] establishes that, under Assumptions A1-A4, the optimal contract, $y_{n+1} - y_n$, is non-increasing in n. Since $V(\bullet)$ is concave, it follows that $v_{n+1} - v_n$ is also non-increasing in n. This and Assumption A5 imply condition (4) is satisfied in the Corollary to Proposition 1. **Q.E.D.**

Proof of Proposition 6: Suppose, contrary to the proposition, that the pair $(\hat{a},1)$ is not implementable (where (a,i) means action a and returns structure i). By Proposition 2 of Hermalin and Katz [1991], there exist sets $\mathbf{A}_1 \subseteq \mathbf{A}$ and $\mathbf{A}_2 \subseteq \mathbf{A}$ and a set of positive weights $\{\mu(a),\lambda(a)\}$ summing to one such that

$$\pi^1(\hat{a})' = \sum_{a\in\mathbf{A}_1} \mu(a)\pi^1(a)' + \sum_{a\in\mathbf{A}_2} \lambda(a)\pi^2(a)' \tag{5}$$

and

$$K(\hat{a}) > \sum_{a\in\mathbf{A}_1} \mu(a)K(a) + \sum_{a\in\mathbf{A}_2} \lambda(a)K(a). \tag{6}$$

Post-multiplying (5) by $\mathbf{x}$ and simplifying yields

$$\pi^1(\hat{a})'\mathbf{x} = \sum_{a\in\mathbf{A}_1\cup\mathbf{A}_2} \hat{\mu}(a)\pi^1(a)'\mathbf{x}, \tag{7}$$

where

$$\hat{\mu}(a) = \frac{\mu(a)\cdot \mathbf{1}_{\{a\in \mathbf{A}_1\}} + \lambda(a)\cdot \mathbf{1}_{\{a\in \mathbf{A}_2\}}}{1-\lambda(\hat{a})\cdot \mathbf{1}_{\{a\in \mathbf{A}_2\}}},$$

where $1_{\{\bullet\}}$ is an indicator function. By construction, $\hat{\mu}(a) > 0$ and the sum of $\hat{\mu}(a)$ over a in $(\mathbf{A}_1 \cup \mathbf{A}_2)\backslash\{\hat{a}\}$ is one. Carrying out the same simplification on (6) yields

$$K(\hat{a}) > \sum_{a\in \mathbf{A}_1\cup\mathbf{A}_2} \hat{\mu}(a)K(a). \tag{8}$$

But (7) and (8) contradict the assumption that $\hat{a}$ satisfies the convexity of disutility property.

Q.E.D.

Proof of Proposition 7: Suppose the equilibrium contract, $\mathbf{v}^*$, implements (a^*,i=2). Define $\tilde{\mathbf{v}} \equiv \mathbf{Q}^{\infty\,\prime}\mathbf{v}^*$, and note that $\pi^1(a)'\tilde{\mathbf{v}} = \pi^2(a)'\mathbf{v}^*$.

We will now show that the contract $\tilde{\mathbf{v}}$ implements (a^*,i=1). Consider first the (IR) constraint. Since

$$\pi^1(a^*)'\tilde{\mathbf{v}} - K(a^*) = \pi^2(a^*)'\mathbf{v}^* - K(a^*) = 0,$$

the (IR) constraint is satisfied. Next consider the (IC) constraints. Because the choice of returns structure is endogenous, we have two sets of constraints

$$\{\pi^1(a)'-\pi^1(a^*)'\}\tilde{\mathbf{v}} \le K(a) - K(a^*) \quad \forall a\in \mathbf{A}, \tag{9}$$

and

$$\{\pi^2(a)'-\pi^1(a^*)'\}\tilde{\mathbf{v}} \le K(a) - K(a^*) \quad \forall a \in \mathbf{A} \tag{10}$$

Condition (9) is satisfied because

$$\pi^1(a)' \tilde{\mathbf{v}} = \pi^2(a)'\mathbf{v}^*,$$

and $\mathbf{v}^*$ implements a^* under Π^2. Consider (10). Using the fact that $\mathbf{Q}^\infty$ must be an idempotent matrix (*i.e.*, $\mathbf{Q}^\infty\mathbf{Q}^\infty = \mathbf{Q}^\infty$), we have

$$\pi^2(a)' \tilde{\mathbf{v}} = \pi^1(a)'\mathbf{Q}^{\infty\prime}\mathbf{Q}^{\infty\prime}\mathbf{v}^* = \pi^1(a)'\mathbf{Q}^{\infty\prime}\mathbf{v}^* = \pi^2(a)'\mathbf{v}^*.$$

This, the fact that $\pi^1(a^*)' \tilde{\mathbf{v}} = \pi^2(a^*)'\mathbf{v}^*$, and the fact that $\mathbf{v}^*$ implements a^* under Π^2, imply (10).

Next, we show that $\tilde{\mathbf{y}}$ has lower expected cost than $\mathbf{y}^*$ (where, as before, $y_n = V^{-1}(v_n)$). $\tilde{v}_n = \mathbf{q}_{\bullet n}'\mathbf{v}^*$, where $\mathbf{q}_{\bullet n}$ is the nth column of $\mathbf{Q}^\infty$. Since $V^{-1}(\cdot)$ is a convex function, Jensen's inequality implies that $\tilde{y}_n \le \mathbf{q}_{\bullet n}'\mathbf{y}^*$ for all n. This, in turn, implies that $\pi^1(a^*)' \tilde{\mathbf{y}} \le \pi^2(a^*)'\mathbf{y}^*$. Thus, the expected monetary payments made by the principal to the agent are lower when a^* is implemented under returns structure 1 rather than 2. Since $\pi^1(a^*)'\mathbf{x} = \pi^2(a^*)'\mathbf{x}$, the expected gross returns are identical. It follows that the owners (weakly) prefer implementing (a^*,i=1) to implementing (a^*,i=2). **Q.E.D.**