

CORPORATE DIVERSIFICATION AND AGENCY

**Benjamin E. Hermalin
and
Michael L. Katz**

University of California at Berkeley

10 January 2000

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 the 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

Why, then, do firms undertake activities to diversify risk? Many analysts have suggested diversification is the result of the agency relationship that exists between managers and outside investors.³ The building blocks for an agency-based explanation of diversification are these: Risk-averse owners are 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 would also like to diversify away the risks associated with their firms, but

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 Porter [1987] also makes this point from the perspective of corporate strategy. In recent years, several firms have unbundled their securities, either by outright divestiture of operations or by creating separate classes of tracking stocks, each of whose returns are tied to the performance of a specific division within the firm.

³ We briefly discuss alternative explanations in the concluding section.

⁴ We do not explore the possibility that the firm is controlled by a "large shareholder" who is undiversified. See Zhang [1998] for an analysis of investment choice in the presence of such an owner.

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. Moreover, through the effects on her reputation, a manager's future income may depend on the returns of the one firm that she currently manages. Lacking the opportunity to diversify in other ways (*e.g.*, part-time managerial positions with multiple firms), managers may value within-firm diversification. Intuitively, even diversified shareholders benefit from a reduction in the riskiness of the firm's profit stream when the reduction lowers the compensation level needed to attract and retain (risk-averse) managers, or reduces the distortions in managers' investment decisions.

In analyzing the effects of agency on diversification incentives, however, it is essential to recognize that the relationship between the riskiness of a firm's returns and its manager's well-being is largely endogenous: The relationship is driven by firm's choice of managerial compensation scheme.⁵ As is well known, a risk-neutral principal has incentives to insure a risk-averse agent, but in the presence of a moral hazard the principal must trade off the provision of insurance against the provision of incentives. As is also well known, the terms of this trade-off depend on the degree to which the principal is informed about the agent's actions.⁶ One of our central concerns in the present paper is whether riskiness matters as well.

Several earlier authors have (indirectly) addressed the issues of how diversification affects information and riskiness, and whether riskiness matters. 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 risk reduction benefits shareholders by reducing agency costs.

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.

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

Unfortunately, the 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] argue that diversification can ameliorate agency problems 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. 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. She examines a specific model in which diversification into additional lines of business generates new signals 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. Aron does not formalize the concept of increased informativeness, however. Indeed, in her comparison of diversification and the use of relative performance schemes, she focuses her answer on the correlation of returns across projects which is not directly related to informativeness.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

⁷ As Holmstrom [1979] has shown, informativeness is related to statistical sufficiency, which is different than correlation. Consider, for example, two random variables s_1 and s_2 . As long as there exists a function, $t(\cdot)$, such that $t(s_1) = s_2$, s_1 is a sufficient statistic for s_2 regardless of their

resulting returns.⁸ Financial instruments, such as insurance contracts and foreign exchange options, similarly change the relationship between the firm's returns and the management's characteristics and actions. When executive compensation contracts depend on measures of firm performance, such as the level of sales or profits, the value of the measure is of interest not only as a signal of the agent's effort or ability, but also because it represents the financial returns to ownership. In evaluating corporate investments, 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 a firm may be willing to trade off expected returns for improved information. A more difficult question is how a firm is willing to trade improvements in informativeness against increases in risk, or whether risk even matters.⁹ 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

correlation.

⁸ See Mirrlees [1976] for an early and insightful analysis of the relationship between agency and firm structure.

⁹ In a model of hidden information, rather than the hidden-action case examined in the present paper, Hermalin [1993] shows 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 the riskiness of the firm's returns is increased. DeMarzo and Duffie [1995] examine a similar model and find that certain risk-reducing actions may reduce informativeness, to the manager's advantage.

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 demonstrate through examples 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. We also examine the use of hedging and insurance contracts and suggest that agency does not provide a good explanation of the use of these contracts.

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 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.¹⁰ Moreover, even in the case of an acquisition, shareholders may delegate

¹⁰ There is a sizable—and inconclusive—empirical literature that examines whether managers pursue acquisitions, conglomerate and otherwise, at shareholders' expense. For example, 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 decision to the agent because they lack the information needed to discern whether the transaction valuable to the firm when evaluated solely in terms of its expected payoff. Thus, it is worth investigating what happens when the manager controls the diversification decision after the principal has set the compensation scheme.¹¹ 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. 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.

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

A. The Model

Consider a standard principal-agent model similar to Grossman and Hart's [1983]. A risk-neutral owner 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} . $K(\cdot)$, the disutility-of-action function, is increasing; $V(\cdot)$, the utility of money function, is strictly increasing, strictly concave, and has an unbounded range. The owner's 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 $\mathbf{x} \equiv (x_1, x_2, \dots, x_N)'$ represents the set of possible

¹¹ We do not examine the (to us) less plausible case where the structure is not verifiable and is controlled by the principal.

revenue levels.¹² 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 i th returns structure conditional on action a having been taken. The density over revenue levels conditional on action a is

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

As is well known, the returns in this agency problem play two roles. First, 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

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

and

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

Second, 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.

B. Blackwell Informativeness and Risk in Agency Settings

The notion of Blackwell informativeness provides a starting point for our analysis. This is the notion that, given the choice between two observation strategies—observe the outcome of an experiment or observe the outcome with noise—a rational experimenter would prefer the former to the latter. Hence, if one information structure is a noisy transformation of a second, then the second is more informative in a Blackwell sense.

Consider two standard agency problems that are identical except for the stochastic

¹² The prime (') denotes vector or matrix transpose.

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 (*i.e.*, 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.

Grossman and Hart [1983, Proposition 13] show how the incentive-insurance trade-off in the design of the agent's compensation scheme is affected by an improvement in information in the Blackwell sense. They establish that, if there exists a stochastic transformation matrix \mathbf{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.

While distinct concepts, Blackwell informativeness and riskiness are closely related, and this relationship can make it difficult to determine which property drives agency costs. 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 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 $a \in \mathbf{A}$.¹³ Economically, one can interpret a mean-preserving garbling as follows. Under returns structure Π^1 , the payoff if the n th state

¹³ 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. Because little use can be made (by us) of the 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 .

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

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 Π^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 the following theorem of Blackwell:¹⁴

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, \mathbf{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:¹⁵

¹⁴ Marshall and Olkin [1979, p.417].

¹⁵ Proofs not in the text may be found in the Appendix.

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.*

While the first two results make distinctions between informativeness and riskiness difficult, the following observation helps distinguish between the two concepts:

Remark: *Let Π^1 and Π^2 be two returns structures, where: (i) Π^2 is of full rank and (ii) the expected returns conditional on the action taken are the same under the two returns structures. Π^1 and Π^2 are equally informative in the Blackwell sense if and only if there exists a permutation matrix \mathbf{M} , such that $\pi^2(a) = \mathbf{M}\pi^1(a)$ for all actions a .*

This remark strongly suggests that riskiness is irrelevant: When two returns structures are equally informative, the cost of implementing any given action is the same under both structures, yet the riskiness of the two returns structures could be vastly different.

C. Intuitive Informativeness and Risk in Agency Settings

Like Gjesdal [1981] and Kim [1995], we believe that Blackwell's definition of informativeness is unduly restrictive for agency purposes and thus often fails to provide a ranking in situations in which one information structure is in fact superior to the other. Very loosely speaking, the principal cares only about whether he can distinguish between the agent's taking a single desired action and all other actions, and Blackwell informativeness does not make use of this fact.

It is well known 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.¹⁶ 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

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

It is straightforward to verify that, for both a , $\boldsymbol{\pi}^2(a)$ is less risky than $\boldsymbol{\pi}^1(a)$.¹⁷ Under returns structure 2, 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 returns structure 1, however, obtaining a revenue of 2 is proof that the manager took action $a = 0$, and the owner can thus implement $a = 1$ at the full-information cost. Therefore, the owner prefers the riskier returns structure 1.

Example 1 illustrates the fact that Blackwell informativeness is an imperfect measure of informativeness in an agency problem—while the two returns structures could not be ranked according to the Blackwell criterion, the second structure clearly is more informative in an intuitive sense. Our next result shows that this point is not dependent on the existence of a shifting support.

This result builds on the following notion of informativeness. Suppose that the principal

¹⁶ For a subtle and elegant generalization of the so-called “shifting support” result, see Mirrlees [1999].

¹⁷ Note that there is no *common* stochastic transformation matrix such that $\boldsymbol{\pi}^1(a) = \mathbf{Q}\boldsymbol{\pi}^2(a)$. Because $\pi_2^1(1) = 0$, the middle row of \mathbf{Q} would necessarily consist entirely of zeros, which would be inconsistent with $\pi_2^1(0) = 1/3$. Thus, the returns structures cannot be ranked by the criterion of Blackwell informativeness.

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 \hat{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

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

and

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

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

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

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

because compensation in state n is an increasing function of

$$\mu^i - \lambda^i(a) h_n^i(a, \hat{a}), \tag{2}$$

where μ^i and $\lambda^i(a)$ are positive constants. Proposition 3's two conditions are conditions on the

way the $h_n^i(a, \hat{a})$ are distributed under the two information systems. As such, Proposition 3 is related to Kim's [1995] 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. It is possible to show by example that information structures can satisfy our conditions without their likelihood ratios being ordered by second-degree stochastic dominance as required by Kim.¹⁸

The important point to note is that $\pi^1(\hat{a})$ and $\pi^2(\hat{a})$ may satisfy our conditions even though $\pi^1(\hat{a})$ is riskier than $\pi^2(\hat{a})$, as Example 1 above illustrates. This point can be seen more generally 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 $\mathbf{x} = (0, 100, 200)'$,

$$\boldsymbol{\pi}^1(a_1) = \boldsymbol{\pi}^2(a_1) = (.395, .48, .125)',$$

$$\boldsymbol{\pi}^1(a_2) = \boldsymbol{\pi}^2(a_2) = (.32, .48, .2)',$$

$$\boldsymbol{\pi}^1(a_3) = (.48, .04, .48)',$$

and

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

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

¹⁸ An example can be constructed when there are two states and two actions, a and \hat{a} , such that $\pi_1^1(a) = .7$, $\pi_1^1(\hat{a}) = .3$, $\pi_1^2(a) = .7$, and $\pi_1^2(\hat{a}) = .35$. Details available upon request. Stochastic dominance is not a necessary condition because, although the principal's costs are convex in (2), the weights μ^i and $\lambda^i(a)$ also depend on the information.

structure than under the second one. 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 returns structures being compared while varying the densities associated with other actions. By doing this, riskiness is held constant across the two returns structures but informativeness may vary. The following 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*

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

and

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

*then the cost of implementing \hat{a} is less under returns structure Π^1 than under Π^2 .*¹⁹

The corollary underscores again that when risk and informativeness are not related, informativeness matters and risk does not.

3. RISK AND INFORMATION EFFECTS OF DIVERSIFICATION AND HEDGING

A. 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. There are one or two projects, depending on the case under consideration. The agent chooses whether to apply 0, 1, or

¹⁹ The corollary also follows from Kim's [1995] main proposition suitably modified for a discrete state space.

2 units of effort to each available project. A project succeeds with probability $\gamma\alpha$, where $\gamma \in (0, \frac{1}{2})$ and α is the effort allocated to *that* project; otherwise it fails. The projects' successes are statistically independent. The agent's cost of effort depends solely on the *total* effort, a , not the allocation between projects. $K(a)$ is finite if $a \in \{0, 1, 2\}$ and is infinite otherwise. The owner's gross benefit is $x_S > 0$ from a project that succeeds. We normalize the payoff of failure as 0. If two projects succeed, the owner's gross benefit is $2x_S$. Notice that the expected gross benefit for a given level of total effort, γax_S , is independent of how that effort is allocated across projects. 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.²⁰

Because the projects are *ex ante* identical and independent, there is no loss in generality from assuming that if the manager conducts only one project, she conducts the first. When both projects are available, the manager chooses actions from the set $\{(0,0), (1,0), (2,0), (1,1)\} \equiv \mathbf{A}_2$, where a pair represents effort allocated to the first and second project, respectively. Observe that the manager's choice set when only one project is available, \mathbf{A}_1 , is equivalent to the subset consisting of the first three elements of \mathbf{A}_2 .

Because \mathbf{A}_1 and \mathbf{A}_2 are of different sizes, the resulting information structures cannot be directly compared using Blackwell informativeness or the methods of Proposition 3. However, the problem can be analyzed by other means.

Lemma 1: *Suppose $K(\cdot)$ is affine (or concave) on $\{0, 1, 2\}$. The cost of inducing $(2,0)$ when the*

²⁰ 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.

agent's action set is \mathbf{A}_1 is equal to the cost of inducing $(2,0)$ when the action set is $\{(0,0), (2,0)\}$.

Proof: Let $\mathbf{v} \equiv (v_0, v_1) \equiv [V(y_0), V(y_1)]$ denote a contract stated in terms of the agent's utility. Then the incentive compatibility (IC) constraint for the agent to choose $(2,0)$ instead of $(0,0)$ is

$$2\gamma v_1 + (1-2\gamma)v_0 - K(2) \geq v_0 - K(0),$$

which is equivalent to

$$\gamma(v_1 - v_0) \geq [K(2) - K(0)]/2. \quad (3)$$

Using the fact that $[K(2) - K(0)]/2 \geq K(2) - K(1)$, inequality (3) implies

$$2\gamma v_1 + (1-2\gamma)v_0 - K(2) \geq \gamma v_1 + (1-\gamma)v_0 - K(1),$$

which is the IC constraint for the choice of $(2,0)$ instead of $(1,1)$. **Q.E.D.**

Proposition 4: *Suppose $K(\cdot)$ is affine (or concave) on $\{0,1,2\}$. The cost of implementing $a = 2$ is independent of whether diversification is feasible.*

Proof: Suppose the choice sets are $\{(0,0), (2,0)\}$ and $\{(0,0), (1,1)\}$. The following stochastic matrix is a garbling that maps the former returns structure into the latter:

$$\begin{pmatrix} 1 & \frac{\gamma}{2} & 0 \\ 0 & 1-\gamma & 0 \\ 0 & \frac{\gamma}{2} & 1 \end{pmatrix}$$

Because adding new possible actions cannot reduce the principal's cost of inducing an existing action, it follows that the cost of inducing $a = 2$ under \mathbf{A}_2 is no less than the minimum of the cost when the choice set is $\{(0,0), (1,1)\}$ and the cost when the choice set is $\{(0,0), (2,0)\}$. By the first part of this proof, this minimum is no less than the cost when the choice set is $\{(0,0), (2,0)\}$.

Thus, by Lemma 1, the cost of inducing $a = 2$ under \mathbf{A}_2 is no less than the cost under \mathbf{A}_1 . On the other hand, because there is a shifting support when the agent chooses $(1,1)$ —this is the only way

she could get two successes—the principal can costlessly force the agent *not* to take (1,1).

Therefore, the cost of inducing (2,0) under \mathbf{A}_2 equals the cost of inducing it under \mathbf{A}_1 . **Q.E.D.**

We have shown that there is no advantage to making diversification feasible, but—because the principal can costlessly block it—there is no disadvantage either. It is nonetheless informative to consider the contrived case where the principal must choose between diversification and non-diversification; that is between the action sets $\mathbf{A}_d \equiv \{(0,0),(1,0),(1,1)\}$ and \mathbf{A}_1 , respectively. When $K(\cdot)$ is affine or concave, the existence of the non-trivial garbling that maps $\{(0,0), (2,0)\}$ into $\{(0,0), (1,1)\}$ implies that the principal would be strictly worse off given the latter two-element action set than the former because the agent is strictly risk averse and $a = 2$ is not a least-cost action for the agent. As Lemma 1 showed, the cost of implementing $a = 2$ given the former two-element action set is the same as the cost under the larger set \mathbf{A}_1 . Because adding undesired actions cannot reduce the principal’s cost, the cost of implementing $a = 2$ in the latter two-element action set is no greater than under the larger set \mathbf{A}_d . By transitivity, implementing $a = 2$ without diversification (*i.e.*, implementing (2,0) under \mathbf{A}_1) costs strictly less than implementing $a = 2$ with diversification (*i.e.*, implementing (1,1) under \mathbf{A}_d).

Now suppose K convex. An argument similar to that used to prove Lemma 1 establishes:

Lemma 2: *Suppose $K(\cdot)$ is convex on $\{0,1,2\}$. The cost of inducing (2,0) when the agent’s action set is \mathbf{A}_1 is equal to the cost of inducing (2,0) when the action set is $\{(1,0), (2,0)\}$.*

We can show by example that it can be cheaper to induce $a = 2$ when the choice set is $\{(1,0), (2,0)\}$ than when it’s $\{(1,0), (1,1)\}$.²¹ By Lemma 2, the cost under \mathbf{A}_1 is equal to the cost under the first action set. Because adding undesired actions cannot reduce the principal’s cost of

²¹ By numerical simulation, it can be shown that when $V(y) = \ln(y)$, $\gamma \in (1/20, 9/20)$, $K(0) = 0$, $K(2) = 2$, and $K(1) = \theta$, $\theta \in (1/10, 9/10)$, then it’s cheaper to induce $a = 2$ in the first case than it is in

inducing an existing desired action, the cost of inducing (1,1) under \mathbf{A}_d is no less than the cost when the choice set is $\{(1,0), (1,1)\}$. Therefore, this example demonstrates that it can be strictly cheaper to induce $a = 2$ without diversification (*i.e.*, under \mathbf{A}_1) than by diversifying (*i.e.*, under \mathbf{A}_d) regardless of whether $K(\cdot)$ is affine, concave, or convex.

In the situations identified above, *conditional* on $a = 2$, diversification is actually riskier than focusing on a single project.²² Moreover, diversification reduces informativeness. By splitting effort, the probability of being successful falls for any given project. And success serves as a signal of high effort. Thus, although now there are two projects to observe, the cumulative effects of the decrease in the informativeness of each project outweighs the information gained by having two projects.

This conclusion may appear at odds with Aron's [1988] findings. Her article appears to show that diversification helps ameliorate the incentive problem. The difference lies in our assumptions about whether diversifying lessens the impact of the agent's action on a given project or not. We assume that the *cost* of diversification is that it reduces the probability that any given project will succeed; that is, by splitting a given level of total effort among multiple projects, the probability that any one project succeeds will fall. In contrast, Aron essentially assumes that the probability of success is not diminished, but the reward associated with success is.²³

When diversification leaves the probability of success unaffected and simply lowers the

the second.

²² It is readily shown that this result is not the consequence of having assumed independence between the projects.

²³ In our setting, this approach would correspond to having the probability of success of each project be γa , where a is the agent's total effort. The payoffs from a project's success would be $x_S/2$ if two projects are chosen and x_S if only one project is undertaken.

payoff from success, diversification increases informativeness. This is not surprising: The sample size doubles without any degradation of the signal. On the other hand, this type of diversification strikes us as more likely to be the exception than the rule. In many situations, we would expect that division of effort among multiple projects (*e.g.*, different R&D directions) would reduce the probability that any given project would succeed. In that case, diversification can decrease informativeness. When diversification changes the quality of the performance signals, the impact of diversification on informativeness—and thus the impact on the principal’s equilibrium payoffs—is ambiguous.

B. Hedging

Many corporations use financial instruments, such as insurance and foreign exchange futures contracts to hedge risk. As with corporate diversification, economists have pointed to agency as an explanation of why firms with risk neutral owners would engage in such behavior.²⁴ Let $p_i(\omega)$ denote the return from holding one unit of hedging security i when the state is ω , and let $\mathbf{p}(\omega)$ denote the corresponding vector of payoffs for all hedging securities. ω comprises the verifiable variables on which the payoffs of the financial contracts depend, and $\mathbf{p}(\cdot)$ maps these variables into payoffs. Let \mathbf{m} denote the firm’s holdings of hedging securities. Then, the firm’s payoff is $z = x + \mathbf{m}'\mathbf{p}(\omega)$, where as before x is the firm’s revenue. Note that if the elements of $\mathbf{p}(\omega)$ are independent of the agent’s action, a , then there can be no benefit to contracting on z instead of x — x is a sufficient statistic for a .²⁵ Indeed, if $\mathbf{p}(\omega)$ is noisy, then z is less informative than x and the firm’s owner is worse off contracting on z instead of x .

There is no need to take non-zero positions in these securities even when the distribution

²⁴ For analyses of incentives to reduce risk through financial contracts, see Smith and Stulz [1985] and DeMarzo and Duffie [1995].

of $p(\omega)$ depends on a , and thus the realization of $p(\omega)$ might serve as an informative signal of the agent's action. Because ω is verifiable, the owner can condition the agent's compensation on z *without purchasing any securities*. In fact, because there is little reason to expect that the optimal contract would be a function of the *sum* of x and $\mathbf{m}'p(\omega)$, contracting on z is very likely worse than contracting directly on x and ω . Thus, in our model, if he has the information needed to do so, the owner should forbid the purchase of any hedging securities that offer negative expected returns (*e.g.*, insurance contracts). If, however, the agent has better information about the expected returns, then it may be inefficient for the owner to control \mathbf{m} directly. In such settings, hedging could be a factor in agency problems. We explore issues of who controls the information structure in the next section.

4. WHO CONTROLS THE INFORMATION STRUCTURE?

So far, we have assumed that the owner decides whether the firm should diversify or not. In many situations, this is a reasonable assumption: A merger or acquisition often requires shareholder approval, for example, and is of sufficient importance that it can pay owners to become informed about the transaction. On the other hand, there are diversification decisions (*e.g.*, the number of paths to pursue in developing a new product) that managers make with little direct oversight. For there to be a meaningful distinction between situations in which the principal chooses the information structure and those in which the agent chooses, the principal must not be able to make the agent's compensation contingent on the choice of information structure in the latter case. While this inability may be 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.

²⁵ See Holmstrom [1979].

We can embed the choice of returns structure in a standard agency problem in which 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 owner can specify the returns structure in the contract (*i.e.*, the diversification choice is verifiable) the *direct-choice problem*. 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 owner indirectly chooses the returns structure through the incentive scheme they offer to his manager.²⁶

Obviously, the owner's expected profit in the indirect-choice problem is no greater than his expected profit in the direct-choice problem. The first question we ask is: When is his expected profit no less? 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). Recall $v_i \equiv V(y_i)$, 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:*

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

²⁶ The indirect-choice model is an example of a multi-task principal-agent model. For more on multi-task principal-agent problems see Holmstrom and Milgrom [1991].

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

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 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 [1983] 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 that the following conditions are satisfied:*

- A1 (No Shifting Support): $\pi_n^i(a) > 0$ for all $a \in \mathbf{A}$, $n \in \mathbf{N}$, and $i \in \{1,2\}$.
- A2 (Monotone Likelihood Ratio Property): For all a and \hat{a} in \mathbf{A} , $K(a) < K(\hat{a})$ implies that $\pi_n^i(a) / \pi_n^i(\hat{a})$ is non-increasing and convex in n and $i \in \{1,2\}$.
- A3 (Concavity of Distribution Function Property): $K(a) = \gamma K(\hat{a}) + (1-\gamma)K(\tilde{a})$ ($\gamma \in [0,1]$) implies $\mathbf{P}^i(a) \leq \gamma \mathbf{P}^i(\hat{a}) + (1-\gamma)\mathbf{P}^i(\tilde{a})$, where $\mathbf{P}^i(a)$ is the cumulative distribution function vector over revenues induced by action a (i.e., $\mathbf{P}^i(a) = (\pi_1^i(a), \pi_1^i(a) + \pi_2^i(a), \dots, \pi_1^i(a) + \dots + \pi_N^i(a))'$), $i \in \{1,2\}$.
- 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 owner an expected profit equal to what he 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. In such cases, the manager prefers the less informative returns structure and the owner's expected profit is strictly less in the indirect-choice problem. 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 $\pi(a)'x = \sum_{j \in J} \lambda_j \pi(a_j)'x$ implies that $K(a) \leq \sum_{j \in J} \lambda_j K(a_j)$, where $\{\lambda_j\}$ is a set of non-negative weights summing to one and J is an index set for actions other than a .²⁷*

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 .

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 \hat{a} 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 \hat{a} and induces the manager to choose the more informative returns structure.*

The convexity of disutility property is a more stringent property than implementability in the

²⁷ This property is related to the Concavity of Distribution Function Property above. It differs in that it applies to any convex combination of actions and is based on expected values rather than first-degree stochastic dominance.

direct-choice problem.²⁸

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

Example 3: Suppose $V(y) = \ln(y)$, $\mathbf{A} = \{0,1\}$, $K(0) = 0$, $K(1) = 1$, $\boldsymbol{\pi}^1(0) = (.5, .25, .25)'$, $\boldsymbol{\pi}^1(1) = (.4, .05, .55)'$, 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 $\boldsymbol{\Pi}^2 = \mathbf{Q}\boldsymbol{\Pi}^1$, where \mathbf{Q} is the mean-preserving garbling

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

Under either information structure, the optimal contract for implementing $a = 0$ is to pay the agent $y = 1$ for all outcomes. The owner's 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 owner's expected profit is 10.48. Therefore, the owner will induce the agent to choose the less informative information structure, $i=2$.

Under either returns structure, the optimal contract for implementing $a = 1$ is one that punishes the manager for achieving $x = 12$ (since $x = 12$ is a relatively rare event if $a = 1$, but a

²⁸ Implementability of a in the direct-choice problem requires only that there exists *some* vector \mathbf{v} such that, if $\boldsymbol{\pi}(a)' \mathbf{v} = \sum_{j \in \mathbf{J}} \lambda_j \boldsymbol{\pi}(a_j)' \mathbf{v}$, then $K(a) \leq \sum_{j \in \mathbf{J}} \lambda_j K(a_j)$, where $\{\lambda_j\}$ is a set of non-negative weights summing to one.

relatively more common event if $a = 0$). Consequently, the manager's utility is a convex function of revenue, so she prefers the riskier—and therefore less informative—returns structure.

In this example, because the two returns structures are very “close” to one another and there is relatively little loss of information. Would the owner implement the less informative returns structure if the two returns structures were “far apart”? The following proposition shows the answer is no when one interprets 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}^∞ as n goes to ∞ . In the indirect-choice problem with returns structures Π^1 and $\Pi^2 \equiv \mathbf{Q}^\infty \Pi^1$, the owner will choose a contract that induces the agent to choose the more informative returns structure, Π^1 .*

5. CONCLUSION

We explored agency 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, owners of the firm potentially care about the returns’ mean, 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 *per se*, one might observe a positive correlation between diversification and risk reduction. To explore this issue further, we proposed alternative 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 3, 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.²⁹ 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 because a single failing project may be "saved" by other (successful) projects, or it

²⁹ For an early example of a model in which this occurs, see Lewellen [1971].

may drag all of them down.³⁰ 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.³¹ 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 administrative ones, not the loss in the value of firm's productive activities.

³⁰ For more on this point, see Higgins' [1971] discussion of Lewellen [1971].

³¹ Several authors have sought to model these issues formally. See, for example, Aghion and Bolton [1992], Harris and Raviv [1990], and Hart [1991].

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.
- DeMarzo, Peter M., and Darrell Duffie. 1995. "Corporate Incentives for Hedging and Hedge Accounting." *The Review of Financial Studies* **8**: 743-771.
- 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.
- 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.
- Mirrlees, James A. 1976. "The Optimal Structure of Incentives and Authority within an Organization." *Bell Journal of Economics* **7**:105-31.
- _____ 1999. "The Theory of Moral Hazard and Unobservable Behaviour: Part I." *Review of Economic Studies* **66**:3-21.
- 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.
- Smith, Clifford W., and Rene M. Stulz. 1985. "The Determinants of Firm's Hedging Policies." *Journal of Financial and Quantitative Analysis*, **20**:391-405.
- Zhang, Guochang. 1998. Ownership concentration, risk aversion and the effect of financial structure on investment decisions." *European Economic Review*, **42**:1751-1778.

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, $\Pi^2(\mathbf{R} - \mathbf{I})'\mathbf{x} = \mathbf{0}$. 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.

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

and

$$\sum_{i \in \mathbf{N}} \{\pi_i(a) - \pi_i(\hat{a})\}v_i \leq K(a) - K(\hat{a}) \quad \forall a \in \mathbf{A}.^{32} \quad (\text{IC})$$

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

Proof of Lemma A1: Suppose that, contrary to the hypothesis of the lemma,

$$\pi_m(a) - \pi_m(\hat{a}) > 0 > \pi_n(a) - \pi_n(\hat{a}) \quad \forall a \in \mathbf{A} \setminus \{\hat{a}\} \quad \text{and} \quad v_m^* > v_n^*.$$

Define \mathbf{v}^a by

³² 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]).

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

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

$$\begin{aligned} \sum_{i \in \mathbf{N}} \{\pi_i(a) - \pi_i(\hat{a})\} v_i^a &= \sum_{i \in \mathbf{N}} \{\pi_i(a) - \pi_i(\hat{a})\} v_i^* + \{v_0 - v_m^*\} \{\pi_m(a) - \pi_m(\hat{a})\} + \{v_0 - v_n^*\} \{\pi_n(a) - \pi_n(\hat{a})\} \\ &< \sum_{i \in \mathbf{N}} \{\pi_i(a) - \pi_i(\hat{a})\} v_i^* \leq K(a) - K(\hat{a}) \quad \forall a \in \mathbf{A}. \end{aligned}$$

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 . Thus, 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

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

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

$$v_0 \leq \min_{i \in \mathbf{N} \setminus \mathbf{S}} v_i^*.$$

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

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

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

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

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

It remains to show that the principal prefers contract under returns structure $\mathbf{\Pi}^1$ to contract \mathbf{v}^* under returns structure $\mathbf{\Pi}^2$. Define $\mathbf{y} \equiv (y_1, \dots, y_N)'$, 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 n th 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,

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

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

$$\min_{n \in \mathbf{N}} \pi_n^2(\hat{a}) / \pi_n^1(\hat{a}) = 1. \quad \mathbf{Q.E.D.}$$

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) denotes 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

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

and

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

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

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

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

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

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 $\tilde{\mathbf{v}} \equiv \mathbf{Q}^\infty' \mathbf{v}^*$, 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

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

and

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

Condition (9) is satisfied because

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

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

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

This, the fact that $\boldsymbol{\pi}^1(a^*)' \tilde{\mathbf{v}} = \boldsymbol{\pi}^2(a^*)' \mathbf{v}^*$, and the fact that \mathbf{v}^* implements a^* under $\boldsymbol{\Pi}^2$, 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 n th column of \mathbf{Q}^∞ . Since $V^{-1}(\cdot)$ is a convex function, Jensen's inequality implies that $\tilde{y}_n \leq \mathbf{q}_{\bullet n}' \mathbf{y}^*$ for all n . This, in turn, implies that $\boldsymbol{\pi}^1(a^*)' \tilde{\mathbf{y}} \leq \boldsymbol{\pi}^2(a^*)' \mathbf{y}^*$.

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