Performance Measure Congruity and Diversity in Multi-Task Principal/Agent Relations
Author(s): Gerald A. Feltham and Jim Xie
Source: *The Accounting Review*, Vol. 69, No. 3 (Jul., 1994), pp. 429-453
Published by: American Accounting Association
Stable URL: https://www.jstor.org/stable/248233
Accessed: 31-03-2020 19:19 UTC

The Accounting Review
Vol. 69, No. 3
July 1994
pp. 429–453

Performance Measure Congruity and Diversity in Multi-Task Principal/Agent Relations

Gerald A. Feltham
University of British Columbia
Jim Xie
University of Alberta

SYNOPSIS AND INTRODUCTION: Accounting numbers are frequently used in evaluating management performance, and performance evaluation is an important ingredient in motivating managers. Three significant factors generally create difficulties in developing performance measures for a given manager. First, the actions and strategies implemented by the manager are not observable directly, so the manager cannot be compensated directly for his input into the firm. Second, the full consequences of the manager's actions are not observable, in large part because the impact of those actions extend beyond his subunit of the firm and beyond his time as manager of that subunit. Third, uncontrollable events influence the consequences that are observed.

The agency theory literature has explored extensively the implications of the nonobservability of the manager's actions and the fact that performance measures are influenced by unobservable, uncontrollable events. However, this literature has given only limited attention to the fact that performance measures frequently are incomplete or imperfect representations of the economic consequences of the manager's actions.[1]

On the other hand, discussions of performance evaluation in management accounting texts often raise issues regarding the incompleteness and imperfectness of the accounting numbers that are used as performance measures. For example,

[1] These issues are not completely ignored in the agency theory literature. See, for example, the multiperiod models examined by Magee (1978; 1986, ch. 14), Baiman and Noel (1985), Dye (1988), Kanodia et al. (1989), and Bushman and Indjejikian (1993).

Jerry Feltham is thankful for financial support of this research by the Social Sciences and Humanities Research Council of Canada. Jim Xie is thankful for financial support from the Royal Bank Research Fellowship of the University of Alberta. We are grateful for comments on earlier drafts by an associate editor, two anonymous referees, Rick Young (our discussant at the 1993 AAA meetings), Martin Wu, Michael Gibbins, Chu Zheng, and participants in workshops at the University of Alberta, the University of British Columbia (Department of Economics), the University of Minnesota, Northwestern University, and the UBCOW accounting research workshop at the University of Washington.

Submitted March 1993.
Accepted February 1994.

divisional accounting profit is described as a short-term financial measure that may induce managers to ignore the future economic consequences of their current actions.[2] More generally, management accounting texts discuss various problems that arise in inducing managers to have goals that are *congruent* with those of the firm's owners.[3] These discussions typically follow one of two tacks. First, most texts discuss alternative methods for measuring various accounting numbers. For example, discussions of divisional profit measures often consider direct costing versus absorption costing, the elimination of allocated fixed costs, market versus cost based transfer prices, and the inclusion of interest charges for assets used. The objective here is to create a *single* measure that is as congruent with the firm's objectives as possible.

Second, some texts discuss the use of additional, often nonfinancial, performance measures. Kaplan and Atkinson (1989, 536) refer to General Electric and McDonald's as leaders in the use of such measures, and Anthony et al. (1992, 651) provide the following summary of their measures.

> McDonald's evaluated its store managers on product quality, service, cleanliness, sales volume, personnel training, and cost control.
>
> When General Electric decentralized in the 1950s, it identified multiple measures of divisional performance: profitability, market position, productivity, product leadership, personnel development, employee attitudes, and public responsibility.

This paper uses an agency theory model to explore the economic impact of variations in performance measure congruence and the use of multiple performance measures to deal with both problems of goal congruence and the impact of uncontrollable events on performance measures. To address the congruency issues, we use a multidimensional representation of the manager's actions. Most of the agency theory literature has examined models in which the manager's action space is either single dimensional or finite. Our approach is similar to the multi-task model examined by Holmstrom and Milgrom [HM] (1991). Our analysis differs from theirs in that we focus on performance measure issues and consider measures that may be influenced by more than one element of the manager's action.

The key characteristics of a single performance measure are its congruence with the principal's expected gross payoff and its noisiness (due to uncontrollable events). The first-best result is achieved if, and only if, the performance measure is *perfectly congruent and noiseless.* A contract based on a noncongruent measure induces suboptimal effort allocation across tasks, whereas performance measure noise results in suboptimal effort intensity.

We characterize the value of providing additional performance measures, and illustrate the use of additional measures to reduce risk and noncongruity (due to myopia and window dressing). The value of an additional measure is zero if, and only if, the existing measures constitute a sufficient statistic for the additional measure with respect to the manager's action.

[2] See, for example, Magee (1986, 267), Horngren and Foster (1991, 892-894), and Anthony et al. (1992, 649-651).
[3] See Kaplan and Atkinson (1989, 534-540) for a general discussion of this issue.

The terminal value of the firm is not contractible information if it is realized subsequent to the contract termination date. However, the market price (at the contract termination date) of a publicly traded firm is contractible information. The analysis demonstrates that while price efficiently aggregates investor information for valuation purposes, it is not likely to be an efficient aggregation for incentive purposes. Hence, there is a loss of efficiency if the price is used as the sole performance measure. Of course, it can be a valuable performance measure if it contains otherwise noncontractible information.

Key Words: *Multi-task agencies, Multiple performance measures, Congruity, Price based compensation.*

THE remainder of the paper is organized as follows. The basic multi-task principal/agent model is specified in section I, along with a characterization of the first-best contract. The second-best results with a single performance measure are examined in section II. A general characterization of the value of additional performance measures is provided in section III, while special cases of noise reduction and congruity improvement are examined in section IV. The impact of using the market price (at the contract termination date) as a performance measure is examined in section V and concluding remarks are provided in section VI.

I. Basic Model

An agent (a manager) controls n activities that influence the payoff to the principal (representing the owners of the firm employing the agent). These activities can pertain to different divisions, different products, or to management of a variety of activities such as marketing, production, personnel development, and product research. The set of feasible activity levels is represented by an n-dimensional set of non-negative real numbers, denoted $\mathbf{A} \equiv \mathbf{R}^{n+}$, and the agent's action choice is represented by the vector[4] $\boldsymbol{a} = (a_1,...,a_n)^t$.

The principal is risk neutral (e.g., the firm's owners are well diversified and all risks are firm-specific) and his gross payoff (prior to the agent's compensation) is denoted x. The relation between the agent's action and x is assumed to take the form

$$x = B(\boldsymbol{a}) + \varepsilon_x,$$

where $B(\boldsymbol{a})$ is the expected gross payoff and $\varepsilon_x \sim N(0,\sigma_x^2)$ is a random uncontrollable component of the gross payoff. The variance of the uncontrollable component is assumed to be independent of the agent's action choice.

The agent is risk averse and incurs a direct personal cost, denoted $C(\boldsymbol{a})$, if he implements action $\boldsymbol{a}$. His preferences are represented by a negative exponential utility function $u(W) = -e^{-rW}$, where r is the agent's absolute risk aversion and W is *his compensation minus personal cost.*

We restrict $B(\boldsymbol{a})$ and $C(\boldsymbol{a})$ to simple forms that induce closed form interior solutions. In particular, we assume $B(\boldsymbol{a})$ is linear and $C(\boldsymbol{a})$ is separable and quadratic:[5]

$$B(\boldsymbol{a}) = \boldsymbol{b}^t\boldsymbol{a} = b_1a_1 + ... + b_na_n$$

[4] Bold face is used for vectors and matrices, all vectors are column vectors, and "t" denotes transpose.

[5] HM provide a number of interesting insights by considering a variety of benefit functions $B(\boldsymbol{a})$ and personal costs $C(\boldsymbol{a})$. We use a simple linear/quadratic form since this is adequate to provide a number of interesting insights into the impact of alternative performance measures. $B(\boldsymbol{a})$ and $C(\boldsymbol{a})$ could also include fixed components without changing the analysis.

$$C(\boldsymbol{a}) = \frac{1}{2}\boldsymbol{a}^t\boldsymbol{a} = \frac{1}{2}[a_1{}^2 + ... + a_n{}^2],$$

where $\boldsymbol{b} = (b_1,...,b_n)^t$, $b_j \geq 0$, represents the payoff per unit of effort in the two tasks.

If $\boldsymbol{a}$ is publicly observable and contractible, then there is no incentive problem. As is well known, the first-best contract consists of a fixed wage w^* that is sufficient to obtain the agent's acceptance of the contract given that he must provide the first-best action $\boldsymbol{a}^*$, i.e., $w^* = C(\boldsymbol{a}^*) + w^o$, where w^o is the agent's reservation wage. Given the assumed forms for $B(\boldsymbol{a})$ and $C(\boldsymbol{a})$, the first-best action is characterized by the following first-order condition:

$$a_j^* = b_j, \quad j = 1,...,n. \tag{1}$$

That is, effort is expended in each task up to the point at which its marginal cost to the agent (which equals a_j) is equal to its marginal benefit to the principal (which equals b_j). The cost to the agent is relevant to the principal because he must reimburse the agent for that cost. The first-best contract provides an *expected surplus* (the principal's expected gross payoff minus the agent's personal cost) of

$$V^* \equiv B(\boldsymbol{a}^*) - C(\boldsymbol{a}^*) = \frac{1}{2}\boldsymbol{b}^t\boldsymbol{b} = \frac{1}{2}\left[b_1^2 + \ldots + b_{\mathrm{m}}^2\right]. \tag{2}$$

If $\boldsymbol{a}$ is not publicly observable, then the principal faces an incentive problem. We assume that the incentive contract must be based on a vector of publicly reported performance measures, denoted $\boldsymbol{y} = (y_1,...,y_m)^t$, that are produced by performance measurement system $\boldsymbol{\eta}$. As discussed in the introduction, these performance measures can include both accounting and nonaccounting numbers.

We assume that the principal's gross payoff is *not contractible information* (i.e., x is not, in general, an element of $\boldsymbol{y}$), which is consistent with settings in which the agent manages a subunit of the firm or his contract terminates prior to the full realization of the consequences of his actions. In the final section of the paper, we consider a setting in which the firm's market price at the contract termination date is used as a performance measure.[6]

For simplicity and following HM, we assume $\boldsymbol{a}$ has only a mean-shifting effect on these measures. We further assume that there is a linear relation between the agent's actions and the expected levels of the performance measures. Hence,

$$\boldsymbol{y} = \boldsymbol{\mu}\ \boldsymbol{a} + \boldsymbol{\varepsilon},$$

i.e., $$y_i = \mu_{i1}a_1 + ... + \mu_{in}a_n + \varepsilon_i \qquad i = 1,...,m, \tag{3}$$

where $\boldsymbol{\mu} = [\ \mu_{ij}\ ]_{mxn}$ is a matrix of performance measure parameters and $\boldsymbol{\varepsilon} = (\varepsilon_1,...,\varepsilon_m)^t$ is an mx1 vector of normally distributed random variables with mean zero and covariance matrix $\boldsymbol{\Sigma}$.[7] Unless stated otherwise, we assume throughout our analysis that $\boldsymbol{\Sigma}$ is positive definite (p.d.) and,

[6] Bushman and Indjejikian (1993) provide a model in which $\boldsymbol{y} = (y_1,y_2)$, where y_1 is interpreted an accounting number and y_2 is the firm's post-action market price. The market price reflects investors' rational expectations given the accounting number and information they have privately acquired.

[7] Much of HM's analysis focuses on the impact of alternative forms of $C(\boldsymbol{a})$ and they simplify the impact of the performance measures by assuming that each measure is influenced by a single task (although there need not be a measure for each task), i.e., $\mu_{ij} = 0$, for $i \neq j$ and $\mu_{ii} = 1$ if there is a measure for task i (and 0 otherwise). In contrast, we are interested in measures, such as accounting reports, that are influenced by the effort expended in more than one task. Hence, we allow for the possibility that μ_{ij} may be nonzero for any i,j.

HM point out that there is no loss of generality in assuming there is a single performance measure for each task, since the tasks can be defined such that agent chooses the expected performance level for each measure. However, that approach is not very useful when one is examining the impact of alternative performance measures.

hence, Σ^{-1} exists. The vector $\boldsymbol{\varepsilon}$ represents uncontrollable events influencing the performance measures.

We *exogenously* restrict our analysis to contracts in which the wage payment is a *linear* function of the performance measure(s) $\boldsymbol{y}$:[8,9]

$$w(\boldsymbol{y}) = \beta + \boldsymbol{v}^t \boldsymbol{y} = \beta + \sum_{i=1}^{m} v_i \left[\mu_{i1} a_1 + \ldots + \mu_{in} a_n + \varepsilon_i\right], \tag{4}$$

where β is the fixed component of compensation and $\boldsymbol{v}=(v_1, ..., v_m)^t$ is an $m\times1$ vector of incentive compensation coefficients. We can view v_i as the price paid per unit of performance measure i that is produced, whereas β is a fixed payment that is independent of the performance measures.

The key element of the contract is its incentive component $\boldsymbol{v}$, since the fixed component β merely serves to ensure that the compensation contract is accepted. The incentive component is used to induce positive effort by the agent, but in so doing imposes risk. The total variance in the agent's income is $\boldsymbol{v}^t\Sigma\,\boldsymbol{v}$, and he must be paid a risk premium of $\frac{1}{2} r\, \boldsymbol{v}^t\Sigma\,\boldsymbol{v}$ to compensate for that risk. Hence, if the agent is offered incentive contract $\boldsymbol{v}$, which induces action $\boldsymbol{a}$, then the expected cost to the principal is $\frac{1}{2} r\, \boldsymbol{v}^t\Sigma\,\boldsymbol{v} + C(\boldsymbol{a}) + w^o$. The reservation wage is immaterial to the choice of $\boldsymbol{v}$ and $\boldsymbol{a}$. Hence, with respect to the choice $\boldsymbol{v}$ and $\boldsymbol{a}$, the principal's problem, for performance measurement system $\boldsymbol{\eta}$, can be expressed as the maximization of the *expected surplus*[10]

$$V(\boldsymbol{a}, \boldsymbol{v}, \boldsymbol{\eta}) \equiv B(\boldsymbol{a}) - [\tfrac{1}{2} r \boldsymbol{v}^t \textstyle\sum \boldsymbol{v} + C(\boldsymbol{a})], \tag{5}$$

subject to incentive constraint

$$\boldsymbol{a} = \boldsymbol{\mu}^t \boldsymbol{v}, \tag{6}$$

which is the first-order condition for the agent's decision problem.[11]

The solution to the principal's problem is[12]

$$\boldsymbol{v}^\dagger = \boldsymbol{Q\mu b} \tag{7}$$

and

$$\boldsymbol{a}^\dagger = \boldsymbol{\mu}^t \boldsymbol{v} = \boldsymbol{Db}, \tag{8}$$

where

$$\boldsymbol{Q} \equiv \left[\boldsymbol{\mu\mu}^t + r\boldsymbol{\Sigma}\right]^{-1}$$

and

$$\boldsymbol{D} \equiv \boldsymbol{\mu}^t \boldsymbol{Q\mu}.$$

[8] HM also restrict their analysis to linear contracts. They make reference to the dynamic model used by Holmstrom and Milgrom (1987) to specify a setting in which linear contracts are optimal. We do not appeal to their analysis since their setting does not fit "naturally" into a setting, such as ours, in which there are multiple performance measures that are observed by both the principal and the agent. Hence, we make no claim that a linear contract is optimal.

Gjesdal (1981; 1982) provides a general characterization of optimal contracts in a multi-task setting. Matsumura (1988) examines optimal contracts in a setting in which the agent allocates time across tasks (e.g. sales of products). She considers both the simultaneous and sequential choice of effort levels.

[9] A linear contract can be interpreted as consisting of a linear aggregation of the performance measures and a linear incentive contract based on the aggregate performance measure. See Banker and Datar (1989) and Amershi et al. (1990) for analyses of conditions under which linear aggregation of performance measures is optimal in a single-task setting. Generally, even if it is optimal to linearly aggregate performance measures, the optimal contract is not a linear function of the aggregate measure.

[10] See HM for a more detailed development of the argument that the principal's problem is equivalent to maximizing the expected surplus.

[11] The agent's objective is to maximize $\boldsymbol{v}^t \boldsymbol{\mu a}$ - $C(\boldsymbol{a})$. The incentive constraints are the first-order conditions to this problem.

[12] Substituting equation (6) into equation (5) suppresses $\boldsymbol{a}$ and eliminates the incentive constraints, so that the principal's problem can

The resulting second-best expected surplus is

$$V(\boldsymbol{\eta}) = \frac{1}{2}\boldsymbol{b}^t\boldsymbol{D}\boldsymbol{b}. \tag{9}$$

We now examine how the characteristics of $\boldsymbol{\eta} \equiv (\boldsymbol{\mu}, \Sigma)$ influence $\boldsymbol{v}^\dagger$, $\boldsymbol{a}^\dagger$, and $V(\boldsymbol{\eta})$. In this exploration we make comparisons between the second-best and first-best levels of effort and surplus. That is, we compare $\boldsymbol{a}^\dagger$ to $\boldsymbol{a}^*$ and we examine the *loss of surplus* resulting from $\boldsymbol{\eta}$, where the loss in surplus is

$$L(\boldsymbol{\eta}) \equiv V^* - V(\boldsymbol{\eta}) = \frac{1}{2}\boldsymbol{b}^t[\boldsymbol{I} - \boldsymbol{D}]\boldsymbol{b}. \tag{10}$$

II. Single Performance Measure Congruity and Precision

Many incentive contracts are based on a single performance measure ($m = 1$), even if the agent's action is multidimensional. This section demonstrates that the loss in surplus in this setting can be viewed as a function of the performance measure's *congruity* and *precision*. Congruity refers to the degree of congruence between the impact of the agent's action on his performance measure and on the principal's expected gross payoff, while precision refers to the noise in the performance measure. A simple example of differences in congruity is the evaluation of a salesman, who sells several products, on the basis of total sales dollars versus total contribution margin. We generally view maximization of the total contribution margin to be more congruent with the firm's objectives than maximization of total sales, since the former recognizes the incremental costs of the sales whereas the latter does not. Similarly, residual income is generally viewed as more congruent with the firm's objectives than divisional profit since the former considers the cost of the capital tied up in divisional assets. Of course, all of the above measures are inherently noisy in that they are influenced by uncertain, uncontrollable events.

In the two-task ($n = 2$) setting, performance measurement system η is characterized by $\boldsymbol{\mu} = [\,\mu_{11}\ \mu_{12}\,]$ and $\Sigma = \sigma_1^2$. To avoid degeneracy we assume $b_1,\mu_{11} > 0$ and $b_2,\mu_{12} \geq 0$. From first order conditions (7) and (8) we obtain

$$v^\dagger = \frac{b_1\mu_{11} + b_2\mu_{12}}{\mu_{11}^2 + \mu_{12}^2 + r\sigma_1^2}$$

and

$$a_j^\dagger = v^\dagger\mu_{1j}, \quad j = 1,2.$$

Observe that the denominator of the incentive function's slope parameter, $v^\dagger$, is an increasing function of $r\sigma_1^2$, which implies that $v^\dagger$ and $a_j^\dagger$, all j, are decreasing functions of $r\sigma_1^2$. For both tasks,

---

be expressed as:

$$V(\eta) = \underset{\boldsymbol{v}}{\text{maximize}}\ \boldsymbol{b}^t\boldsymbol{\mu}^t\boldsymbol{v} - \tfrac{1}{2}\left[r\boldsymbol{v}^t\Sigma\,\boldsymbol{v} + \boldsymbol{v}^t\boldsymbol{\mu}\boldsymbol{\mu}^t\boldsymbol{v}\right].$$

Now the incentive contract is the sole object of choice and the induced action is implicit in the analysis. Differentiating $V(\boldsymbol{\eta})$ with respect to $\boldsymbol{v}$ provides the first-order condition (7). Substituting (7) into (6) provides (8).

This analysis assumes $\boldsymbol{Q}$ exists, which is true as long as there is no performance measure y_i such that either $\mu_{ij} = r\sigma_{ik} = 0$ for all $j=1,\dots,n$ and $k = 1,\dots,m$, or $\mu_{ij} = \sum_{\ell\neq i}\xi_\ell\mu_{\ell j}$ and $r\sigma_{ik} = r\sum_{\ell\neq i}\xi_\ell\sigma_{i\ell}$ for some $\{\xi_\ell\}_{\ell\neq i}$. Any such measure can be eliminated without consequence, since the measure is either a constant or redundant. The existence of $\boldsymbol{Q}$ follows from the fact that $\boldsymbol{\mu}\boldsymbol{\mu}^t$ and $r\Sigma$ are always positive semidefinite, and their sum is positive definite as long as the above conditions do not hold (the inverse of a positive definite matrix always exists).

an increase in risk aversion or noise reduces the intensity of the effort induced by the second-best contract.

Further observe that the relative effort expended in the two tasks is equal to their relative impact on the performance measure (irrespective of r and σ_1^2):

$$\frac{a_2^\dagger}{a_1^\dagger} = \frac{\mu_{12}}{\mu_{11}}.$$

In the first-best action, on the other hand, the relative effort expended in the two tasks is proportional to their relative contributions to the principal's expected gross payoff:

$$\frac{a_2^*}{a_1^*} = \frac{b_2}{b_1}.$$

The following specifies a measure of noncongruity between a performance measure and the principal's expected gross payoff. We refer to a performance measure as being perfectly congruent with the principal's gross payoff if there is zero noncongruity.

Definition: The *measure of noncongruity* for performance measure y_i, relative to the principal's gross payoff, is[13]

$$\delta_i \equiv [b_1\mu_{i2} - b_2\mu_{i1}]^2.$$

Performance measure y_i is *perfectly congruent* with $\boldsymbol{b}$ if its measure of noncongruity is zero.

Observe that y_i is perfectly congruent if, and only if, there exists a constant $\lambda_i \neq 0$ such that $\mu_{ij} = \lambda_i b_j$, j = 1,2.

The impact of noncongruity and noise on the principal's expected net payoff is reflected in the following specification of the loss of surplus (which follows from equation (10)):

$$\mathrm{L}(\boldsymbol{\eta}) = \frac{\delta_1 + \frac{1}{2} r\sigma_1^2 (b_1^2 + b_2^2)}{\mu_{11}^2 + \mu_{12}^2 + r\sigma_1^2}. \tag{11}$$

The first term, δ_1, reflects the loss due to noncongruity, and the second term, $\frac{1}{2} r\sigma_1^2 (b_1^2 + b_2^2)$, reflects the loss due to performance measure noise.

From equation (11) it is obvious that a single performance measure can only achieve the first-best result if it is both perfectly congruent *and* noiseless (or the agent is risk neutral).

Proposition 1: If there is a single performance measure ($m = 1$), then $\mathrm{L}(\boldsymbol{\eta}) = 0$ if, and only if, $r\sigma_1^2 = 0$ and there exists a constant λ such that $\mu_{1j} = \lambda b_j$, $j = 1,2$.

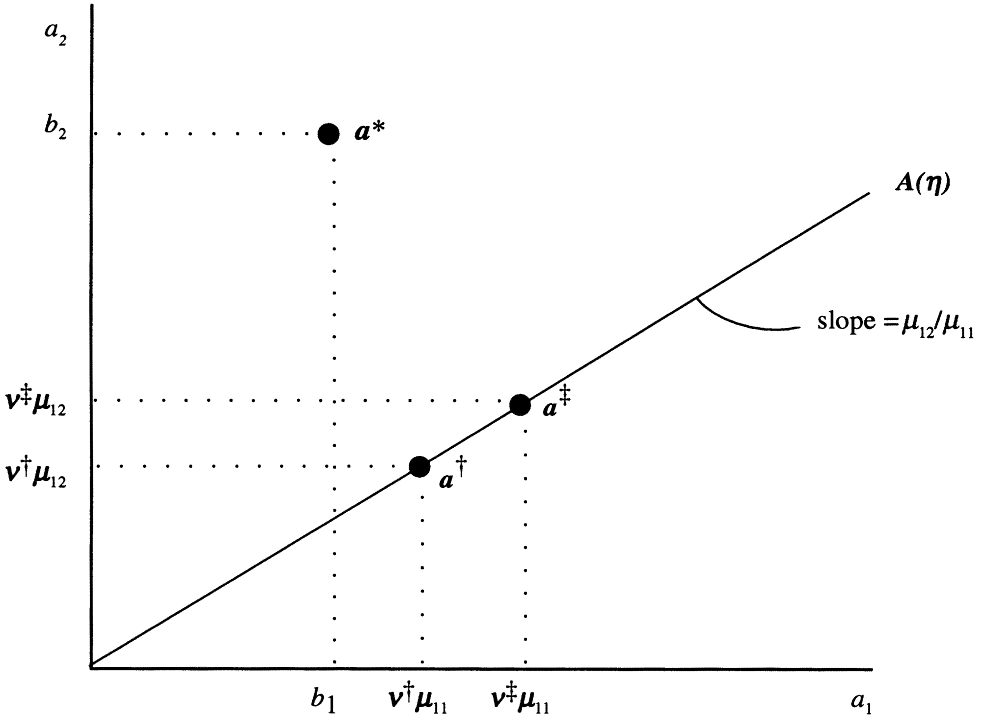
Risk neutrality *or* a noiseless performance measure have been sufficient conditions for achievement of the first-best result in most models examined in the principal/agent literature, but those models typically assume that the action space is either single dimensional or finite. Furthermore, many models in the literature presume that $y_1 = x$, which immediately implies that the performance measure is perfectly congruent. The key point here is that if the gross payoff to the principal is not contractible information, then risk neutrality or a noiseless performance measure are not sufficient to achieve the first-best result. The performance measure must also be perfectly congruent.

[13] For $n > 2$, the measure of noncongruity for y_i is

$$\delta_i = \sum_{j=1}^{n-1} \sum_{k=j+1}^{n} \left[b_j\mu_{ik} - b_k\mu_{ij}\right]^2.$$

Figure 1

Impact of Noncongruity and Imprecision of a Single Performance Measure



$$a_j^{\ddagger} = v^{\ddagger}\mu_{1j} \qquad v^{\ddagger} = [\mu_{11}b_1 + \mu_{12}b_2] \div [\mu_{11}^2 + \mu_{12}^2]$$

$$a_j^{\dagger} = v^{\ddagger}\mu_{1j} \qquad v^{\dagger} = [\mu_{11}b_1 + \mu_{12}b_2] \div [\mu_{11}^2 + \mu_{12}^2 + r\sigma_1^2]$$

$$\boldsymbol{\mu a}^{\dagger} < \boldsymbol{\mu a}^{\ddagger} = \boldsymbol{\mu a}^{*}$$

Figure 1 depicts the impact of a noncongruent, imprecise performance measure. For purposes of the illustration, assume the agent is a salesman and the two dimensions refer to the efforts for two products. The first-best action is $\boldsymbol{a}^* = \boldsymbol{b}$, where b_j is the expected contribution margin per unit of effort in selling product j. Assume the incremental costs are difficult to measure and the firm uses sales dollars as the performance measure. The line denoted $\boldsymbol{A(\eta)}$ represents the set of actions that can be implemented with performance measurement system $\boldsymbol{\eta}$; it is the set of effort levels that minimize the agent's personal cost of providing each possible performance level (sales total) and it has a slope of μ_{12}/μ_{11} (the relative number of sales dollars per unit of sales effort for the two products). The total sales that would result if the agent selected the first-best effort is $y_1^* = \mu_{11}b_1 + \mu_{12}b_2$. Action $\boldsymbol{a}^{\ddagger}$ represents the least costly action to the agent that will produce sales level y_1^*. This is the optimal second-best action if the performance measure is noiseless or the agent is risk neutral (i.e., $r\sigma_1^2 = 0$). If sales is a noisy measure (which is almost certainly the case) and the agent is risk averse, then the optimal second-best action, denoted $\boldsymbol{a}^{\dagger}$, is less than but proportional to $\boldsymbol{a}^{\ddagger}$. Hence, noncongruity of the performance measure causes a shift in the relative amount of effort put into selling the two products (in this case, relatively more effort is put into selling product one than is first-best) and the imprecision of the sales measure (with risk aversion) results in weaker incentives than those that would induce the first-best performance level.

Observe that noise in a performance measure does not preclude inducement of the first-best action, it only makes it costly (because of the risk premium that must be paid). Noncongruity, on the other hand, makes it impossible to induce the first-best action. One can view much of the discussion of the trade-off between accounting and market based performance measures as a discussion of the relative merits of congruity and precision. Market based measures may be more congruent but less precise (i.e., contain more noise) than accounting measures because the former reflect information about the future consequences of current actions but are influenced by more uncontrollable events.

III. The Value of Additional Performance Measures

Table 1 illustrates the value of adding performance measures. In this section we provide a general statement of the value of additional performance measures relative to an existing set. The subsequent section considers three special cases.

Let $\boldsymbol{\eta} = (\boldsymbol{\mu}_1, \Sigma_{11})$ represent a performance measurement system that reports m_1 measures and let $\boldsymbol{\eta}_+ = (\boldsymbol{\mu}_+, \Sigma_+)$ represent a system that reports $m_+ = m_1 + m_2$ measures, where

$$\boldsymbol{\mu}_+ = \begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix}$$

and

$$\Sigma_+ = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}.$$

The following proposition provides a formal specification of $\pi(\boldsymbol{\eta}_+, \boldsymbol{\eta})$, the incremental expected value of the additional performance measures provided by $\boldsymbol{\eta}_+$.

Proposition 2: If $\boldsymbol{Q}_+ \equiv [\boldsymbol{\mu}_+\boldsymbol{\mu}_+^t + r\Sigma_+]^{-1}$ exists, then

$$\pi(\boldsymbol{\eta}_+,\boldsymbol{\eta}) \equiv \mathrm{V}(\boldsymbol{\eta}_+) - \mathrm{V}(\boldsymbol{\eta}) = \tfrac{1}{2}\boldsymbol{T}^t\boldsymbol{P}\boldsymbol{T} \geq 0. \tag{12}$$

where $\boldsymbol{P} \equiv [\boldsymbol{H}_{22} - \boldsymbol{H}_{21}\boldsymbol{Q}\boldsymbol{H}_{12}]^{-1}$ is a positive definite matrix,

$$\boldsymbol{T} \equiv (\boldsymbol{\mu}_2 - \boldsymbol{H}_{21}\boldsymbol{Q}\boldsymbol{\mu}_1)\boldsymbol{b}^t,$$

$$\boldsymbol{Q} = [\boldsymbol{\mu}_1\boldsymbol{\mu}_1^t + r\Sigma_{11}]^{-1},$$

and $\boldsymbol{H}_{ij} \equiv \boldsymbol{\mu}_i\boldsymbol{\mu}_j^t + r\boldsymbol{\Sigma}_{ij}, \quad i,j = 1,2.$

Non-negativity of the incremental value of additional performance measures follows directly from the fact that the principal can always assign zero incentive to the additional measures. The positive definiteness of $\boldsymbol{P}$ and expression (12) imply that the additional measures provided by $\boldsymbol{\eta}_+$ have zero value if, and only if, $\boldsymbol{T} = 0$.

Table 1
Multiple Performance Measures Example

payoff parameters: $\boldsymbol{b} = (2, 4)^t$ | $\boldsymbol{a}^* = (2, 4)^t$ | $V^* = 10$ | agent's risk aversion: $r = 1$

performance measurement system η	least cost incentive contract to implement the first-best action	second-best incentive contract $\boldsymbol{v}^\dagger$	second-best action $\boldsymbol{a}^\dagger$	second-best surplus $\mathbf{V}(\eta)$
$m=1$: $\boldsymbol{\mu}=[5\ 3]$, $\boldsymbol{\Sigma}=[5]$	not implementable	$v=.5$	$a_1=2.5$ $a_2=1.5$	6.205
$m=2$: $\boldsymbol{\mu}=\begin{bmatrix}5 & 3\\ 3 & 3\end{bmatrix}$, $\boldsymbol{\Sigma}=\begin{bmatrix}5 & 0\\ 0 & 25\end{bmatrix}$	$v_1=-1$ $v_2=2.333$	$v_1=.467$ $v_2=.158$	$a_1=2.808$ $a_2=1.875$	6.557
$m=3$: $\boldsymbol{\mu}=\begin{bmatrix}5 & 3\\ 3 & 3\\ 2 & 4\end{bmatrix}$, $\boldsymbol{\Sigma}=\begin{bmatrix}5 & 0 & 0\\ 0 & 25 & 0\\ 0 & 0 & 900\end{bmatrix}$	$v_1=-.864$ $v_2=2.017$ $v_3=.136$	$v_1=.463$ $v_2=.157$ $v_3=.008$	$a_1=2.802$ $a_2=1.891$	6.584

A given set of performance measures may not be sufficient to implement all possible actions. Hence, increasing the number of performance measures may increase the set of implementable actions, and thereby may result in the implementation of a more preferred action. In addition, increasing the number of performance measures may reduce the risk that must be imposed to induce a particular implementable action. Expression (12) specifies the value of additional measures due to both these possible sources. (Appendix B considers a given set of performance measures and identifies the set of implementable actions and the minimum cost contract for implementing each action.)

To explore the conditions under which additional performance measures have *zero* value, we consider the conditions under which measures from $\boldsymbol{\eta}$ are a sufficient statistic for the measures from $\boldsymbol{\eta}_+$, with respect to the actions $\boldsymbol{a} \in \boldsymbol{A}$.[14]

Lemma 1: The measures provided by performance measurement system η are a *sufficient statistic* for the measures provided by $\boldsymbol{\eta}^+$, with respect to the actions $\boldsymbol{a} \in \boldsymbol{A}$, if there exists an $m_2 \times m_1$ constant matrix Ξ such that

$$\boldsymbol{\mu}_2 = \Xi\boldsymbol{\mu}_1 \text{ and } \Sigma_{21} = \Xi\Sigma_{11}. \tag{13}$$

If condition (13) holds, then $\boldsymbol{T} = 0$ and proposition 2 directly implies that these conditions are sufficient for the additional performance measures to have zero value.[15]

Proposition 3:[16] If $\boldsymbol{Q}_+$ exists, then $\pi(\boldsymbol{\eta}_+, \boldsymbol{\eta}) = 0$ if, and only if, there exists an $m_2 \times m_1$ constant matrix Ξ such that condition (13) holds.

The above result is an extension of a well-known result in the agency theory literature (e.g., see Holmstrom (1979)). The key point here is that in a multi-task setting consideration must be given to both the impact of the action on the expected performance measures and the measures' statistical properties. This is reflected in the fact that condition (13) has two parts. The first requirement (the existence of Ξ such that $\boldsymbol{\mu}_2 = \Xi\boldsymbol{\mu}_1$) implies that *the additional measures do not expand the set of implementable actions*. The second requirement ($\Sigma_{21} = \Xi\Sigma_{11}$) focuses on the statistical (covariance) properties of the measure. In a single dimensional setting, the latter is all that is required (if the *direction* of the impact of the agent's action on the mean of each measure is appropriate).

Intuitively, condition (13) implies that the second set of performance measures can be viewed as noisy linear functions of the first measures, i.e., if $y_1 = \boldsymbol{\mu}_1\boldsymbol{a} + \varepsilon_1$, $y_2 = \boldsymbol{\mu}_2\boldsymbol{a} + \varepsilon_2$, and condition (13) holds, then $y_2 = \Xi y_1 + noise$, where ε_1 and *noise* are independently distributed. Hence, the second set of measures are redundant and would create unnecessary additional risk if used.

To illustrate the preceding discussion, again consider the example in table 1, but now assume that the covariance matrix for the three performance measure case is

[14] See Amershi (1988) for a general discussion of sufficient statistics as they relate to information economic and principal-agent analyses.

[15] If expression (13) holds, then $\boldsymbol{T} = (\boldsymbol{\mu}_2 - \boldsymbol{H}_{21}\boldsymbol{Q}\boldsymbol{\mu}_1)\boldsymbol{b} = \Xi(\boldsymbol{\mu}_1 - [\boldsymbol{\mu}_1\boldsymbol{\mu}_1^t + r\Sigma_{11}]\boldsymbol{Q}\boldsymbol{\mu}_1)\boldsymbol{b} = 0$, since $[\boldsymbol{\mu}_1\boldsymbol{\mu}_1^t + r\Sigma_{11}]\boldsymbol{Q} = \boldsymbol{I}$.

[16] This result is similar to the classical value of additional information result in Holmstrom's (1979) analysis of a single-task setting. He assumes $\boldsymbol{\eta}$ reports $y_1 = x$ and $\boldsymbol{\eta}_+$ reports $\boldsymbol{y}_+ = (x, y)^t$. He establishes that if condition $f(y_1,y_2 \mid \boldsymbol{a}) = g(y_1,y_2)\, h(y_1 \mid \boldsymbol{a})$ for either all $\boldsymbol{a}$ or no $\boldsymbol{a}$, then y has incremental value if, and only if, this condition is false. Amershi and Hughes (1989) and Amershi et al. (1990) demonstrate that one must take care in asserting that additional information is valuable if the existing information is not a sufficient statistic for the two sets of signals. The concerns they raise do not arise in our setting because of our use of normal distributions.

$$\boldsymbol{\Sigma} = \begin{bmatrix} 5 & 0 & -5 \\ 0 & 25 & \frac{175}{3} \\ -5 & \frac{175}{3} & 900 \end{bmatrix}$$

The third measure is now redundant, since condition (13) holds with $\Xi = [\ \text{-}1\ 7/3\ ]$. That is, $y_3 = \text{-}y_1 + (7/3)y_2 + noise$, where $noise \sim N(0, 1460/3)$. In table 1, the third measure has incremental value because ε_3 is independent of ε_1 and ε_2, and therefore y_3 can be used to reduce risk even though it is very noisy. However, with the above correlation structure, the optimal contract is $v^\dagger = (.467, .158, 0)^t$; y_3 cannot be used to reduce risk and hence it is ignored (even though it is perfectly congruent with $\boldsymbol{b}$).

IV. Special Cases of Risk Reduction and Congruity Improvement

Additional performance measures have incremental value either because they reduce the risks imposed on the agent or because they provide diverse performance measures that can be used to induce actions that are more congruent with the principal's gross payoff. We now consider some special settings that provide further insight into the two different roles of additional performance measures.

In each of the settings examined below, we consider the impact of increasing the number of performance measures from one (system $\boldsymbol{\eta}$) to two (system $\boldsymbol{\eta}_+$) when there are two-tasks ($n = 2$) and the agent is risk averse ($r > 0$). The two systems are characterized as follows:

$$\boldsymbol{\eta} = (\boldsymbol{\mu}, \Sigma) \qquad \boldsymbol{\mu} = [\ \mu_{11}\ \mu_{12}\ ] \qquad \Sigma = \sigma_1^2$$

$$\boldsymbol{\eta}_+ = (\boldsymbol{\mu}_+, \Sigma_+) \qquad \boldsymbol{\mu}_+ = \begin{bmatrix} \mu_{11} & \mu_{12} \\ \mu_{21} & \mu_{22} \end{bmatrix} \qquad \Sigma_+ = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$$

Information about Uncontrollable Events

Risk is created by uncontrollable events that influence performance measures. We consider a setting in which the first measure is influenced by the agent's actions, but the second is not (i.e., $\mu_{1j} > 0$ and $\mu_{2j} = 0$, for $j = 1,2$). The second measure can be represented as $y_2 = \varepsilon_2$. It does not change the set of implementable actions and its key characteristic is its correlation with the primary measure. It has zero value if the two measures are uncorrelated ($\rho = 0$), but it has *positive value if they have some correlation* ($\rho \neq 0$). The stronger the correlation, the stronger are the incentives applied to both measures. If the two measures are perfectly correlated ($\rho = \pm 1$), then the first-best result can be achieved *if* the primary measure is perfectly congruent with the principal's gross payoff.

The following proposition formalizes the above discussion. Let $v^\dagger$ and $\boldsymbol{a}^\dagger$ represent the optimal contract and action based on $\boldsymbol{\eta}$ and let $v^\ddagger$ and $\boldsymbol{a}^\ddagger$ represent the optimal contract and action based on $\boldsymbol{\eta}_+$.

Proposition 4: Assume $\boldsymbol{\eta}$ provides a single performance measure y_1 with $\mu_{1j} > 0$, all j, and $r\sigma_1^2 > 0$. System $\boldsymbol{\eta}_+$ provides both y_1 and y_2, with $\mu_{2j} = 0$, all j.

(a) $\pi(\eta_+, \eta) > 0$ if, and only if, $\rho \neq 0$.

(b) $L(\eta_+) = 0$ if, and only if, $\rho^2 = 1$ and $\mu_{1j} = \lambda b_j$ for some $\lambda > 0$.

(c) $$v_1^\ddagger = \frac{b_1\mu_{11} + b_2\mu_{21}}{\mu_{11}^2 + \mu_{12}^2 + r(1-\rho^2)\sigma_1^2} \text{ and } v_2^\ddagger = -\rho\frac{\sigma_1}{\sigma_2}v_1^\ddagger.$$

There are two commonly used sources of information about uncontrollable events: investigation of the factors influencing realized performance levels and the performance levels of other divisions or competitors. The use of the first type of information involves implementation of what accountants call the *controllability principle*. Antle and Demski (1988) provide a principal/agent analysis of this principle and Merchant (1987) provides some survey data on the extent of its use in practice. A basic point made by Antle and Demski is that it is not always optimal to merely eliminate the uncontrollable events from the performance measure, since the identified events may provide information about the unidentified events. This can be illustrated by assuming *uncontrollable events* (ε_1) = *identified events* (ε_2) + *unidentified events*. If the identified and unidentified events are independent, then $\rho = \sigma_2/\sigma_1$ and $v_2^\ddagger = -v_1^\ddagger$, i.e., the controllability principle is implemented and the contract can be viewed as paying an incentive wage $v_1^\ddagger$ based on the *net performance measure* $y^\ddagger \equiv y_1 - y_2$. However, if the identified and unidentified events are positively correlated, then $\rho > \sigma_2/\sigma_1$ and $-v_2^\ddagger > v_1^\ddagger$, i.e., a greater adjustment is made than merely eliminating the identified uncontrollable events.

The use of information from other divisions or competitors is frequently manifested in the use of *relative performance measures*. Holmstrom (1982) provides a principal/agent analysis of relative performance measures, Antle and Smith (1986) provide empirical evidence on its use based on statistical analyses of publicly reported compensation and performance data for a large number of firms, and Maher (1987) provides some survey data about the extent of its use. In our analysis, y_2 can be interpreted as the *difference between the actual and expected performance of another firm or set of firms*. If there is a positive correlation between y_1 and y_2, then the second-best contract places a positive incentive on the agent's performance ($v_1^\ddagger > 0$) and subtracts an "insurance" adjustment based on the difference between the competitor's actual and expected performance ($v_2^\ddagger < 0$). An alternative interpretation is to view the contract as specifying a net performance measure $y^\ddagger \equiv y_1 - (\rho\sigma_1/\sigma_2)y_2$, and an incentive wage of $v_1^\ddagger y^\ddagger$. Observe that the adjustment to y_1 based on y_2 depends on only the correlation between the two measures and the relative size of the noise in those two measures (as represented by the standard deviation of that noise). Interestingly, that *adjustment is independent of the congruence* of y_1 with respect to the principal's gross payoff. Of course, the congruence of y_1 does influence $v_1^\ddagger$.

Myopic Performance Measures

The use of accounting earnings as the sole performance measure is often criticized because it does not fully reflect the economic consequences of the agent's actions. This myopia can arise either because the accounting measure is short-run in its orientation, ignoring the future benefits of such activities as product and personnel development, or because it focuses on a single division, ignoring externalities with respect to other divisions. While the earnings number can be modified to make it more congruent,[17] such modifications are often limited by the other uses of the earnings number. Hence, the primary means for dealing with myopia have been either to supplement the

[17] Magee (1978) examines a two-task setting in which the one task influences current operating profit and the other influences holding gains and losses that are realized in the future. His comparison of historical cost versus current cost income measures can be interpreted as an analysis of a single performance measure that is varied in its congruity and noise.

earnings measure with other measures that focus on future effects or externalities, or to use market price based incentives (if the firm's ownership is publicly traded). The first approach is illustrated by General Electric's use of measures of product leadership and personnel development, and by McDonald's use of quality, service and cleanliness (QSC) measures.[18] Bushman and Indjejikian (1993) provide a principal/agent analysis of the second approach, assuming that the market price reflects both accounting information that may be myopic and private investor information that is perfectly congruous with the principal's gross payoff. We examine the use of market price as a performance measure in the next section.

A myopic performance measure is represented by a setting in which $b_1 = \mu_{11} > 0$ and $b_2 > \mu_{12} = 0$. The loss in surplus from using only y_1 as a performance measure is

$$L(\eta) = \frac{1}{2}\left[\frac{r\sigma_1^2}{b_1^2 + r\sigma_1^2} b_1^2 + b_2^2\right].$$

The loss results from the costly risk imposed in motivating the first task and the failure to induce any effort in the second task. The second component of that loss can be reduced by introducing a second performance measure that independently reports on the second task,[19] i.e., $\mu_{21} = 0$, $\mu_{22} = b_2$, and $\sigma_{12} = 0$. The incremental value of the second measure is

$$\pi(\eta_+, \eta) = \frac{1}{2}\left[1 - \frac{r\sigma_2^2}{b_2^2 + r\sigma_2^2}\right] b_2^2$$

and the optimal incentive contract is $v_j^{\ddagger} = b_j^2 / [b_j^2 + r\sigma_j^2]$, $j = 1,2$.

The value of the second measure increases as its precision ($1/\sigma_2^2$) increases, since increased precision reduces the risk premium that must be paid and makes it optimal to use stronger incentives. If both measures are noiseless, then $v_1^{\ddagger} = v_2^{\ddagger} = 1$ and the first-best result is obtained.

Window Dressing

Performance measures, including accounting earnings, are often subject to *manipulation* in the sense that the agent can take actions that improve his performance measure but contribute little or nothing to the principal's gross payoff. We refer to this as *window dressing*, and represent it by $\mu_{11} = b_1 > 0$ and $\mu_{12} > b_2 = 0$. The loss in surplus from using only y_1 as a performance measure is

$$L(\eta) = \frac{1}{2} b_1^2 \frac{\mu_{12}^2 + r\sigma_1^2}{b_1^2 + \mu_{12}^2 + r\sigma_1^2}.$$

The loss is due to the costly risk imposed to induce positive effort in the productive act (a_1) plus the cost created by the nonproductive act (a_2). With a single performance measure, the two acts must be induced in fixed proportions: $a_1/a_2 = \mu_{11}/\mu_{12}$. Consequently, any inducement of a_1

[18] McDonald's use of the QSC measure reflects the belief that the quality, service, and cleanliness of a given store has future payoff consequences both for that store and for McDonald's stores that might be frequented by the same customers.

[19] We now have one performance measure for each task, which permits implementation of all possible actions. Holmstrom and Milgrom (1991) assume this to be the case in much of their analysis, and Paul (1992) assumes that this type of structure underlies the information impounded in the market price.

necessarily induces a_2. This has two negative effects. First, to obtain contract acceptance the agent must be compensated for his personal cost of the nonproductive act. Second, because of this cost, weaker incentives are used than would be the case if $\mu_{12} = 0$, thereby inducing less productive effort.

The loss of surplus due to the nonproductive act can be reduced by introducing a second performance measure that independently reports on either the nonproductive or the productive act. If $\mu_{21} = 0$, $\mu_{22} > 0$, and $\sigma_{12} = 0$, then the incremental value of the second measure is

$$\pi(\eta_+, \eta) = \frac{1}{2} b_1^2 \frac{\mu_{12}^2 b_1^2 Q^2}{1 + r(\sigma_2 / \mu_{22})^2 - \mu_{12}^2 Q},$$

where

$$Q = \left[b_1^2 + \mu_{12}^2 + r\sigma_1^2\right]^{-1}.$$

The incremental value of y_2 increases with μ_{12} because that increases the noncongruity of the first measure and it increases with μ_{22}/σ_2 because that increases the second measure's relative precision about the nonproductive act. The optimal contract imposes a penalty on y_2, which permits the use of stronger incentives on y_1, thereby inducing more productive effort and less nonproductive effort. While it is feasible to eliminate all effort in the nonproductive act, it will not be optimal to do so unless the second measure is noiseless. If both measures are noiseless, then $v_1^\ddagger = 1$, $v_2^\ddagger = -\mu_{12}/\mu_{22}$, and the first-best result is obtained.

If $\mu_{21} > 0$, $\mu_{22} = 0$, and $\sigma_{12} = 0$, then the incremental value of the second measure is

$$\pi(\eta_+, \eta) = \frac{1}{2} b_1^2 \frac{\left[1 - b_1^2 Q\right]^2}{1 + r(\sigma_2 / \mu_{21})^2 - b_1^2 Q}.$$

Again, the incremental value of y_2 increases with μ_{12} because that increases the noncongruity of the first measure, and it increases with μ_{21}/σ_2 because that increases the second measure's relative precision about the productive act. Observe that the second measure is perfectly congruent with the principal's expected payoff. Hence, if it is also noiseless, the first measure is ignored ($v_1^\ddagger = 0$) and the second measure is used ($v_2^\ddagger = b_1/\mu_{21}$) to obtain the first-best result. However, if the second measure is noisy, then it is optimal to use the first measure to reduce the risk that is imposed on the agent, but that risk reduction comes at the expense of compensating the agent for the nonproductive act.

V. Market Price as a Performance Measure

In a publicly traded firm, the market price of the firm's equity at the contract termination date represents the market's expectations with respect to the equityholders' (principal) final net payoff. We now consider the impact of using that market price (adjusted for the manager's compensation) as a performance measure.

A crucial point to recognize is that the market price will reflect both *the information received by investors and their equilibrium beliefs about the agent's action*. The investors' beliefs about the agent's action are based on the incentives under which he is operating, and are not influenced by the information they receive. The investors' information will be impounded in the price to the extent that it influences their beliefs about ε_x (the uncontrollable component of x).

To illustrate these points, assume that $\boldsymbol{y}$ now represents the information received by investors prior to the contract termination date. Recall that the principal's gross payoff is $x = \boldsymbol{b}^t\boldsymbol{a} + \varepsilon_x$, where $\boldsymbol{a}$ is the action selected by the agent and ε_x represents uncontrollable events that influence the principal's payoff (with zero mean and variance σ_x^2). Let $\Sigma_{xy} = [\ \sigma_{xi}\ ]_{1\times m}$ represent the covariance between x and $\boldsymbol{y}$. If the *investors believe the agent has selected action* $\hat{\boldsymbol{a}}$, then their prior belief about x and $\boldsymbol{y}$ is

$$\begin{bmatrix} x \\ \boldsymbol{y} \end{bmatrix} \sim \mathrm{N}\left(\begin{bmatrix} \boldsymbol{b}^t\hat{\boldsymbol{a}} \\ \boldsymbol{\mu}\hat{\boldsymbol{a}} \end{bmatrix}, \begin{bmatrix} \sigma_x^2 & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma \end{bmatrix}\right).$$

The (gross) market price at the contract termination date is equal to the investors' expectation about x conditional on $\boldsymbol{y}$ and $\hat{\boldsymbol{a}}$, which is[20]

$$p = \mathrm{E}[x \mid \boldsymbol{y}, \hat{\boldsymbol{a}}] = \boldsymbol{b}^t\hat{\boldsymbol{a}} + \Sigma_{xy}\Sigma^{-1}[\boldsymbol{y} - \boldsymbol{\mu}\hat{\boldsymbol{a}}]. \tag{14}$$

From the agent's perspective, the distribution of $\boldsymbol{y}$ depends on the action he selects, i.e., $\boldsymbol{y} \sim \mathrm{N}(\mu\boldsymbol{a},\Sigma)$. Consequently, if the market price is used as a performance measure, the agent views that measure as taking the following form:

$$p = [\boldsymbol{b}^t - \Sigma_{xy}\Sigma^{-1}\boldsymbol{\mu}]\ \hat{\boldsymbol{a}} + \Sigma_{xy}\Sigma^{-1}\boldsymbol{\mu}\boldsymbol{a} + \varepsilon_p, \tag{15}$$

where $\varepsilon_p = \Sigma_{xy}\Sigma^{-1}\boldsymbol{\varepsilon} \sim \mathrm{N}(0,\sigma_p^2)$,

$$\sigma_p^2 \equiv \mathrm{Var}[p] = \Sigma_{xy}\Sigma^{-1}\Sigma_{yx}.$$

That is, as a performance measure, the price consists of a fixed component that depends on the investors' belief about the agent's action, a component that varies with the agent's action (because of its impact on $\boldsymbol{y}$), and a random component (due to the uncontrollable events influencing $\boldsymbol{y}$). In equilibrium, the agent's action choice equals the investors' belief $\hat{\boldsymbol{a}}$. However, the investors cannot observe $\boldsymbol{a}$ and, hence, $\boldsymbol{a}$ influences p only through its impact on $\boldsymbol{y}$.

We now consider the impact of using the market price as the sole performance measure, instead of using the underlying information. The *market price* is effectively based on the following *aggregate representation of the investors' information* (see equation (14)):

$$y_x = \Sigma_{xy}\Sigma^{-1}\boldsymbol{y}.$$

On the other hand, if $\boldsymbol{y}$ is contractible information, then the second-best incentive contract is effectively based on the following *aggregate representation of the agent's performance measures* (see equation (7)):

$$y_a = [\boldsymbol{Q}\boldsymbol{\mu}\boldsymbol{b}]^t\boldsymbol{y}$$

where

$$\boldsymbol{Q} = [\boldsymbol{\mu}\boldsymbol{\mu}^t + r\Sigma]^{-1}.$$

The market price can be used as the sole performance measure, without any loss of surplus, *if the signal weights in y_x are proportional to those in y_a.*

[20] See Raiffa and Schlaifer (1961) for derivation of the conditional mean from a joint normal distribution.

Proposition 5: Assume $\boldsymbol{y}$ is publicly reported, $r > 0$, and Σ is positive definite. There is no loss in surplus from using the market price as the sole performance measure, instead of using $\boldsymbol{y}$, if, and only if,

$$\Sigma_{xy}\Sigma^{-1} = \lambda[\boldsymbol{Q\mu b}]^t \qquad \text{for some } \lambda \neq 0. \tag{16}$$

To provide insight into condition (16), we consider some simple two-task settings with one and two public signals. With one signal

$$\Sigma = \sigma_1^2 \qquad \Sigma_{xy} = \sigma_{x1},$$

and with two signals

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} \qquad \Sigma_{xy} = \begin{bmatrix} \sigma_{x1} & \sigma_{x2} \end{bmatrix}.$$

Corollary: Assume $\boldsymbol{y}$ is publicly reported and $r\sigma_i > 0$, all i.

(a) Suppose $m = 1$ and $\mu_{1j} > 0$, all j. There is no loss in surplus from using p as the sole performance measure if, and only if, $\sigma_{x1} \neq 0$.

(b) Suppose $m = 2$, and $\mu_{1j} > 0$, $\mu_{2j} = 0$, all j. There is no loss in surplus from using p as the sole performance measure if, and only if, $\sigma_{x1} \neq 0$ and $\sigma_{x2} = 0$.

(c) Suppose $m = 2$, $\mu_{jj} = b_j$ all j, $\mu_{ij} = 0$ all $i \neq j$, and $\rho = 0$. There is no loss in surplus from using p as the sole performance measure if, and only if,

$$\frac{\sigma_{xi}}{\sigma_i^2} = \lambda \frac{b_i^2}{b_i^2 + r\sigma_i^2} \quad \text{for all } i \text{ and some } \lambda \neq 0. \tag{17}$$

In case (a) there is a single public report, such as accounting earnings. The corollary states that p and y_1 are equivalent performance measures if y_1 provides investors with information about the *uncontrollable events* influencing the terminal value of the firm. With a single signal, the congruity of that signal is immaterial. Observe that p has no value as a performance measure if y_1 is a noisy measure of $\boldsymbol{a}$ and the noise is due to measurement error that tells investors nothing about ε_x. In that setting it is important to use y_1, and not p, as the performance measure.

Case (b) is similar to the risk reduction example in the preceding section. Signal y_1 can be interpreted as accounting earnings, which is influenced by the agent's actions, and y_2 can be interpreted as information about uncontrollable events that influence x and/or y_1. Again, p and y_1 are interchangeable performance measures if y_1 provides investors with information about the uncontrollable events influencing the terminal value of the firm. The key issue is whether y_2 can be ignored if p is used as the performance measure. Somewhat surprisingly, the answer is yes if, and only if, y_2 provides information about the uncontrollable events influencing y_1 but provides *no direct information about the uncontrollable events influencing* x, e.g., y_2 reports measurement errors in y_1. Observe that if $\sigma_{x2} = 0$, then y_2 provides information about the *noise* in y_1 and that information is used in the same way for both evaluating performance and predicting the value of the firm: $y_a \propto y_x \propto y_1 - (\rho\sigma_1/\sigma_2)y_2$.[21] If $\sigma_{x2} \neq 0$, then y_2 provides direct information about the x. This information improves the relation between p and x, but it creates additional noise in the relation between $\boldsymbol{a}$ and p. This is seen most clearly in the setting in which $\rho = 0$ and $\sigma_{x2} \neq 0$. In this setting, y_2 provides information about uncontrollable events that affect x but not y_1, e.g., information about market-wide events that will influence the terminal value of the firm but have not influenced current accounting earnings. Hence, y_2 is ignored in the optimal incentive contract ($y_a \propto y_1$), but it influences the price ($y_x \propto (\sigma_{x1}/\sigma_1^2)y_1 + (\sigma_{x2}/\sigma_2^2)y_2$). Hence, p is a *strictly noisier* performance measure than y_1. Of course, p can be used to obtain the same surplus as y_1 if both p and y_2 are used. Observe that the above results are independent of the congruity of the first signal.

In case (c) there are two independent signals—one for each task. Paul (1992) explores this type of setting. He interprets it as representative of a firm in which the agent manages several divisions and publicly reports information about the profitability of each division.[22] Since the scaling of a signal is arbitrary, we assume that each signal is scaled so that $\mu_{jj} = b_j$. An efficient aggregate performance measure uses divisional weights that reflect the noise in each signal and the rate at which the agent's effort generates value in each division ($y_a = y_1 b_1^2/(b_1^2 + r\sigma_1^2) + y_2 b_2^2/(b_2^2 + r\sigma_2^2)$). Price, on the other hand, uses weights that reflect only the relation between the noise in the divisional signal and the payoff that will ultimately be generated by that division ($y_p = y_1\sigma_{x1}/\sigma_1^2 + y_2\sigma_{x2}/\sigma_2^2$). Other than a setting in which the two tasks are equivalent (i.e., $b_1 = b_2$, $\sigma_1 = \sigma_2$, and $\sigma_{x1} = \sigma_{x2}$), there do not appear to be any contexts in which condition (17) is satisfied (other than by coincidence). For example, assume the signals report the terminal value generated by each task (i.e., $x = y_1 + y_2$). In that setting, $\sigma_{xi} = \sigma_i^2$ and condition (17) holds if, and only if, $b_1/b_2 = \sigma_1/\sigma_2$. That is, the relative noise in the two signals must equal the relative value generated by the two tasks.

The preceding discussion suggests that if there are multiple tasks and multiple public signals that are influenced by the agent's action, then it is unlikely that the market price provides an efficient single performance measure, even though it provides an efficient prediction of the firm's future value.[23] Nonetheless, we find extensive use of market price as a performance measure, in that managers often hold stock, or stock options, in their own firm. There are two potential explanations for the use of price as a performance measure. First, the price may impound investor information that is not publicly reported but is useful in evaluating the agent's performance. For example, Bushman and Indjejikian (1993) provide a rational expectations model in which price impounds private investor information and a public report (e.g., an earnings report).[24] The optimal incentive contract is based on both the market price and the public report. Second, while the price

[21] "$\propto$" denotes "proportional to," i.e., the term on the left is equal to the term on the right times some constant. In this setting,

$$\Sigma_{xy}\Sigma^{-1} = \left[\boldsymbol{v}_x \quad \rho\tfrac{\sigma_1}{\sigma_2}\boldsymbol{v}_x\right] \qquad \boldsymbol{v}_x = \frac{\sigma_{x1}}{\sigma_1^2(1-\rho^2)}$$

$$b^t\boldsymbol{\mu}^t\boldsymbol{Q} = \left[\boldsymbol{v}_a \quad \rho\tfrac{\sigma_1}{\sigma_2}\boldsymbol{n}_a\right] \qquad \boldsymbol{v}_a = \frac{b_1\mu_{11} + b_2\mu_{12}}{\mu_{11}^2 + \mu_{12}^2 + r(1-\rho^2)\sigma_1^2}$$

[22] Paul (1991) considers two actions and interprets them as pertaining to short and long run cash flows. He demonstrates that overemphasis on either short run or long run actions can occur, depending upon which type of information has the most pronounced effect on price.

[23] Paul (1991,1992) also emphasizes this point. Dye (1985) makes a similar point in the context of discussing the potential impact of disclosure of payoff information received by the agent after he has taken his action. Gjesdal (1981) demonstrates that the optimal aggregation of information in valuation decisions is likely to differ from the optimal aggregation used in performance evaluation.

[24] In Bushman and Indjejikian (1993) the investors receive private information that is equal to x plus *noise*. The in-

may not be a perfect aggregate performance measure, it is readily observed and may provide a cost-effective aggregate measure when contracting costs are taken into consideration.

Interestingly, price is *not* necessarily, nor even likely to be, a perfectly congruent performance measure. Examination of pricing relation (15) reveals that while $\boldsymbol{b}^t\hat{\boldsymbol{a}}$ is an important ingredient in the price, the congruity of the price as a performance measure depends on $\Sigma_{xy}\Sigma^{-1}\boldsymbol{\mu}$ and this may not be proportional to $\boldsymbol{b}$. The congruity of the price depends on the congruity and diversity of the information received by investors and the weights they give that information in their prediction of x.

VI. Concluding Remarks

Our analysis has examined the role of multiple performance measures in influencing the direction and intensity of an agent's effort. A perfectly congruent performance measure will induce the first-best direction, but it will not be used to induce the first-best intensity unless it is noiseless. Noisy performance measures create risk, and the agent must be compensated for that risk. The stronger the incentives used, the larger the risk premium that must be paid. Hence, noise weakens the incentives that are used. If the basic measure is perfectly congruent, then the primary role for additional performance measures is to reduce the risk that must be imposed on the agent. However, the use of noncongruent performance measures would produce a second-best action that differs from the first-best action in both direction and intensity.

A single noiseless performance measure cannot be used to achieve the first-best result unless it is perfectly congruent. The direction induced by a single noncongruent measure can be improved by using a set of diverse performance measures. The first-best result can be achieved with multiple noiseless performance measures if they span the first-best direction. Of course, if the additional measures are noisy, then the second-best action may differ from the first-best action in both direction and intensity (even though the first-best action is feasible).

The market price at the contract termination date is a potentially useful performance measure, either as a replacement for some other measure or because it is influenced by noncontractible information. The market price reflects investor beliefs at the trading date and those beliefs are influenced by the information they receive. In particular, the price reflects the investors' equilibrium beliefs about the actions taken by the manager and the information they have about the uncontrollable events that will influence the terminal value of the firm. The investors are not concerned with incentives when they use their information to set the price, and hence price is not likely to be an efficient aggregate performance measure. That is, the relative weighting given to signals in the price need not equal the weightings those signals would be given in the second-best incentive contract. Barring exceptional circumstances, a second-best incentive contract requires that the price be supplemented with other measures, even though those measures are impounded in the price.

---

vestors also receive an accounting report, which may take one of two forms. In the first case, the accounting report also equals x plus noise, which implies that the price (which impounds both types of information) is perfectly congruent with the principal's gross payoff. In the second case, the accounting report equals the payoff generated by the effort in one task, plus noise. Since the first measure is perfectly congruent and the second is not, the price is not perfectly congruent with the principal's gross payoff.

Appendix A
Proofs

Q.E.D.

Proposition 1:

Sufficiency: $\mu_1 = \lambda \boldsymbol{b}^t$ implies that $\delta_1 = 0$. $\lambda \neq 0$ implies that $\boldsymbol{Q}$ is finite, even if $r\sigma_1^2 = 0$. Hence, $L(\eta) = 0$ for the stated conditions, follows directly from equation (11).
Necessity: All terms in the numerator and denominator are nonnegative. Hence, zero loss only occurs if both δ_1 and $r\sigma_1^2\boldsymbol{b}^t\boldsymbol{b}$ equal zero. $\delta_1 = 0$ implies $b_1/b_2 = \mu_{11}/\mu_{12}$, which implies the existence of λ such that $\mu_{1j} = \lambda b_j$. Since $\boldsymbol{b}^t\boldsymbol{b} > 0$, $r\sigma_1^2\boldsymbol{b}^t\boldsymbol{b} = 0$ only if $r\sigma_1^2 = 0$. **Q.E.D.**

Proposition 2:

From equation (9), $V(\eta_+) - V(\eta) = \tfrac{1}{2}\boldsymbol{b}[\boldsymbol{D}_+ - \boldsymbol{D}]\boldsymbol{b}^t$,

where $\boldsymbol{D}_+ = \boldsymbol{\mu}_+^t\boldsymbol{Q}_+\boldsymbol{\mu}_+$

and $\boldsymbol{Q}_+ = [\boldsymbol{\mu}_+\boldsymbol{\mu}_+^t + r\Sigma_+]^{-1}$.

Observe that $\boldsymbol{H} \equiv \boldsymbol{\mu}_+\boldsymbol{\mu}_+^t + r\Sigma_+ = \begin{bmatrix} \boldsymbol{H}_{11} & \boldsymbol{H}_{12} \\ \boldsymbol{H}_{21} & \boldsymbol{H}_{22} \end{bmatrix}$.

It then follows that

$$\boldsymbol{Q}_+ = \begin{bmatrix} \boldsymbol{Q}+\boldsymbol{Q}\boldsymbol{H}_{12}\boldsymbol{P}\boldsymbol{H}_{21}\boldsymbol{Q} & -\boldsymbol{Q}\boldsymbol{H}_{12}\boldsymbol{P} \\ -\boldsymbol{P}\boldsymbol{H}_{21}\boldsymbol{Q} & \boldsymbol{P} \end{bmatrix},$$

which in turn implies

$$\boldsymbol{D}_+ = \boldsymbol{\mu}^t\boldsymbol{Q}\boldsymbol{\mu} + \boldsymbol{\mu}^t\boldsymbol{Q}\boldsymbol{H}_{12}\boldsymbol{P}\boldsymbol{H}_{21}\boldsymbol{Q}\boldsymbol{\mu} - \boldsymbol{\mu}_2^t\boldsymbol{P}\boldsymbol{H}_{21}\boldsymbol{Q}\boldsymbol{\mu} - \boldsymbol{\mu}^t\boldsymbol{Q}\boldsymbol{H}_{12}\boldsymbol{P}\boldsymbol{\mu}_2 + \boldsymbol{\mu}_2^t\boldsymbol{P}\boldsymbol{\mu}_2.$$

Therefore, $\boldsymbol{b}^t[\boldsymbol{D}_+ - \boldsymbol{D}]\boldsymbol{b} = \boldsymbol{b}^t(\boldsymbol{\mu}_2 - \boldsymbol{H}_{21}\boldsymbol{Q}\boldsymbol{\mu})^t\boldsymbol{P}(\boldsymbol{\mu}_2 - \boldsymbol{H}_{21}\boldsymbol{Q}\boldsymbol{\mu})\boldsymbol{b} = \boldsymbol{T}^t\boldsymbol{P}\boldsymbol{T}$.

To prove that $\boldsymbol{P}$ is positive definite, we first observe that both $\boldsymbol{\mu}_+\boldsymbol{\mu}_+^t$ and Σ_+ are both positive semidefinite (and Σ_+ is positive definite if there are no noiseless performance measures). With $r \geq 0$, it then follows that $\boldsymbol{H}$ is positive semidefinite. Since $\boldsymbol{Q}_+ = \boldsymbol{H}^{-1}$ is assumed to exist, it follows that both $\boldsymbol{H}$ and $\boldsymbol{H}^{-1}$ are positive definite. Finally, the fact that $\boldsymbol{H}^{-1}$ is positive definite implies that $\boldsymbol{P} = \boldsymbol{H}_{22} - \boldsymbol{H}_{21}\boldsymbol{H}_{11}^{-1}\boldsymbol{H}_{12}$ is positive definite. **Q.E.D.**

Lemma 1:

In general, $f(\boldsymbol{y}_1, \boldsymbol{y}_2|\boldsymbol{a}) = g(\boldsymbol{y}_2|\boldsymbol{y}_1, \boldsymbol{a})h(\boldsymbol{y}_1|\boldsymbol{a})$, where g is the conditional density for $\boldsymbol{y}_2$ given $\boldsymbol{y}_1$ and $\boldsymbol{a}$, and h is the marginal density function for $\boldsymbol{y}_1$ given $\boldsymbol{a}$. Sufficiency is satisfied by showing that condition (13) implies the conditional mean is independent of $\boldsymbol{a}$ (this parameter has no impact on the variance):

$$E\left[y_2|y_1, \boldsymbol{a}\right] = \boldsymbol{\mu}_2 a + \Sigma_{21}\Sigma_{11}^{-1}\left[y_1 - \boldsymbol{\mu}_1\boldsymbol{a}\right] = \Xi\left(\boldsymbol{\mu}_1\boldsymbol{a} + \Sigma_{21}\Sigma_{11}^{-1}\left[y_1 - \boldsymbol{\mu}_1\boldsymbol{a}\right]\right) = \Xi\, y_1.$$

See Raiffa and Schlaifer (1961) for a general derivation of the conditional mean and variance for normal distributions. **Q.E.D.**

Proposition 3:

$\pi(\boldsymbol{\eta}_+,\boldsymbol{\eta}) = 0$ if, and only if, $\boldsymbol{T} = \mathbf{0}$. The sufficiency of the conditions follows from

$$\begin{aligned} \boldsymbol{T} &= (\boldsymbol{\mu}_2 - \boldsymbol{H}_{21}\boldsymbol{Q}\boldsymbol{\mu})\boldsymbol{b} \\ &= (\Xi\boldsymbol{M} - (\Xi\boldsymbol{\mu}\boldsymbol{\mu}^t + r\Xi\Sigma)\boldsymbol{Q}\boldsymbol{\mu})\boldsymbol{b} \\ &= \Xi(\boldsymbol{\mu} - (\boldsymbol{\mu}\boldsymbol{\mu}^t + r\Sigma)\boldsymbol{Q}\boldsymbol{\mu})\boldsymbol{b} = \mathbf{0}. \end{aligned}$$

The necessity of the conditions is demonstrated as follows.

$$\boldsymbol{T} = \mathbf{0} \quad \Rightarrow \quad \boldsymbol{\mu}_2 - \boldsymbol{\mu}_2\boldsymbol{\mu}^t\boldsymbol{Q}\boldsymbol{\mu} = r\Sigma_{21}\boldsymbol{Q}\boldsymbol{\mu},$$

$$\Rightarrow \quad \boldsymbol{\mu}_2 = r\Sigma_{21}\boldsymbol{Q}\boldsymbol{\mu}[\boldsymbol{I} - \boldsymbol{\mu}^t\boldsymbol{Q}\boldsymbol{\mu}]^{-1} = r\Sigma_{21}\boldsymbol{Q}\boldsymbol{\mu}[\boldsymbol{I} + \boldsymbol{\mu}^t(r\Sigma)^{-1}\boldsymbol{\mu}] = r\Sigma_{21}\boldsymbol{Q}[\boldsymbol{I} + \boldsymbol{\mu}\boldsymbol{\mu}^t(r\Sigma)^{-1}]\boldsymbol{\mu}.$$

Let $\quad \Xi = r\Sigma_{21}\boldsymbol{Q}[\boldsymbol{I} + \boldsymbol{\mu}\boldsymbol{\mu}^t(r\Sigma)^{-1}].$

Hence, $\Xi\boldsymbol{\mu} = \boldsymbol{\mu}_2$ follows directly and

$$\Xi\Sigma = r\Sigma_{21}\boldsymbol{Q}[\boldsymbol{I} + \boldsymbol{\mu}\boldsymbol{\mu}^t(r\Sigma)^{-1}]\Sigma = \Sigma_{21}\boldsymbol{Q}[r\Sigma + \boldsymbol{\mu}\boldsymbol{\mu}^t] = \Sigma_{21}.$$

Q.E.D.

Proposition 4:

If $\mu_{11}>0$, $\mu_{12}>0$, and $\mu_{21}=\mu_{22}=\mathbf{0}$, then $\pi(\eta_+,\eta) = \frac{1}{2}T^2P$,

where $\quad P = [\Delta Q]^{-1},$

$$T = -\,r\rho\sigma_1\sigma_2\boldsymbol{Q}\boldsymbol{\mu}\boldsymbol{b},$$

$$Q = [\boldsymbol{\mu}\boldsymbol{\mu}^t + r\sigma_1^2]^{-1},$$

$$\Delta = r\sigma_2^2[\boldsymbol{\mu}\boldsymbol{\mu}^t + r\sigma_1^2(1-\rho^2)].$$

(i) If $r\sigma_1 > 0$ and $\boldsymbol{\mu} > 0$, then $P > 0$ and $T = 0$ if, and only if, $\rho = 0$ (note: $\sigma_2 = 0 \Rightarrow \rho = 0$).

(ii) $\mathrm{L}(\eta_+) = \dfrac{1}{2}\dfrac{r\sigma_2^2}{\Delta}\left[r\sigma_1^2(1-\rho^2)\boldsymbol{b}\boldsymbol{b}^t + \delta_1\right].$

Hence, L = 0 if, and only if $\rho^2 = 1$ and $\delta_1 = 0$ (which requires $\boldsymbol{\mu} = \boldsymbol{\lambda b}$ for some λ).

(iii) $\boldsymbol{v}^{\ddagger} = \boldsymbol{Q}_+\boldsymbol{\mu}_+\boldsymbol{b} = \dfrac{\mathrm{r}}{\Delta}\boldsymbol{\mu}\boldsymbol{b}\begin{bmatrix} \sigma_2^2 \\ -\rho\sigma_1\sigma_2 \end{bmatrix}.$

Proposition 5:

Observe that an incentive contract v_p that is based solely on p can be replicated by a contract based on $\boldsymbol{y}$ by letting $v^t = v_p\Sigma_{xy}\Sigma^{-1}$. Since $v^\ddagger = \boldsymbol{Q\mu b}$ is the unique optimal contract based on $\boldsymbol{y}$, it follows that the incentive contract based on p produces the same payoff to the principal as $\boldsymbol{v}^\ddagger$ if, and only if, $\boldsymbol{v}^{\ddagger t} = \boldsymbol{v}_p\Sigma_{xy}\Sigma^{-1}$. The stated condition then holds with $\lambda = 1/\boldsymbol{v}_p$ **Q.E.D.**

Appendix B
Minimum Cost Contracts for the Set of Implementable Actions

Grossman and Hart [GH] (1983) use a two step approach to solve the principal's problem in a standard agency problem. They first determine the minimum cost incentive contract for inducing each action $\boldsymbol{a} \in \boldsymbol{A}$, and then they identify the action for which the difference between the expected benefit and the expected compensation cost is maximized. The GH approach can be adapted to our setting, but it must be modified because in a multi-task setting it may not be possible to induce all actions. For example, in figure 1 only the actions on the line denote $\boldsymbol{A}(\boldsymbol{\eta})$ are implementable. With multiple performance measures, the set of implementable actions is

$$\boldsymbol{A}(\boldsymbol{\eta}) \equiv \{\, \boldsymbol{a} \mid \text{for some } v_1,\ldots,v_m,\ a_j = \sum_{i=1}^{m} v_i\mu_{ij} \text{ for all } a_j > 0, \text{ and } \sum_{i=1}^{m} v_i\mu_{ij} \le 0 \text{ for all } a_j = 0 \,\}.$$

An interior action (i.e., $a_j > 0$ all j) is implementable if it is *spanned* by the set of performance measure coefficients. This can be interpreted as a setting in which there exist performance measure weights such that an *aggregate measure* can be formed that is perfectly congruent with the desired action. The following is a formal statement of this spanning condition.

Lemma B.1:

(a) Interior action $\boldsymbol{a} \in \boldsymbol{A}(\boldsymbol{\eta})$ if, and only if, rank($\boldsymbol{\mu}$) = rank($\boldsymbol{M}(\boldsymbol{a})$), where $\boldsymbol{M}(\boldsymbol{a}) \equiv [\,\boldsymbol{\mu}^t\ \boldsymbol{a}\,]^t$.

(b) If Σ is positive definite, rank($\boldsymbol{\mu}^t\Sigma\boldsymbol{\mu}$) = rank($\boldsymbol{\mu}$). Hence, $\boldsymbol{\mu}^t\Sigma\boldsymbol{\mu}$ is non-singular if rank($\boldsymbol{\mu}$) = n.

Proof:

(a) This follows from the requirements for a solution to a system of linear equations.

(b) Since Σ is positive definite, there exists a non-singular $m\times m$ matrix $\boldsymbol{N}$ such that $\Sigma = \boldsymbol{N}^t\boldsymbol{N}$ and rank($\boldsymbol{N}$) = m. From the rank preservation theorem (Ortega, 1987, 61), rank($\boldsymbol{\mu}^t\boldsymbol{N}^t$) = rank($\boldsymbol{N\mu}$) = rank($\boldsymbol{\mu}$). Furthermore, rank($\boldsymbol{\mu}^t\boldsymbol{N}^t\boldsymbol{N\mu}$) = rank($\boldsymbol{N\mu}$) (Ortega, 1987, 63).

For any implementable action $\boldsymbol{a}$ there may be one or more incentive contracts that can be used to induce that action. If the performance measures are noiseless ($\Sigma = \boldsymbol{0}$) or the agent is risk neutral (r = 0), then the choice among the feasible incentive contracts is immaterial (assuming β is selected so that the contract pays the agent C($\boldsymbol{a}$) + w^o). Furthermore, if the performance measures are noiseless and the first-best action is implementable, then the first-best result can be achieved with either a linear incentive contract or a penalty contract.

Proposition B.1: Assume $\Sigma = \boldsymbol{0}$ and $\boldsymbol{a}^* = \boldsymbol{b} \in \boldsymbol{A}(\boldsymbol{\eta})$. The first-best result can be achieved with either

(a) any linear contract ($\beta, \boldsymbol{v}$) such that $\boldsymbol{\mu}^t\boldsymbol{v} = \boldsymbol{b}$ and β = C($\boldsymbol{b}$) + w^o - $\boldsymbol{v}^t\boldsymbol{\mu b}$, or

(b) a penalty contract in which $w(\boldsymbol{y}) = C(\boldsymbol{b}) + w^o$ if $\boldsymbol{y} = \boldsymbol{y}^* \equiv \boldsymbol{\mu b}$, and zero otherwise.

Proof:

(a) Lemma B.1(a) establishes that $\boldsymbol{b} \in \boldsymbol{A}(\boldsymbol{\eta})$ implies the existence of a vector $\boldsymbol{v}$ such that $\boldsymbol{\mu}^t\boldsymbol{v} = \boldsymbol{b}$. It follows from first order condition (8) that the agent chooses $\boldsymbol{a} = \boldsymbol{\mu}^t\boldsymbol{v} = \boldsymbol{b}^t$.

(b) Given the penalty contract, the agent chooses $\boldsymbol{a} \in \boldsymbol{A}^*(\boldsymbol{y}^*) \equiv \arg\min C(\boldsymbol{a})$ subject to $\boldsymbol{y}^* = \boldsymbol{\mu a}$. The first-order conditions for the agent's problem establish that the agent chooses $\boldsymbol{a} = \boldsymbol{\mu}^t\boldsymbol{v}$, where $\boldsymbol{v}$ is an $m\times 1$ vector of Lagrange multipliers for the m performance level constraints. The definition of $\boldsymbol{y}^*$ implies that $\boldsymbol{a}^*$ satisfies the performance level constraints, and the existence of the desired Lagrange multipliers follows from the implementability of $\boldsymbol{a}^*$ and Lemma B.1. **Q.E.D.**

Approach (a) can also be used if the agent is risk neutral, if the performance measures are noisy. Approach (b) is the multi-task equivalent of the penalty contract used when there is moving support in the single-dimensional action case. The key here is that either congruent or multiple performance measures are required to implement an effective penalty contract when a variety of effort combinations can be used to attain a given performance level. The first-best action must be the *least cost action* for implementing the specified performance levels.

If the performance measures are noisy and the agent is risk averse, then the least cost contract is the feasible contract that minimizes the risk imposed on the agent. The following proposition identifies the optimal contract to implement action $\boldsymbol{a}$ when Σ is positive definite and there are an infinite number of contracts that can be used to induce $\boldsymbol{a}$. (From Lemma (B.1), $\boldsymbol{a} \in \boldsymbol{A}(\boldsymbol{\eta})$ if, and only if, rank($\boldsymbol{M}(\boldsymbol{a})$) = rank($\boldsymbol{\mu}$), where $\boldsymbol{M}(\boldsymbol{a}) = [\, \boldsymbol{\mu}^t\ \boldsymbol{a}\,]^t$. If this condition holds, there is at least one incentive contract $\boldsymbol{v}$ that induces $\boldsymbol{a}$. The least cost contract can be expressed as a closed form solution if rank($\boldsymbol{\mu}$) = n, but not if rank($\boldsymbol{\mu}$) < n.)

Proposition B.2: Assume Σ is positive definite and consider implementable, interior action $\boldsymbol{a}$.

(a) If rank($\boldsymbol{\mu}$) = m, then there is a unique contract for implementing $\boldsymbol{a}$ and, if rank($\boldsymbol{\mu}$) = $\boldsymbol{n}$, that contract is:

$$\boldsymbol{v} = [\boldsymbol{\mu}^t]^{-1}\boldsymbol{a}. \tag{B.1}$$

(b) If rank($\boldsymbol{\mu}$) < $\boldsymbol{m}$, then there are an infinite number of contracts that can be used to induce $\boldsymbol{a}$ and, if rank($\boldsymbol{\mu}$) = $\boldsymbol{n}$, the least cost contract is:

$$\boldsymbol{v} = \Sigma^{-1}\boldsymbol{\mu}[\boldsymbol{\mu}^t\Sigma^{-1}\boldsymbol{\mu}]^{-1}\boldsymbol{a}. \tag{B.2}$$

Proof:

Lemma B.1(a) establishes that the implementability of $\boldsymbol{a}$ implies rank($\boldsymbol{\mu}$) = rank($\boldsymbol{M}(\boldsymbol{a})$).

(a) If rank($\boldsymbol{\mu}$) = rank($\boldsymbol{M}(\boldsymbol{a})$) = $m \le n$, then it is well known that there is a unique solution to the system of $\boldsymbol{n}$ equations in m unknowns. If $m = n$, then $\boldsymbol{\mu}^{-1}$ exists and equation (B.1) is obtained from condition (8).

(b) If rank($\boldsymbol{\mu}$) = rank($\boldsymbol{M}(\boldsymbol{a})$) < m, then there are only rank($\boldsymbol{\mu}$) independent equations for m unknowns and m -rank($\boldsymbol{\mu}$) > 0 variables can be chosen arbitrarily. Hence, there exist infinitely many incentive contracts that can implement $\boldsymbol{a}$. The one that minimizes $\boldsymbol{v}^t\Sigma\boldsymbol{v}$ can be obtained by solving the following problem:

$$\min_{v} \quad \boldsymbol{v}^t\Sigma\boldsymbol{v} \quad \text{subject to } \boldsymbol{\mu}^t\boldsymbol{v} = \boldsymbol{a}. \tag{B.3}$$

The Lagrangian for this problem is

$$\boldsymbol{v}^t\Sigma\boldsymbol{v} - \Lambda^t(\boldsymbol{\mu}^t\boldsymbol{v} - \boldsymbol{a}),$$

where Λ is an $n\times 1$ vector of Lagrange multipliers. The first-order condition is

$$2\Sigma\boldsymbol{v} = \boldsymbol{\mu}\Lambda \Leftrightarrow \boldsymbol{v} = \tfrac{1}{2}\Sigma^{-1}\boldsymbol{\mu}\Lambda. \tag{B.4}$$

If rank($\boldsymbol{\mu}$) = n, then $\boldsymbol{\mu}^t\Sigma^{-1}\boldsymbol{\mu}$ is non-singular (see Lemma B.1(b)) and we can solve for Λ by substituting $\boldsymbol{v}$ into condition (B.3) to get

$$\tfrac{1}{2}\boldsymbol{\mu}^t\Sigma^{-1}\boldsymbol{\mu}\Lambda = \boldsymbol{a} \Leftrightarrow \Lambda = 2[\boldsymbol{\mu}^t\Sigma^{-1}\boldsymbol{\mu}]^{-1}\boldsymbol{a}.$$

Substituting for Λ in (B.4) provides expression (B.2). **Q.E.D.**

To illustrate the above result, consider the two-task example in table 1. The single performance measure case is the same as in figure 1; the first performance measure cannot be used to implement the first-best action. With two noncongruent performance measures, there is a unique contract for implementing each action, including the first-best action. While the first-best action can be implemented in this setting, the second-best contract implements a different action. This arises from the tradeoff between the loss from noncongruent actions and the cost of risk. With three noncongruent performance measures there are an infinite number of contracts that can implement any action. The least cost contract minimizes the risk imposed on the agent. Hence, even though the third measure is perfectly congruent with the first-best action, it is given little weight because it is very noisy. Again, risk considerations result in a second-best action that differs from the first-best action.

References

Amershi, A. H. 1988. Blackwell informativeness and sufficient statistics with applications to financial markets and multiperson agencies. In *Economic Analysis of Information and Contracts*, edited by G. A. Feltham, A. H. Amershi, and W. T. Ziemba. Boston: Kluwer Academic Publishers, 25–91.

———, R. D. Banker, and S. M. Datar. 1990. Economic sufficiency and statistical sufficiency in the aggregation of accounting signals. *The Accounting Review* 65 (January): 113–130.

———, and J. S. Hughes. 1989. Multiple signals, statistical sufficiency, and Pareto orderings of best agency contracts. *RAND Journal of Economics* 20 (Spring): 102–112.

Anthony, R. N., J. Dearden, and V. Govindarajan. 1992. *Management Control Systems*, Seventh Edition. Boston: Irwin.

Antle, R., and J. S. Demski. 1988. The controllability principle in responsibility accounting. *The Accounting Review* 62 (October): 700–718.

———, and A. Smith. 1986. An empirical investigation of the relative performance evaluation of corporate executives. *Journal of Accounting Research* 24 (Spring): 1–39.

Baiman, S., and J. Noel. 1985. Noncontrollable costs and responsibility accounting. *Journal of Accounting Research* 23 (Autumn): 486–501.

Banker, R. D., and S. M. Datar. 1989. Sensitivity, precision, and linear aggregation of signals for performance evaluation. *Journal of Accounting Research* 27 (Spring): 21–39.

Bushman, R. M. and R. J. Indjejikian. 1993. Accounting income, stock price and managerial compensation. *Journal of Accounting and Economics* 16 (January/April/July): 3–23.

Dye, R. A. 1985. Disclosure of nonproprietary information. *Journal of Accounting Research* 23 (Spring): 123–145.

———. 1988. Earnings management in an overlapping generations model. *Journal of Accounting Research* 26 (Autumn): 195–235.

Gjesdal, F. 1981. Accounting for stewardship. *Journal of Accounting Research* 19 (Spring): 208–231.

———. 1982. Information and incentives: The agency information problem. *Review of Economic Studies* (July): 373–390.

Grossman, S., and O. D. Hart. 1983. An analysis of the principal-agent problem. *Econometrica* 51 (January): 7–45.

Holmstrom, B. 1979. Moral hazard and Observability. *Bell Journal of Economics* 10 (Spring): 74–91.

———. 1982. Moral hazard in teams. *Bell Journal of Economics* 13 (Autumn): 324–340.

———, and P. Milgrom. 1987. Aggregation and linearity in the provision of intertemporal incentives. *Econometrica* 55: 303–328.

———, and ———. 1991. Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics and Organizations* 7: 24–52.

Horngren, C. T., and G. Foster. 1991. *Cost Accounting: A Managerial Emphasis*, Seventh Edition. Englewood Cliffs, New Jersey: Prentice Hall.

Kanodia, C., R. Bushman, and J. Dickhaut. 1989. Escalation errors and sunk cost effect: An explanation based on reputation and information asymmetries. *Journal of Accounting Research* 27 (Spring): 59–77.

Kaplan, R., and A. A. Atkinson. 1989. *Advanced Management Accounting*, Second Edition. Englewood Cliffs, New Jersey: Prentice Hall.

Magee, R. P. 1978. Accounting measurement and employment contracts: Current value reporting. *Bell Journal of Economics* 9 (Spring): 145–158.

———. 1986. *Advanced Managerial Accounting*. New York: Harper and Row.

Maher, M. W. 1987. The use of relative performance evaluation in organizations. In *Accounting and Management: Field Study Perspectives*, edited by W. J. Bruns, Jr. and R. S. Kaplan. Boston: Harvard Business School Press, 295–315.

Matsumura, E. M. 1988. Sequential choice under moral hazard. In *Economic Analysis of Information and Contracts*, edited by G. A. Feltham, A. H. Amershi, and W. T. Ziemba. Boston: Kluwer Academic Publishers, 221–245.

Merchant, K. A. 1987. How and why firms disregard the controllability principle. In *Accounting and Management: Field Study Perspectives*, edited by W. J. Bruns, Jr. and R. S. Kaplan. Boston: Harvard Business School Press, 316–338.

Ortega, J. M. 1987. *Matrix Theory: A Second Course*. New York: Plenum Press.

Paul, J. M. 1991. Managerial myopia and the observability of future cash flows. Working Paper, University of Michigan, Ann Arbor, Michigan.

———. 1992. On the efficiency of stock-based compensation. *Review of Financial Studies* 5: 471–502.

Raiffa, H., and R. Schlaifer. 1961. *Applied Statistical Decision Theory*. Boston: Graduate School of Business, Harvard University.