DOI: 10.1111/j.1475-679X.2007.00246.x
Journal of Accounting Research
Vol. 45 No. 3 June 2007
Printed in U.S.A.

Performance Measure Congruity and the Balanced Scorecard

JÖRG BUDDE*

Received 15 September 2005; accepted 20 December 2006

ABSTRACT

This paper studies the incentive effects of a balanced scorecard within a multitask agency framework under both formal and relational contracts. First, the main characteristics of the balanced scorecard are analyzed with respect to performance measure congruity. It is shown that under complete verifiability, a properly designed balanced scorecard is capable of perfectly aligning the interests of owners and employees by means of an explicit contract. I then investigate whether subjective performance evaluation is beneficial when not all the scorecard measures are contractible. It emerges that congruity of the contractible scorecard measures constrains a purely implicit incentive contract, but the first-best solution may still be obtained through a combination of formal and relational contracts. Furthermore, a purely explicit contract in most cases can be improved by incorporating subjective rewards.

1. *Introduction*

The last decade has witnessed the rise of nonfinancial performance measures. Since in a dynamic environment financial measures may not properly represent a firm's prospects, several methods have been proposed to capture the long-term effects of managerial activities. The balanced scorecard, which is the most prominent of these new concepts, provides a framework in which both financial and nonfinancial success measures are linked by the the firm's *strategy*.

Following Kaplan and Norton [1992, 1996], the primary objective of the balanced scorecard is to identify, communicate, and implement strategy

*University of Bonn, Germany.

within a firm. It has been broadly accepted, however, that organizational enforcement requires the tying of employee compensation to scorecard measures (e.g., Kaplan and Norton [2001, p. 151]). Both Kaplan and Norton [2001, particularly chap. 10] and Banker, Potter, and Srinivasan [2000] provide empirical evidence of this assertion, through examples of alteration in employee behavior or the organization's financial performance, after remuneration had been related to nonfinancial measures.

Although the general importance of such incentive provision is hardly disputed, its implementation is still a matter of controversy. Kaplan and Norton [1996, p. 81] consider explicit, formula-based incentive plans with some skepticism. They suspect that "... unintended or unexpected consequences [could] arise from the way the targets for the measures are achieved." In particular, they presume that defining a personal performance score as the weighted sum of multiple performance measures may result in extensive payments, even if the employee has focused on only some of the objectives. To obtain a more balanced provision of managerial effort, they propose that incentive payments should only be granted if prespecified threshold levels are met for a whole set of critical measures. They go on to suggest that considering the deviations between budgeted and realized amounts, and thereby improving the manager's performance assessment, might be an even more promising approach. In their estimation the results of this assessment could provide a basis for subjective rewards (Kaplan and Norton [2001, p. 267]).[1]

The present paper analyzes potential economic effects supporting the two procedures just described. In particular, I study (1) why unintended consequences could result from tying a manager's compensation to scorecard measures and (2) whether subjective performance evaluation is capable of alleviating this deficiency. The first issue, once generally highlighted as "the folly of rewarding A, while hoping for B" (Kerr [1975]), has long been confirmed in the economics literature and has been studied extensively in a multitask agency framework (e.g., Holmström and Milgrom [1991], Feltham and Xie [1994]). By applying this model to the present topic, the question of why the balanced scorecard should give only a *distorted* picture of performance can be addressed. Because scorecard measures are by definition intended to provide a *balanced* image of the value-generating factors, the mere fact of performance distortion is somewhat controversial. I provide a formal definition of scorecard *balance*, and compare it to the property of congruity as defined in the multitask agency setting. It is shown that balance is a sufficient condition for performance measure congruity, and first best can be achieved if the agent is risk neutral. This result, however, builds on

---

[1] Alternatively, the weighting of measures could be left to the principal's discretion. Ittner, Larcker, and Meyer [2003] provide a field study of how weights are chosen in a major financial services company. Contrary to Kaplan and Norton's [1996] intention for scorecard systems, however, they find that most of the weight is given to financial measures.

the assumption that all scorecard measures can be used in an explicit contract. Since contractibility could be in doubt for nonfinancial performance measures, I also investigate whether incongruity of the verifiable measures may be compensated for by using a subjective performance evaluation. To this purpose, I adapt the model of relational contracts proposed by Baker, Gibbons, and Murphy [1994] to the multitask agency setting. I extend their work by considering multidimensional performance measures for which the design of an aggregate performance score becomes part of the contracting issue. Although this modification considerably changes the results of the model, it shows that subjective rewards are not always credible.

My findings have significant implications for the use of a balanced scorecard for incentive purposes. First, the analogy between balance and congruity emphasizes the role of the firm's *strategy* with regard to incentive purposes. Only if the strategy is properly reflected by the scorecard measures may an alignment of interests be achieved. Perfect alignment, however, requires either that all scorecard measures can be used in an explicit contract or that the principal's discount rate is not too high, so that a relational contract can be applied. This contradicts Kaplan and Norton's [1996] postulate of subjective rewards, and poses a challenge to the alignment of interests in cases where the balanced scorecard is not entirely contractible.

The balance of performance measures has already been studied in accounting research related to agency theory. Datar, Kulp, and Lambert [2001] show that a performance measure with high sensitivity is not necessarily assigned a high weight by the optimal contract, even if it is perfectly congruent with the firm's outcome. Properties of the balanced scorecard are also considered by Dikolli and Kulp [2002], who consider an incentive contract based on multiple, interrelated measures. Smith [2002] studies the weight put on a nonfinancial performance measure in an incentive contract if the agent is able to shift credit between financial and nonfinancial performance. Dikolli [2001] analyzes strategic performance measures with respect to the agent's employment horizon. In his quasi-two-period model, a nonfinancial performance measure becomes valuable as a leading indicator of future financial success because the principal and agent discount future success at different rates.

All of these papers focus on the tradeoff between risk and congruity. In Dikolli's [2001] model, for example, under risk neutrality the agent's discount factor (if different from zero) is readily adjusted to the principal's objective. Risk considerations make the nonfinancial performance measure a valuable tool, even if principal and agent discount at an equal rate.

The present paper abstracts from risk-sharing issues in order to work out another aspect, namely, the interplay of verifiable and nonverifiable performance measures in the case of relational contracts. Since in this respect the paper refers to the analysis of Baker, Gibbons, and Murphy [1994], it is straightforward to use the same set of basic assumptions in order to delineate clearly the central results of the paper. Like most of the papers on relational contracts, Baker, Gibbons, and Murphy [1994] assume that both parties are risk neutral.[2] Under risk neutrality, the properties of the balanced scorecard are more clearly analyzed with respect to the alignment of interests. Moreover, verifiability as a further characteristic of performance measures can be incorporated by extending the analysis to a multiperiod setting, which is the main focus of my analysis. Since in this regard the paper extends and challenges the analysis of Baker, Gibbons, and Murphy [1994] it contributes not only to the literature on strategic performance measurement, but also to that regarding relational contracts.

The remainder of this paper is structured as follows: Section 2 describes the one-period agency framework in which the first research question is analyzed. In section 3, the balanced scorecard is analyzed with regard to performance measure congruity. Section 4 provides a multiperiod extension of the initial agency model and studies the use of subjective rewards. Section 5 concludes and discusses directions for further research. All proofs are given in the appendix.

2. *One-Period Framework*

As a starting point, consider a situation where a principal hires an agent to perform a certain set of tasks once on his behalf. The agent's activity $\mathbf{a} \in \mathbb{R}^n$ has multiple aspects and cannot be legally enforced. It includes not only actions that have immediate financial consequences, but also activities that mainly affect future cash flow in the long term. The principal's objective is to maximize a certain value V net of wage payments. V is assumed to be a linear function $V(\mathbf{a}) = \mathbf{d}'\mathbf{a}$ of the agent's action $\mathbf{a}$, where $\mathbf{d} = (d_i, \ldots, d_n)'$ is the vector of each activity's marginal product. To introduce the problem of distorted performance measurement, V is assumed to be not contractible.

Since neither $\mathbf{a}$ nor the principal's objective V can be contracted on, the principal has to apply some other performance measure for incentive compensation. When describing the balanced scorecard, the set of measures $\mathbf{P} = (P_1, \ldots, P_m)'$ is assumed to be multidimensional. Some measures represent the current period's profit, while others are leading indicators of future profits. Each measure may be influenced by each of the agent's actions. Under the assumption of constant marginal products, P_j is defined as $P_j(a_1, \ldots, a_n, \epsilon_j) = \mathbf{y}_j'\mathbf{a} + \epsilon_j$, with marginal products $\mathbf{y}_j = (y_{j1}, \ldots, y_{jn})'$ and a noise term $\epsilon_j \sim N(0, \sigma^2_{\epsilon_j})$. Following the terminology of Banker and

[2] Pearce and Stacchetti [1998] study relational contracts with a risk-averse agent. Contrary to the present paper, they take the principal's gross profit to be contractible, which allows them to concentrate on risk-sharing issues. Without this assumption, the tradeoff of precision, credibility, *and* congruity has to be considered. I concentrate on the last two points, designating the analysis of a risk-averse manager to future research.

Datar [1989], I refer to y_{ji} as the *sensitivity* of performance measure j to action i. Summarizing these relations, the performance measures are described by $\mathbf{P} = \mathbf{Y}'\mathbf{a} + \epsilon$, where $\mathbf{Y} = (\mathbf{y}_1, \ldots, \mathbf{y}_m)$ is the $(n \times m)$ matrix of the sensitivities and $\epsilon = (\epsilon_1, \ldots, \epsilon_m)'$ is the vector of noise terms. Based on $\mathbf{P}$, the principal offers a linear incentive contract $S = s_f + \mathbf{P}'\mathbf{s}$ in which s_f is a fixed salary, and $\mathbf{s} = (s_1, \ldots s_m)'$ is the vector of compensation parameters defining performance-related payments.

By choosing $\mathbf{a}$, the agent incurs a private cost $C(\mathbf{a})$. For computational convenience, I assume C to be of the form $C(\mathbf{a}) = \frac{1}{2}\mathbf{a}'\mathbf{a}$ (cf. Itoh [1991], Feltham and Xie [1994], Baker [2000]).[3] To focus on the implications of performance measure congruity, both principal and agent are assumed to be risk neutral.[4] Thus, the agent's utility under compensation S and action choice $\mathbf{a}$ is given by $U^A(S, \mathbf{a}) = S - C(\mathbf{a})$, and his reservation utility is denoted by U^R.

The principal's contracting problem in this model is identical to that analyzed by Feltham and Xie [1994], except for the agent's risk neutrality. By choosing the contract parameters $s_1, \ldots, s_m$ the principal maximizes the expected total surplus $V(\mathbf{a}) - C(\mathbf{a})$, subject to the the incentive compatibility constraint $\mathbf{a} = \mathbf{Y}'\mathbf{s}$, which is the first-order condition of the agent's action choice problem. The optimal contract parameters are given by Feltham and Xie [1994, p. 433]:

$$\mathbf{s^0} = (\mathbf{Y}'\mathbf{Y})^{-1}\mathbf{Y}'\mathbf{d}, \tag{1}$$

resulting in a second-best effort $\mathbf{a}^0 = \mathbf{Y}\mathbf{s}^0$ and total profit

$$\Pi^0 = \Pi(\mathbf{s^0}) = \frac{1}{2}\mathbf{d}'\mathbf{Y}(\mathbf{Y}'\mathbf{Y})^{-1}\mathbf{Y}'\mathbf{d}. \tag{2}$$

Total profit is identical to $(\mathbf{y}^0)'\mathbf{y}^0/2$, where $\mathbf{y}^0 = \mathbf{Y}\mathbf{s}^0$ is the sensitivity of the aggregate performance measure $\mathbf{P}^0 = \mathbf{P}\mathbf{s}^0$. Among all possible aggregates $\mathbf{P}\mathbf{s}$, $\mathbf{y}^0$ describes that measure that is "closest" to V. This closeness has been referred to as the *congruity* of a performance measure, and several metrics of congruity have been suggested (Feltham and Xie [1994], Feltham and Wu [2000], Datar, Kulp, and Lambert [2001]). I refer to Baker [2000, 2002], who proposes the cosine of $\mathbf{d}$ and $\mathbf{y}$ as a metric of congruity for a single performance measure with sensitivities $\mathbf{y}$. A related index is derived by simply comparing the second-best total surplus, equation (2), to the total surplus

[3] More generally, one can consider C to have the quadratic form $\mathbf{a}'\mathbf{C}\mathbf{a}$, which accounts for potential complementarities or substitutabilities, as studied by Dikolli and Kulp [2002]. The resulting effects, however, can also be captured by a redefinition of tasks resulting in modified marginal products $\mathbf{d}$.

[4] Of course, assuming a risk-neutral agent provokes the standard argument of selling the firm to the manager. I rule out this alternative because V does not necessarily need to be the value of the firm as a whole, but may also represent the manager's *contribution* to firm value, which cannot be separated from the remaining assets and sold to the manager.

$\Pi^{FB} = \frac{1}{2}\mathbf{d}'\mathbf{d}$ (see Feltham and Xie [1994, p. 433]) under first-best conditions. Relating $\Pi(\mathbf{s^0})$ to Π^{FB}, congruity can be defined as

$$\phi(\mathbf{d}, \mathbf{Y}) = \frac{\Pi(\mathbf{s^0})}{\Pi^{FB}} = \frac{\mathbf{d}'\mathbf{Y}(\mathbf{Y}'\mathbf{Y})^{-1}\mathbf{Y}'\mathbf{d}}{\mathbf{d}'\mathbf{d}} = \frac{\mathbf{d}'\mathbf{y}^0}{\mathbf{d}'\mathbf{d}}. \tag{3}$$

The congruity index $\phi(\mathbf{d}, \mathbf{Y})$ is known as the Rayleigh quotient (cf. Johnson, Riess, and Arnold [1993, p. 433]). It is identical to the squared cosine of the angle between $\mathbf{d}$ and $\mathbf{y}^0$. It generalizes Baker's [2000, 2002] measure to multiple performance measures, and by squaring scales it to the unit interval. Moreover, from equation (3), it is apparent that the second-best total profit $\Pi(\mathbf{s}^0) = \phi\Pi^{FB}$ is a linear function of the congruity. This greatly facilitates the notation of the multiperiod model analyzed in section 4.

3. Properties of the Balanced Scorecard

So far, I have not considered any specific properties of strategic performance measurement systems. For the balanced scorecard, however, Kaplan and Norton [1992] propose a specific structure of measures: four "perspectives" in which particular properties of measures can be derived from the firm's *strategy*. For my purposes, the most interesting of these properties is the attribute "balanced." This is described by Kaplan and Norton [1992, p. 73] as follows: "By forcing senior managers to consider all the important operational measures together, the balanced scorecard lets them see whether improvement in one area may have been achieved at the expense of another." The logic behind this statement is that the firm's strategy links all the different perspectives and their measures, thereby eventually characterizing the firm's long-term objective. By this means actions that are not value-adding, but improve the measured performance, can be ruled out.

To analyze the economic consequences of balance in the present agency framework, I define it as a property that ensures that no actions exist that affect only one performance measure without also affecting V:

DEFINITION (BALANCE). *A system* $\mathbf{P} = \mathbf{Ya} + \epsilon$ *of performance measures is balanced with respect to an objective* $V = \mathbf{d}'\mathbf{a}$ *if each variation* $\boldsymbol{\alpha}$ *of* $\mathbf{a}$ *which changes one performance measure affects at least one other measure or the principal's objective:* $\mathbf{d}'\boldsymbol{\alpha} = 0 \wedge \mathbf{y}_j'\boldsymbol{\alpha} = 0 \ \forall\, j \in \{1, \ldots, m\}\backslash\{k\} \Rightarrow \mathbf{y}_k'\boldsymbol{\alpha} = 0.$

With this property, the scorecard cannot indicate an *overall* improvement if the firm's objective is not affected. For incentive purposes, this implies that under a properly designed incentive contract, a manager should not be able to *game* his performance score, that is, he cannot enhance his remuneration without simultaneously affecting the principal's objective V.

Using this definition, balance should be obtainable by simply keeping track of *all* numbers that might affect the firm's objective. However, the resulting system of performance measures is hardly manageable for either decision making or incentive contracting. Therefore, another postulate is that the balanced scorecard should be restrained to a *minimal* number of relevant measures. Only these *strategic* measures (in contrast to *diagnostic* measures, which are only used to detect potential malfunctions) ought to be used for strategic planning and control. With regard to the present model, a minimum requirement for this property is that no measure's effort impact, $\mathbf{y}_j'\mathbf{a}$, can be derived from other measures, that is, the vectors of marginal products have to be linearly independent:[5]

DEFINITION (MINIMALITY). *A system* **P** *of performance measures is* minimal *if the matrix* **Y** *has full rank m.*

An immediate consequence of minimality is that the number, m, of performance measures must not exceed the number, n, of actions taken by the agent. Furthermore, for $m = n$ the first-best solution can be obtained because in this case the scorecard is capable of inducing *any* effort profile the principal desires. In general, it seems realistic to assume that $m < n$ and therefore the first-best effort in general cannot be induced. However, since the balanced scorecard is designed to provide a balanced picture of the firm's long-term objectives, it seems plausible that its measures may nevertheless be capable of inducing the first-best effort levels. Using the formal definition of balance, the following conjecture can be proved:

PROPOSITION 1. *Let* **P** *be balanced with respect to the principal's objective V. Then the following statements hold:*

(i) The first-best solution can be achieved by a linear contract based on **P**.
(ii) If **P** *is also minimal, all measures have nonzero value in the optimal contract.*

Proposition 1 shows how the proposed properties of a balanced scorecard are related to performance measure congruity. Balance is sufficient for the implementation of first-best effort levels, whereas minimality additionally guarantees that no measure is superfluous for contracting. First best is obtained by representing the firm's objective V by the scorecard measures. This is always possible if the scorecard is balanced.

Thus, as a first implication, I find no theoretical justification for Kaplan and Norton's [1996] concern that formula-based contracts may provide distorted incentives. While such contracts permit "substantial incentive compensation to be paid if the business unit overachieves on a few objectives even if it falls short on others" (Kaplan and Norton [1996, p. 82]), discrepancies of pay and performance only occur by chance, not as a result of effort misallocation—provided an optimal contract is applied and based on a balanced scorecard. Consequently, in order to find support for their assertion, the assumptions of the previous model have to be scrutinized. The following section studies conditions under which their objection might prevail, and provides resources to overcome the resulting problems.

---

[5] This condition also justifies the formal derivation of the second-best contract in equation (1), where I assumed that $\mathbf{Y}'\mathbf{Y}$ is invertible. A sufficient condition for this is that **Y** has full rank (See Berck and Sydsaeter [1993, p. 107]).

4. *The Use of Subjective Rewards under a Balanced Scorecard*

4.1 VERIFIABILITY OF PERFORMANCE MEASURES IN EXPLICIT CONTRACTS

Part (ii) of Proposition 1 makes it clear that every measure of a minimal and balanced scorecard is needed to induce the desired effort levels. This becomes important if we relax the assumption that all measures can be used in an explicit incentive contract. A necessary condition for such contracts to be applicable is that the respective performance information is not only observable to the contracting parties, but can also be verified by an outside arbiter, such as a court, to prevent the principal from refusing the wage payment ex post.

For financial performance measures, verifiability is ensured by accounting standards. Nonfinancial measures lack such legal foundations and thus from the outset cannot be assumed to be contractible. Admittedly, strategic performance measurement requires substantiating corporate goals by quantifiable measures. Even if the particular measures are defined precisely, however, they are often based on procedures such as customer or employee surveys that are not regulated by law as financial transactions are. For example, a study quoted by Kaplan and Norton [2001, p. 254] finds that companies using the balanced scorecard for incentive purposes experienced particular difficulties in quantifying measures based on the *customer* and *learning and growth* perspectives. Such measures also varied significantly across the companies. It therefore seems plausible that at least some of the nonfinancial measures in **P** cannot explicitly be contracted on. This is an important limitation of contracting, because according to Proposition 1 (ii) the optimal contract puts nonzero weight on *each* performance measure of a balanced scorecard. Thus, if not all scorecard measures can be included in the contract, the remaining measures will not be perfectly congruent:

COROLLARY 1. *Let the performance measurement system* **P** *be minimal and balanced. If not all measures of* **P** *are verifiable, the first-best effort cannot be induced by an explicit contract.*

Apart from this qualitative result, Proposition 1 (ii) also implies that each measure of a balanced and minimal scorecard has a positive impact on congruity, which is therefore strictly increasing in the number of verifiable measures. The precise impact of each measure on congruity, however, may vary in size and may even depend on which other measures are verifiable. Thus, while profit is a linear function of congruity, congruity in general is not linear in the number of verifiable performance measures.

Regardless of the extent of the resulting losses, the general question arising from the deficiency of a partly subjective scorecard is whether the nonverifiable part of the information can be used for incentive purposes in a different way. Kaplan and Norton [1996, p. 82] argue that by carefully discussing the balanced scorecard results with the manager, additional information on the drivers of performance can be obtained. Since such a dialogue rarely results in verifiable data, however, they propose the use of discretionary rewards.[6]

In the economics literature, subjective rewards have mainly been studied as an example of relational contracts. These "informal agreements and unwritten codes of conduct" (Baker, Gibbons, and Murphy [2002]) can be based on outcomes that are observed ex post by the contracting parties alone. At the same time, they have to be self-enforcing: It must be in the parties' best interest to keep the agreement. Since in a one-shot relationship the principal always gains from refusing any voluntary payment, relational contracts are usually analyzed in a dynamic setting where the two parties repeatedly enter into a contract.[7] The next section applies this framework to analyze how subjective rewards can be used to improve a formal contract.

4.2 SUBJECTIVE PERFORMANCE EVALUATION IN A MULTIPERIOD AGENCY

The following analysis is related to Baker, Gibbons, and Murphy [1994], who apply Bull's [1987] model of an infinitely repeated agency relationship to the question of how formal and relational contracts can optimally be combined. In their model, the subjective assessment of firm value is complemented by a verifiable performance measure, the marginal product of which is privately observed by the agent after contracting. For both measures, a payment is determined ex ante: an enforceable, formal contract for the verifiable, yet distorted measure, and a relational contract for the nonverifiable, subjective evaluation. Under the assumption that both parties apply Grim-trigger strategies in the repeated game,[8] the latter is self-enforcing if the principal's future benefits from an ongoing relational contract exceed the amount he could save by once refusing the payment of the implicit contract.

[6] As evidence of this practice, Global Financial Services replaced its formula-based bonus system by a system that also contains qualitative evaluations (Ittner, Larcker, and Meyer [2003, 726f.]). Beyond this example, the empirical study of Gibbs et al. [2004, p. 429] supports that in departments with complex tasks, subjective bonuses are used to mitigate distortions caused by objective performance measures that are incomplete, short-term focussed, or susceptible to manipulation. More generally, Hayes and Schaefer [2000] find empirical evidence that firms use discretionary rewards if observable measures are less predictive for future returns. This confirms findings by Bushman, Indjejikian, and Smith [1996, p. 189] that individual performance evaluation increases with the length of product development and product life cycle.

[7] There are only a few articles that analyze the use of subjective performance evaluation in a static setting. In this environment, credibility can be achieved by involvement of a third party (MacLeod [2003]) or by the use of bonus pools (Baiman and Rajan [1995], Rajan and Reichelstein [2006]).

[8] Roughly speaking, this is the case if both parties employ cooperative strategies as long as no party has defected, but never forgive a party once it has defected. As can be deduced from the analysis of Abreu [1988], this assumption can be made without loss of generality for the following results.

Consequently, the credibility of subjective rewards depends critically on the principal's fallback position after a potential defection. If the latter is a formal contract, the surplus from this contract determines the credibility of the bonus payment. Baker, Gibbons, and Murphy [1994] show that a sufficiently congruent objective performance measure can completely rule out the use of a subjective performance evaluation. This suggests that in the present model a high congruity of the verifiable performance measures could hinder subjective rewards, eventually precluding first-best if not all scorecard measures are verifiable.

To prove this conjecture, I translate the analysis of Baker, Gibbons, and Murphy [1994] to the multitask model with multiple performance measures. To this purpose, I assume that the one-shot game of section 2 is repeated infinitely often. The verifiable performance indicators are used in a formal contract, whereas the entire scorecard can be applied in a relational contract. To account for Kaplan and Norton's [1996, p. 82] proposal of using threshold levels in compensation agreements, the latter is assumed to be a bonus contract.[9] If the scorecard is balanced, by Proposition 1 this bonus contract can be based on a performance score that is perfectly congruent to the principal's objective, as is assumed by Baker, Gibbons, and Murphy [1994]. While the nonverifiable firm value in the analysis of Baker, Gibbons, and Murphy [1994] is either one or zero, in this work it can take on arbitrary values in $\mathbb{R}$. The relational contract in the present model therefore has to fix not only the bonus payment, but also a threshold of performance for which it is due. The agent's probability of receiving the bonus depends on this level as well as on the distribution of the performance score. Consequently, the credibility of a bonus payment is also subject to the precision of the performance measure.

Since the focus of the present analysis is on performance measure congruity rather than precision, I henceforth analyze the benchmark case in which the principal attains unverifiable, precise information on the agent's impact, $\hat{P}_j = \mathbf{y}_j'\mathbf{a}$, on each performance measure P_j. This is an extreme interpretation of Kaplan's and Norton's [1996] opinion that a discussion of the scorecard results provides an improved understanding of the manager's performance. It turns out that even in this optimistic scenario, the first-best solution is not always achievable. The consequences of noisy performance measurement are discussed shortly, in subsection 4.4.

4.3 THE USE OF SUBJECTIVE REWARDS UNDER PRECISE PERFORMANCE EVALUATION

If a perfectly congruent and precise system of performance measures is available, the value added V is de facto observable. Without loss of generality,

[9] Of course, a linear contract can be considered instead. This, however, reduces the credibility of the relational contract. I comment on this issue when I compare my results to that of Baker, Gibbons, and Murphy [1994].

it can therefore be assumed that a relational contract can be based on V. A formal contract, in contrast, can only be written on the verifiable part of the scorecard. To distinguish the set of $l(l \leq m)$ verifiable measures from the whole set, $\mathbf{P}$, of scorecard measures, I denote them by $\mathbf{P}_v$. The verifiable measures can be described by the $(n \times l)$ matrix $\mathbf{Y}_v$ of sensitivities and the vector ϵ_v of noise terms. According to the analysis of section 2, the second-best formal contract is based on an aggregate performance score $\mathbf{y}^0 = \mathbf{Y}_v(\mathbf{Y}_v'\mathbf{Y}_v)^{-1}\mathbf{Y}_v'\mathbf{d}$, resulting in a total profit of $\Pi^0 = (\mathbf{y}^0)'\mathbf{y}^0/2 = \phi\Pi^{FB}$, where for notational convenience the congruity $\phi(\mathbf{d}, \mathbf{Y}_v)$ is simply denoted by ϕ. In the following analysis, ϕ is critical to the principal's fallback position after a potential defection, which can be either a purely explicit contract or no production at all. A purely explicit contract is beneficial to the principal if his net utility, $U^0 = \phi\Pi^{FB} - U^R$, from that contract is nonnegative, that is,

$$\phi \geq \frac{U_R}{\Pi^{FB}} \equiv \hat{\phi}. \tag{4}$$

The congruity ϕ also determines the effectiveness of a formal contract when used in combination with subjective rewards. To distinguish the two effects, I first derive conditions under which the first-best solution can be obtained by purely subjective rewards. Subsequently, it is shown that first best can also be implemented by a combination of subjective rewards and a formal contract. Finally, I examine how such a hybrid contract can be used even if the first-best solution cannot be obtained.

Using V as a performance score, first best can be implemented by simply prescribing the first-best gross profit $V^{FB} = \mathbf{d}'\mathbf{a}^{FB}$ and a bonus payment, B, to be paid if the realized profit accords to that target. The least costly way for the agent to meet this target is to choose the first-best effort $\mathbf{a}^{FB} = \mathbf{d}$. To prevent the agent from not working at all, the bonus payment at least has to compensate him for the resulting cost of effort, that is,

$$B \geq C(\mathbf{a}^{FB}) = \frac{1}{2}\mathbf{d}'\mathbf{d} = \bar{B}. \tag{5}$$

To keep the relational contract credible, the bonus payment, B, should be kept to a minimum. Therefore, $B = \bar{B}$, and an additional salary is required to provide the agent's reservation utility. Thus in total, the principal offers a contract, $S = s_f + B$, in which $s_f = U^R$ and the bonus, $B = \bar{B}$, is promised for $V \geq V^{FB}$. The principal's net profit from this contract is $U^{FB} = \Pi^{FB} - U^R$.

Unfortunately, this contract need not be credible. Since the bonus is based on nonverifiable data, the principal ex post could refuse to pay. To prevent him from doing so, the payable bonus $\bar{B}$ must not exceed the net present value of the benefits the principal realizes under an ongoing contract. This benefit depends on the principal's fallback position. It can be calculated by comparing the principal's utility U^{FB} derived from an ongoing first-best relational contract to either his utility U^0 from a purely explicit contract or his utility 0 if he ceases production. Its net present value under a discount rate r is $\sum_{t=1}^{\infty}[U^{FB} - U^0](1+r)^{-t} = [\Pi^{FB} - \Pi^0]/r$ for a formal contract and $\sum_{t=1}^{\infty} U^{FB}(1+r)^{-t} = [\Pi^{FB} - U^R]/r$ for ceasing production. Comparing these values to the payable bonus $\bar{B}$ provides a threshold level of the principal's discount rate for the first-best solution to be implementable:

PROPOSITION 2. *The first-best solution can be obtained by subjective rewards based on a balanced scorecard if the principal's discount rate is smaller than a threshold level*

$$\bar{r} = \begin{cases} 1-\phi & \text{if} \quad \phi \geq \hat{\phi} \\ 1-\hat{\phi} & \text{if} \quad \phi < \hat{\phi}. \end{cases} \tag{6}$$

Proposition 2 is consistent with the results of Baker, Gibbons, and Murphy [1994]. It shows that in the case of a valuable formal contract ($\phi \geq \hat{\phi}$), a high congruity of the verifiable performance measures constricts a purely relational contract by improving the principal's periodical profit $U^0 = \phi\Pi^{FB} - U^R$ after a potential defection. This diminishes the principal's future benefits from the relational contract, which in turn reduces the rate by which this benefit may be discounted. If the congruity exceeds a critical level, $\bar{\phi} \equiv 1 - r$, the principal cannot credibly promise to pay the required bonus because his fallback position is too good.

In the case that a purely explicit contract is not valuable ($\phi < \hat{\phi}$), the feasibility of the first-best solution depends on the agent's reservation utility $U^R = \hat{\phi}\Pi^{FB}$ instead of $\Pi^0 = \phi\Pi^{FB}$. In the critical discount rate, ϕ is therefore replaced by its threshold level $\hat{\phi}$ for which the explicit contract becomes valuable. Thus, in this case, credibility is independent of congruity.

In both cases, a sufficiently impatient principal cannot credibly promise to compensate the agent only by a bonus payment because the bonus is too high. An obvious question is whether first best can still be obtained if subjective rewards are complemented by an explicit contract, thereby reducing the voluntary payment. Under such a *hybrid* contract, the agent's total compensation would take the form $S = s_f + B + \mathbf{P}_v\mathbf{s}$. B again denotes the voluntary bonus payment that is due if a certain performance target is realized, and $\mathbf{s}$ is the vector of contract parameters determining the agent's share of the verifiable performance measures $\mathbf{P}_v$. Obviously, if the target is defined by means of the principal's objective V, the first-best effort cannot be induced. The agent strives for the bonus as well as for the payment from the formal contract. His optimal effort is thus a mixture of the first-best action and the one resulting from a purely formal contract. Since the latter by assumption cannot induce the first-best effort, the action inevitably departs from the desired profile.

Since contrary to the model of Baker, Gibbons, and Murphy [1994] subjective rewards can be based not only on V as a whole, but also on only a subset of the scorecard measures, explicit incentives can be complemented by subjective rewards in a way that implements the overall first-best action. To this purpose, the relational contract has to be based on a performance index comprised of all the desired effects of the agent's action that are not covered by verifiable scorecard measures. This index is denoted $\hat{\mathbf{P}}\mathbf{s}$, where $\mathbf{s}$ is chosen to ensure that the agent chooses the first-best effort, $\mathbf{d}$. The existence of such contract parameters is guaranteed by the balance property. Compared with a purely relational contract, the required bonus payment is reduced, and first best can be implemented for a wider range of discount rates:

PROPOSITION 3. *The first-best solution is obtained by a hybrid contract based on a balanced scorecard if the principal's discount rate is smaller than a threshold level*

$$\check{r} = \begin{cases} 1 & \text{if} \quad \phi \geq \hat{\phi} \\ \dfrac{1-\hat{\phi}}{1-\phi} & \text{if} \quad \phi < \hat{\phi}. \end{cases} \tag{7}$$

The proof of Proposition 3 shows that the optimal hybrid contract for the first-best action includes a formal contract identical to the one applied without subjective rewards. Thus, if the agent abandons the bonus, he chooses the second-best effort profile, $\mathbf{a}^0 = \mathbf{y}^0$, which—as the analysis of section 2 shows—is closest to the first-best effort, $\mathbf{d}$, which is sought to be induced. By this means, the effort cost that has to be covered by the relational contract is kept to a minimum.

Proposition 3 is in stark contrast to the results of Baker, Gibbons, and Murphy [1994] because it shows that under a hybrid contract, congruity of the verifiable measures does not constrain the use of subjective rewards, even if the principal's fallback position is a formal contract. This new result is due to the fact that subjective rewards in the present model are not necessarily based on the principal's objective, V, in contrast to the model of Baker, Gibbons, and Murphy [1994]. By restricting subjective rewards to that portion of the first-best action profile that cannot be induced by the explicit contract, the relational contract only has to compensate the cost of the additional effort $\mathbf{a}^{FB} - \mathbf{a}^0 = \mathbf{d} - \mathbf{y}^0$. This amount can be expressed in terms of congruity, because ϕ measures how well $\mathbf{y}^0$ matches $\mathbf{d}$. Since a fraction ϕ of first-best incentives is already provided by the formal contract, the required bonus given in Proposition 2 is reduced by the factor $1 - \phi$ compared to a purely subjective reward. The critical discount rate thus increases by the factor $1/(1 - \phi)$, which is exactly reflected by equation (7), as opposed to equation (6). This comparison also demonstrates the different role that congruity plays in the contract types of Propositions 2 and 3: While the critical discount rate, $\bar{r}$, for a purely relational contract depends on congruity only if the principal's fallback position is a formal contract, it now varies in ϕ if that contract is not valuable.

The differences between the hybrid contract and a purely relational contract can be clarified by the following example:[10]

---

[10] For ease of exposition, this example assumes that the number of performance measures equals the number of activities.

EXAMPLE. Assume $n = m = 2$ with $\mathbf{d} = (1, 1)'$ and $\mathbf{Y} = \left(\begin{smallmatrix} 2 & 0 \\ 1 & 1 \end{smallmatrix}\right)$. By the full rank of $\mathbf{Y}$, the system $\mathbf{P}$ of performance measures is minimal and balanced. If all measures were contractible, the first-best action $\mathbf{d}$ could be obtained by a pure formal contract $S = s_f + \mathbf{Ps}$ with $\mathbf{s} = (1/2, 1/2)$, rendering a total profit of $\Pi^{FB} = 1$. Now assume that P_2 is not contractible. The optimal pure explicit contract based on P_1 takes the form $S^0 = s_f + P_1 s_1$ with $s_1 = 3/5$, resulting in an action $\mathbf{a}^0 = \mathbf{y}^0 = (6/5, 3/5)$. The total profit is $\Pi^0 = 9/10$, and the congruity is $\phi = 9/10$.

The first-best action might be induced by a pure relational contract. The required bonus equals the agent's cost of effort $C(\mathbf{a}^{FB}) = 1$, which is credible for discount rates $r \leq 1 - \phi = 0.1$. If subjective rewards are accomplished by a formal contract, the bonus is dramatically reduced. The optimal hybrid contract consists of the second-best formal contract S^0 and a bonus for achieving a target of 1/5 for the subjective performance index $\hat{\mathbf{P}}\mathbf{s} = -(1/10)\hat{P}_1 + (1/2)\hat{P}_2$, which reflects the gap $\mathbf{d} - \mathbf{y}^0$ between the first-best and the second-best action. Waiving the bonus, the agent would choose $\mathbf{a}^0$, realizing an expected payment of $s_f + 9/5$ and incurring a cost of $C(\mathbf{a}^0) = 9/10$. If he chooses $\mathbf{a}^{FB}$, his expected payment is $s_f + B + 9/5$ and his cost of effort is $C(\mathbf{a}^{FB}) = 1$. To induce the first-best effort, the bonus therefore has to be $C(\mathbf{a}^{FB}) - C(\mathbf{a}^0) = 1/10$, which is credible for $r \leq 1$. □

This example is illustrated in figure 1.

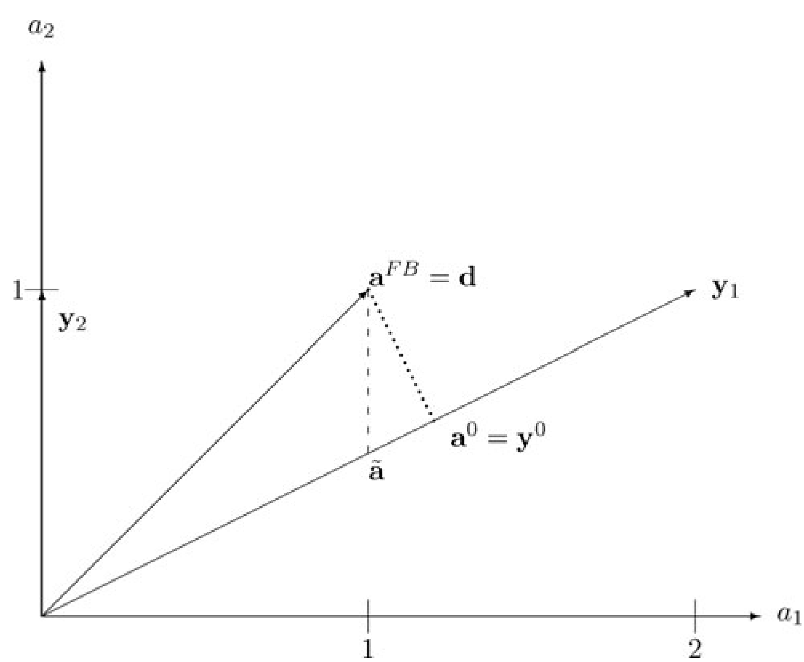

FIG. 1.—Actions resulting from formal and relational contracts. A pure formal contract induces the action $\mathbf{a}^0 = \mathbf{y}^0$. To induce the first-best action $\mathbf{a}^{FB} = \mathbf{d}$ by a hybrid contract, the relational contract has to incentivize the gap $\mathbf{a}^{FB} - \mathbf{a}^0$ (the dotted line). If the relational contract is based only on P_2, it has to close a larger gap, $\mathbf{a}^{FB} - \tilde{\mathbf{a}}$ (the dashed line).

The action $\mathbf{a}^0$ resulting from the formal contract minimizes the required bonus payment $C(\mathbf{a}^{FB}) - C(\mathbf{a}^0)$. To that purpose, the level $a_1^0 = 6/5$ of the first action has to exceed its first-best level of 1. In order to adjust this, the subjective performance index also has to incorporate the verifiable performance measure $\hat{P}^1$. This is counterintuitive because incentives based on P_1 can also be provided by a formal contract. What matters, however, is not the single performance measure, but its optimal combination in both parts of the contract. If the relational contract is entirely based on P_2, first best can only be obtained by a formal contract $S = s_f + P_1/2$ and a target of 1/2 for the subjective performance index $\hat{P}_2$. Since the action $\tilde{\mathbf{a}} = (1, 1/2)$ from the formal contract differs from $\mathbf{d}$ more than $\mathbf{a}^0$ does, the required bonus increases to 1/8.

For $r > \max\{1, (1-\hat{\phi})/(1-\phi)\}$, even a hybrid contract cannot credibly provide incentives for the first-best effort. However, a hybrid contract may still provide a better solution than a purely explicit contract. In this regard, the question arises whether the properties of the combined contract remain valid. Furthermore, it is of interest to determine the level of congruity for which the hybrid contract is credible. Since the answers to these questions are interrelated, they are combined into the following proposition:

PROPOSITION 4. *The optimal hybrid contract based on a balanced scorecard has the following properties:*

(i) The explicit contract is identical to the optimal pure explicit contract.
(ii) The induced action is a convex combination of the first-best action and the action induced by a pure explicit contract.
(iii) A credible hybrid contract exists if the principal's discount rate is smaller than a threshold level

$$\tilde{r} = \begin{cases} \infty & \text{if} \quad \phi \geq \hat{\phi} \\ \dfrac{1-\hat{\phi}}{\hat{\phi}-\phi} & \text{if} \quad \phi < \hat{\phi}. \end{cases} \tag{8}$$

The first part of the proposition is intuitively derived from the previous results. As in the analysis of the first-best situation of Proposition 3, the formal contract is identical to the one offered without subjective rewards. This minimizes the bonus necessary to achieve a certain target of the principal's objective V. Since due to the lack of credibility the first-best effort cannot be induced, the principal aims to approximate this solution as closely as possible. Building on the explicit contract, this is best done by offering subjective rewards for the same performance measure as in the first-best hybrid contract, but with a less ambitious target. The relational contract then provides inducement for as much of the gap $\mathbf{d} - \mathbf{y}^0$ to the first-best action as can credibly be rewarded by a voluntary payment.

When credibility becomes important, effort is therefore continuously shifted from its first-best allocation to the second-best allocation under a purely explicit contract. This adjustment also reduces the bonus required to make the agent accept the implicit part of the contract. Since the bonus only has to compensate the agent for effort in excess of that induced by the explicit contract, it can become arbitrarily small for the convex combination described here. A small bonus is always credible, however, as long as the principal is not infinitely impatient. If a purely formal contract is beneficial, then a hybrid contract must be valuable. If a formal contract on its own is not profitable for the principal, however, a small bonus may not suffice to facilitate the combined contract. To this end, the difference $U^R - \Pi^0 = (\hat{\phi} - \phi)\Pi^{FB}$ between the necessary total profit U^R and the profit Π^0 from a pure formal contract has to be covered by subjective rewards. The bonus can only be reduced to a fraction, $\hat{\phi} - \phi$, of that required in the case of a purely subjective reward, as studied in Proposition 2. Consequently, there exists a critical discount rate up to which the contract is credible. Compared to a purely subjective reward, this threshold is increased by a factor $1/(\hat{\phi} - \phi)$, as reflected in equation (8), which may be contrasted to equation (6).

EXAMPLE—CONT. Assume that in the above example the principal's discount rate is $r = 2$. If, for example, $U^R = \hat{\phi} = 3/5 < \phi = 9/10$, the principal's fallback position is a formal contract. The optimal hybrid contract combines the explicit contract S^0 with a target of 2/15 for the subjective performance index $\hat{P}\mathbf{s}$, 2/3 of the target for first best. The agent's action is $\mathbf{a}^0 + 2/3(\mathbf{d} - \mathbf{a}^0) = (16/15, 13/15)$. The payable bonus 4/90 is $(2/3)^2 = 4/9$ of the bonus required for the first-best effort, whereas the benefit 8/90 from the hybrid contract is 8/9 of its amount under the first-best action. Thus, the voluntary payment is credible for a discount rate twice as high as the threshold $r = 1$ for first best.

If, on the other hand, we take $U^R = 39/40 > \phi$, a purely explicit contract is not valuable. Compared with the first situation, the profit generated by subjective rewards first has to cover the difference, $39/40 - 9/10 = 3/40$, between Π^0 and U^R. Beyond this the above contract renders a benefit of only $8/90 - 3/40 = 5/360$, the relational contract being credible for $r \leq 5/16$. Even a contract modification does not generate a credible hybrid contract. The best the principal can do is to combine the formal contract with a target of 3/40 for the subjective performance index, inducing an action $\mathbf{a}^0 + 3/4$ $(\mathbf{d} - \mathbf{a}^0) = (21/20, 18/20)$. The payable bonus is 9/160, and the benefit from the contract is 3/160. This contract is credible for $r \leq 1/3$. □

Regarding the quality of the verifiable performance measures, the required bonus decreases with the congruity ϕ of verifiable performance measures. If the principal's fallback position is a formal contract, this effect exactly outweighs the improvement of the principal's fallback position, and the existence of a hybrid contract is independent of the congruity ϕ. If a pure formal contract is not profitable, congruity enhances welfare because it no longer improves the fallback position. As in the analysis of Baker, Gibbons, and Murphy [1994], subjective rewards may then create a valuable combination with explicit incentives even though the latter are not beneficial on their own. For this to be the case, the congruity has to exceed a certain threshold $((1 + r)\hat{\phi} - 1)/r$, which is positive if $r > (1 - \hat{\phi})/\hat{\phi} = \check{r}$. For $r < \check{r}$, a credible hybrid contract exists for each level of congruity. For

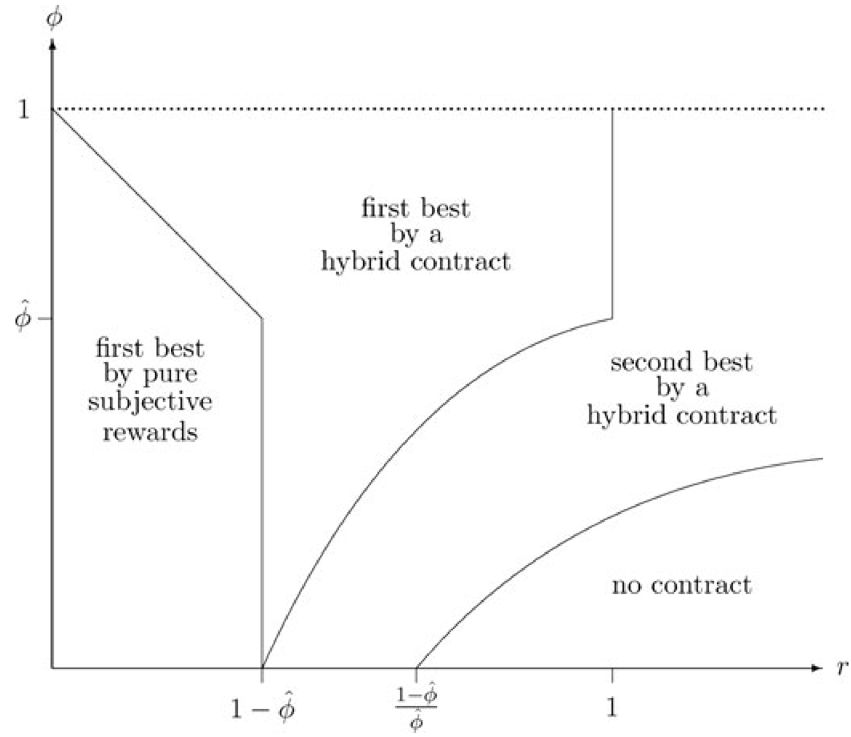

FIG. 2.—Contract type under different combinations of the principal's discount rate r and the congruity ϕ of the verifiable performance measures. The critical congruity $\hat{\phi}$ for which a pure formal contract becomes valuable is assumed to be 0.6. The dotted line indicates the maximum level 1 of congruity. The solid curves delimit the sectors for which the indicated contract types are optimal.

higher values of r, the critical level of congruity increases and approaches $\hat{\phi}$ for an infinitely impatient principal.

The results of Propositions 2 to 4 are illustrated in figure 2.

If the principal is rather patient, first best can be achieved for any congruity ϕ of the verifiable performance measures. If the principal's discount rate is above $1 - \hat{\phi}$, first best can only be obtained if the congruity is high enough. If this is not the case, a second-best hybrid contract results as long as $r \leq (1 - \hat{\phi})/\hat{\phi}$. If the principal's discount exceeds that level, there may be no contract, a second-best or a first-best hybrid contract. First best can no longer be achieved only if $r > 1$.

This variety of scenarios raises the question of how the principal's profit is governed by the congruity of the verifiable scorecard measures. Despite the apparent differences in the type of contract, the answer is unambiguous:

PROPOSITION 5. *The principal's net profit is nondecreasing in ϕ for all levels of the principal's discount rate.*

Proposition 5 partially contradicts the findings of Baker, Gibbons, and Murphy [1994]. While in their analysis a less distorted verifiable performance measure is only beneficial if a pure formal contract is not valuable, congruity in the present model also improves contracting if the principal's fallback position is a formal contract. The difference in these results can be attributed to the more elaborate relational contract assumed in this paper. This can be illustrated by the first-best solution: While in the model of Baker, Gibbons, and Murphy [1994] first best can only be induced by a purely relational contract, which requires a bonus rate of $b = 1$ regardless of the distortion in verifiable performance measures, the required bonus in the present model decreases with the congruity of the verifiable measures. A higher bonus payment makes the agreement less credible, eventually precluding the possibility of implicit incentives.

There are two reasons for this difference in the bonuses. The first is the multiplicity of measures used in the relational contract of this paper. In this regard, improvements to the contract in Baker, Gibbons, and Murphy [1994] require the bonus payment to be tied not only to V, but rather to a combination of V and the verifiable measure P. The second reason is the type of relational contract. Since in Baker, Gibbons, and Murphy [1994] the agent's effort is identical to the probability of achieving the targeted profit, the expected bonus payment is de facto a *linear* function of effort and thus corresponds to a *linear* relational contract in the present multitask setting. A payment $B = \hat{\mathbf{P}}\mathbf{s}$, however, will be double that required for the bonus contract. This is due to the fact that equating marginal cost and marginal compensation under a linear contract inevitably leads to an overcompensation of effort cost by the performance-related payment. For a formal contract, the overcompensation can be balanced by a negative base salary. With regard to the credibility of a relational contract, however, such an adjustment is pointless because only the amount of the variable payment matters. Consequently, a linear relational contract is always less credible than the bonus scheme applied in this paper.

4.4 THE USE OF SUBJECTIVE REWARDS UNDER NOISY PERFORMANCE EVALUATION

In the preceding analysis the principal can offer a forcing relational contract because all performance measures are assessed without noise. If a relational contract is based on a noisy performance index, however, a forcing contract is not feasible because the performance targets are met only with a certain probability. Even if the agent chooses his effort according to the agreement, it is not certain that he will receive the bonus. On the other hand, he may be rewarded even though he has not provided the required effort. When choosing his action, he therefore trades off his cost of effort against the *expected* bonus payment.

This reaction significantly changes the size of the required bonus. If the agent's action choice can be characterized by first-order conditions,[11] then the bonus is determined by the agent's marginal cost and the performance measure's sensitivity to the agent's choice of action (Demougin and Fluet [1998, p. 627]). For normally distributed performance measures, the latter is determined by the variance of the applied performance index. Consequently, in general, the bonus required under a noisy performance evaluation is a function of the performance measure's congruity and also its precision.

I do not analyze this issue in detail here. Based on the above considerations, however, its main consequences may be projected. A generalization of Proposition 2 concerning the use of a purely relational contract should be straightforward. Intuitively, the critical discount rate under noisy performance measurement depends on both the congruity and the precision of the aggregate performance index. If a purely relational contract is not credible, however, the analysis gets more complex.

In this case, the principal has to decide which aggregate performance index is used in the relational contract. Under noisy performance evaluation, this decision is influenced by risk considerations. A performance measure pointing to the first-best action, as in Proposition 3 and 4, is hardly optimal. Due to the lack of precision, such a measure might well result in a high bonus payment. The principal is likely better off choosing a less congruent but more precise performance measure. The optimal contract, of course, accounts for both congruity and precision. To maximize the benefits from the combined contract, the principal should seek the least distorted performance index for the formal contract. To minimize the required bonus, on the other hand, it is of interest to diversify the risk through a combination of measures. Since these two objectives hardly point in the same direction, a subtle tradeoff between the benefits of risk diversification and congruity has to be considered. Since a detailed investigation of this tradeoff requires considerable analysis, I leave it to future research at this point.

5. *Conclusion*

Regarding the use of the balanced scorecard for incentive purposes, several conclusions can be drawn from the present analysis. The first refers to Kaplan and Norton's [1996] initial objection to formula-based incentive contracts. Contrary to their concern, such contracts are capable of providing undistorted incentives if the performance measures meet the properties these authors propose for a balanced scorecard. Under these properties,

[11] Under the present assumption of normally distributed error terms, this is the case provided that the performance index is not too precise.

misalignment of interest may only result from an improper aggregation of performance information. If this alignment is not achieved, the present analysis suggests that the firm should check whether the scorecard and the formula-based incentive contract are properly designed rather than setting targets for a whole set of performance measures as suggested by Kaplan and Norton [1996, p. 81].

The second conclusion concerns the use of subjective rewards based on a balanced scorecard. This paper proves that a lack of verifiability in some of the performance measures justifies the application of relational contracts. Probably the most important result in this respect is that the credibility of such contracts can be considerably improved by restricting informal incentives to that part of the first-best action that cannot be induced by a formal contract. This separation of incentives is made feasible only by the balance property of the scorecard. Since both $\mathbf{d}$ (by the balance property) and $\mathbf{y}^0$ (by the second-best formal contract) can be induced by a contract based on all measures of the scorecard, the difference, $\mathbf{d} - \mathbf{y}^0$, can be as well. Unverifiable measures are not "lost" for purposes of incentive compensation, as each of them plays a critical role in the optimal contract. This can be highlighted by the fact that, provided the scorecard is minimal and balanced, *all* unverifiable measures should be used in the relational contract. This is the case even if its combination with the formal contract does not produce a first-best, but only a second-best, solution.

Concerning the standards that might arise for systems like the balanced scorecard, the message of this paper seems unambiguous at first glance. If standards demand that a larger number of scorecard measures deliver "hard" information that can be verified in court, according to Proposition 5 the agency definitely benefits. This contrasts with earlier results of Baker, Gibbons, and Murphy [1994], who assume coarser subjective performance information than that provided by a balanced scorecard, and reinforces the robustness of the balanced scorecard with respect to credibility concerns. However, since the emerging standardized measures need not coincide with those of the initial scorecard, their consideration might "blow up" the scorecard because no other measure can be replaced without violation of the balance property. Alternatively, in terms of credibility, the firm might pass on perfect balance in order to improve the verifiability of a given number of strategic measures.

Directions for future research mainly regard the use of relational contracts based on imprecise performance information. As addressed in the last section, the design of a subjective performance score then becomes a matter of congruity and noise, as is the case for a risk-averse agent. Unlike the latter situation, however, risk in this context does not necessarily impair the surplus of the agency. Risk enters the contracting problem only through the credibility constraint, which need not be binding. Therefore, the proposed model may serve as a suitable device for studying the balance of performance measures, with respect to both their congruity and precision.

APPENDIX

Proof of Proposition 1. By the definition of balance, the following statement holds:

$$\mathbf{d}'\boldsymbol{\alpha} = 0 \wedge \mathbf{y}_j'\boldsymbol{\alpha} = 0 \quad \forall\, j \in \{1, \ldots, m\}\backslash\{k\} \Rightarrow \mathbf{y}_k'\boldsymbol{\alpha} = 0.$$

The left part of the implication describes $\boldsymbol{\alpha}$, which represents the elements of the nullspace $\mathcal{B}^0$ of the matrix $\mathbf{B} = (\mathbf{d}, \mathbf{y}_1, \ldots, \mathbf{y}_{k-1}, \mathbf{y}_{k+1}, \ldots, \mathbf{y}_m)$. In other words, each $\boldsymbol{\alpha}$ is orthogonal to the subspace $\mathcal{B}$ spanned by $\mathbf{B}$. If $\mathbf{y}_{\mathrm{k}}\boldsymbol{\alpha} = 0$ follows from this for all $\boldsymbol{\alpha} \in \mathcal{B}^0$, then the vector $\mathbf{y}_{\mathrm{k}}$ must be an element of $\mathcal{B}$, that is, there exists a real-valued $(n \times 1)$ vector $\boldsymbol{\lambda} = (\lambda_d, \lambda_1, \ldots, \lambda_{k-1}, \lambda_{k+1}, \ldots, \lambda_n)$ such that $\mathbf{B}\boldsymbol{\lambda} = \mathbf{y}_k$:

$$\mathbf{y}_k = \lambda_d \mathbf{d} + \sum_{j \neq k} \lambda_j \mathbf{y}_j. \tag{A1}$$

Rearranging equation (A1) yields[12]

$$\mathbf{d} = -\sum_{j \neq k} \frac{\lambda_j}{\lambda_d}\mathbf{y}_j + \frac{1}{\lambda_d}\mathbf{y}_k.$$

Choosing $s_j = \lambda_j/\lambda_d$ for all $j \neq k$ and $s_k = 1/\lambda_d$ yields the first-best effort by the agent's action choice $\mathbf{a} = \mathbf{Ys}$.

To prove the second part of the proposition, consider a system $\mathbf{P} = (P_1, \ldots, P_m)$ of performance measures under which the first-best action can be induced with incentive weights $\mathbf{s} = (s_1, \ldots, s_m)$, where $s_k = 0$ for some k, that is, measure k is not used for incentive contracting. From the first part of the proof, it follows that $\mathbf{d}$ is an element of the subspace $\mathcal{Y}_{-k}$ spanned by $(\mathbf{y}_1, \ldots, \mathbf{y}_{k-1}, \mathbf{y}_{k+1}, \ldots, \mathbf{y}_m)$. Furthermore, by the full rank of $\mathbf{Y}$, y_k is not an element of that subspace. Thus, there exist effort variations orthogonal to $\mathcal{Y}_{-k}$ that leave V and P_j unchanged for $j \neq k$, but increase the performance measure P_k. Hence $\mathbf{P}$ cannot be balanced. ■

Proof of Corollary 1. The corollary directly follows from part (ii) of Proposition 1.

Proof of Proposition 2. If the principal's fallback position is a formal contract, the described contract is credible if

$$\bar{B} = \frac{1}{2}\mathbf{d}'\mathbf{d} \leq [\Pi^{FB} - \Pi^0]/r. \tag{A2}$$

Since the payable bonus $\bar{B}$ is identical to the first-best total profit Π^{FB}, and $\Pi^0 = \phi\Pi^{FB}$ from the definition of congruity, equation (A2) simplifies to the conditions

$$\Pi^{FB} \leq [\Pi^{FB} - \phi\Pi^{FB}]/r \quad \text{and} \quad r \leq 1 - \phi.$$

[12] The division by λ_d is not critical. If $\mathbf{P}$ is minimal, $\lambda_d \neq 0$ holds by the full rank of $\mathbf{Y}$. If $\mathbf{P}$ is not minimal, the degrees of freedom when choosing the weights $\boldsymbol{\lambda}$ can be used to ensure $\lambda_d \neq 0$.

If the principal's fallback position is to cease production, the relational contract is credible if

$$\bar{B} = \frac{1}{2}\mathbf{d}'\mathbf{d} \leq [\Pi^{FB} - U^R]/r, \tag{A3}$$

which simplifies to the conditions $\Pi^{FB} \leq [\Pi^{FB} - U^R]/r$ and $r \leq 1 - \frac{U^R}{\Pi^{FB}} = 1 - \hat{\phi}$. ■

Proof of Proposition 3. Consider the combined contract $S = s_f + B + \mathbf{P}_v\mathbf{s}_v$, in which the bonus, B, may be offered for a target, T, of any aggregate $\hat{P} = \mathbf{a}'\mathbf{Ys}$ of precise performance measures. If the agent foregoes this bonus, he optimally chooses an effort $\mathbf{a}_v = \mathbf{Y}_v\mathbf{s}_v$, resulting in an expected utility of

$$\mathrm{E}U(\mathbf{a}_v) = s_f + \mathbf{s}_v'\mathbf{Y}_v'\mathbf{Y}_v\mathbf{s}_v - \frac{1}{2}\mathbf{s}_v'\mathbf{Y}_v'\mathbf{Y}_v\mathbf{s}_v = s_f + \frac{1}{2}\mathbf{s}_v'\mathbf{Y}_v'\mathbf{Y}_v\mathbf{s}_v. \tag{A4}$$

If he strives for the bonus, he faces the following constrained optimization problem:

$$\max_{\mathbf{a}} \; \mathrm{E}U^A = s_f + B + \mathbf{a}'\mathbf{Y}_v\mathbf{s}_v - \mathbf{a}'\mathbf{a}$$
$$\text{s.t. } \mathbf{a}'\mathbf{Ys} \geq T.$$

The resulting action is $\mathbf{a} = \mathbf{Y}_v\mathbf{s}_v + \lambda\mathbf{Ys}$, where $\lambda \geq 0$ is the multiplier of the restriction. If the contract parameters $\mathbf{s}_v$, $\mathbf{s}$ and T are designed such that $\mathbf{a} = \mathbf{d}$ (such parameters exist if the scorecard is balanced), the agent's expected utility is

$$\mathrm{E}U(\mathbf{a}^{FB}) = s_f + B + \mathbf{d}'\mathbf{Y}_v\mathbf{s}_v - \frac{1}{2}\mathbf{d}'\mathbf{d}. \tag{A5}$$

Comparing (A4) to (A5), the bonus that is necessary to induce the first-best action has to fulfill

$$B \geq \frac{1}{2}\mathbf{s}_v'\mathbf{Y}_v'\mathbf{Y}_v\mathbf{s}_v - \mathbf{d}'\mathbf{Y}_v\mathbf{s}_v + \frac{1}{2}\mathbf{d}'\mathbf{d} \equiv \hat{B}. \tag{A6}$$

Since $\hat{B}$ simply measures the distance between $\mathbf{d}$ and $\mathbf{Y}_v\mathbf{s}_v$, the bonus is minimized by setting $\mathbf{s}_v = \mathbf{s}^0$, as in the optimal, purely formal contract. To induce the first-best action, T and $\mathbf{s}$ are chosen such that $\lambda\mathbf{Ys} = \mathbf{d} - \mathbf{y}^0$. The payable bonus becomes $\hat{B} = \mathbf{d}'\mathbf{d}/2 - (\mathbf{y}^0)'\mathbf{y}^0/2 = \Pi^{FB} - \Pi^0$, which is credibly promised if

$$\Pi^{FB} - \Pi^0 \leq [\Pi^{FB} - \Pi^0]/r \quad \text{or} \quad r \leq 1$$

for a valuable, purely formal contract. If the latter is not beneficial, the relational part of the contract is credible if

$$\Pi^{FB} - \Pi^0 \leq [\Pi^{FB} - U^R]/r \quad \text{or} \quad r \leq \frac{\Pi^{FB} - U^R}{\Pi^{FB} - \Pi^0} = \frac{1 - \hat{\phi}}{1 - \phi}.$$

■

Proof of Proposition 4. As was the case for the first-best action, **d**, analyzed in Proposition 3, the implementation of any action **a** requires—according to (A6)—a bonus payment of

$$B \geq \frac{1}{2}\mathbf{s}_v'\mathbf{Y}_v'\mathbf{Y}_v\mathbf{s}_v - \mathbf{a}'\mathbf{Y}_v\mathbf{s}_v + \frac{1}{2}\mathbf{a}'\mathbf{a} \equiv B^a. \tag{A7}$$

The principal therefore faces the following optimization problem:

$$\max_{\mathbf{s_v},\mathbf{a}} \ \mathbf{a}'\mathbf{d} - \frac{1}{2}\mathbf{a}'\mathbf{a}$$

$$\text{s.t. } rB^a \leq \mathbf{a}'\mathbf{d} - \frac{1}{2}\mathbf{a}'\mathbf{a} - (\mathbf{y}^0)'\mathbf{y}^0.$$

The first-order condition with respect to $\mathbf{a}$ is $(\mathbf{d} - \mathbf{a})(1 + \lambda) + \lambda r(\mathbf{Y}_v\mathbf{s}_v - \mathbf{a}) = 0$. The optimal effort to be induced is therefore

$$\mathbf{a} = \underbrace{\frac{\lambda r}{1 + \lambda + \lambda r}}_{\kappa}\mathbf{Y}_v\mathbf{s}_v + \underbrace{\frac{1 + \lambda}{1 + \lambda + \lambda r}}_{1-\kappa}\mathbf{d}.$$

Since λ and r are nonnegative, this is a convex combination of $\mathbf{d}$ and $\mathbf{Y}_v\mathbf{s}_v$. Obviously, the bonus is minimized by choosing $\mathbf{s}_v = \mathbf{s}^0$, and $\mathbf{a} = \kappa\mathbf{y}^0 + (1 - \kappa)\mathbf{d}$.

The principal's expected profit from an action $\mathbf{a} = \kappa\mathbf{y}^0 + (1 - \kappa)\mathbf{d}$ is

$$\begin{aligned} U^h &= \mathbf{a}'\mathbf{d} - \frac{1}{2}\mathbf{a}'\mathbf{a} - U^R \\ &= \kappa\mathbf{d}'\mathbf{y}^0 - (1 - \kappa)\mathbf{d}'\mathbf{d} - \frac{\kappa}{2}(\mathbf{y}^0)'\mathbf{y}^0 - \kappa(1 - \kappa)\mathbf{d}'\mathbf{y}^0 - \frac{1 - \kappa}{2}\mathbf{d}'\mathbf{d} - U^R. \end{aligned}$$

Taking into account that $\mathbf{d}'\mathbf{y}^0 = (\mathbf{y}^0)'\mathbf{y}^0$ for the optimal formal contract, this equals

$$U^h = \frac{\kappa}{2}(\mathbf{y}^0)'\mathbf{y}^0 + \frac{1 - \kappa}{2}\mathbf{d}'\mathbf{d} - U^R = \kappa^2\Pi^0 + (1 - \kappa^2)\Pi^{FB} - U^R. \tag{A8}$$

From (A7), the required bonus is

$$\begin{aligned} B^a &= \frac{1}{2}(\mathbf{y}^0)'\mathbf{y}^0 - \mathbf{a}'\mathbf{y}^0 + \mathbf{a}'\mathbf{a} \\ &= \frac{1}{2}(\mathbf{y}^0)'\mathbf{y}^0 - \kappa(\mathbf{y}^0)'\mathbf{y}^0 - (1 - \kappa)\mathbf{d}'\mathbf{y}^0 - \frac{\kappa}{2}(\mathbf{y}^0)'\mathbf{y}^0 - \kappa(1 - \kappa)\mathbf{d}'\mathbf{y}^0 \\ &\quad - \frac{1 - \kappa}{2}\mathbf{d}'\mathbf{d}, \end{aligned}$$

which for $\mathbf{d}'\mathbf{y}^0 = (\mathbf{y}^0)'\mathbf{y}^0$ equals

$$B^a = \frac{(1 - \kappa)^2}{2}(\mathbf{d}'\mathbf{d} - (\mathbf{y}^0)'\mathbf{y}^0) = (1 - \kappa)^2(\Pi^{FB} - \Pi^0). \tag{A9}$$

For the fallback position of a formal contract, the bonus payment is credible if $rB^a \leq U^h - U^0$, or

$$\begin{aligned} r(1 - \kappa)^2(\Pi^{FB} - \Pi^0) &\leq \kappa^2\Pi^0 + (1 - \kappa^2)\Pi^{FB} - U^R - (\Pi^0 - U^R) \\ &= (1 - \kappa^2)(\Pi^{FB} - \Pi^0), \end{aligned}$$

which holds for $r \leq (1 + \kappa)/(1 - \kappa)$ or $(r - 1)/(r + 1) \leq \kappa \leq 1$. Thus, if $\phi \geq \hat{\phi}$, a credible hybrid contract exists for all levels of the principal's discount rate.

If a purely explicit contract is not valuable, the bonus is credible if $rB^a \leq U^h$ or $r(1-\kappa)^2(\Pi^{FB}-\Pi^0) \leq \kappa^2\Pi^0 + (1-\kappa^2)\Pi^{FB} - U^R$. This is fulfilled for

$$\frac{r}{1+r} - \frac{\sqrt{(1-\phi)(1+r\phi-\hat{\phi}(1+r))}}{(1-\phi)(1+r)} \leq \kappa \leq \frac{r}{1+r} + \frac{\sqrt{(1-\phi)(1+r\phi-\hat{\phi}(1+r))}}{(1-\phi)(1+r)} \qquad \text{(A10)}$$

Such κ exist if the radicand in (A10) is nonnegative, which holds for $r \leq (1-\hat{\phi})/(\hat{\phi}-\phi)$. To prove that $\kappa \in [0, 1]$, one can take into account that the square root in (A10) is maximal for $\phi = \hat{\phi}$, which yields an upper bound $\bar{\kappa} = 1$ and a lower bound $\underline{\kappa} = (r-1)/(r+1)$. The lower bound is nonnegative for $r > 1$. ■

Proof of Proposition 5. Consider the principal's problem of choosing the parameter κ in order to maximize his expected profit (A8) subject to the credibility constraints $rB^a \leq U^h - U^0$ or $rB^a \geq U^h$. With $\Pi^0 = \phi\Pi^{FB}$, for $\phi < \hat{\phi}$ this takes the form:

$$\max_{\kappa} U^h = \kappa^2\phi\Pi^{FB} + (1-\kappa^2)\Pi^{FB} - U^R$$
$$\text{s.t. } rB^a = r(1-\kappa)^2(\Pi^{FB} - \phi\Pi^{FB}) \leq \kappa^2\phi\Pi^{FB} + (1-\kappa^2)\Pi^{FB} - U^R.$$

By the envelope theorem, the total impact of congruity on profit is given by the partial derivative of the Lagrangian with respect to ϕ,

$$\frac{\mathrm{d}U^h}{\mathrm{d}\phi} = \frac{\partial\mathcal{L}}{\partial\phi} = \kappa^2\Pi^{FB}(1+\lambda) + \lambda r\Pi^{FB}(1-\kappa)^2,$$

which is nonnegative because $\lambda \geq 0$. For $\phi \geq \hat{\phi}$, the credibility constraint simplifies to $r \leq (r-1)/(r+1)$ and is independent of ϕ (see the proof of Proposition 4). The total impact of congruity is therefore

$$\frac{\mathrm{d}U^h}{\mathrm{d}\phi} = \kappa^2\Pi^{FB} \geq 0.$$

■

REFERENCES

ABREU, D. "On the Theory of Infinitely Repeated Games with Discounting." *Econometrica* 56 (1988): 383–96.

BAIMAN, S., AND M. V. RAJAN. "The Informational Advantages of Discretionary Bonus Schemes." *The Accounting Review* 70 (1995): 557–79.

BAKER, G. P. "The Use of Performance Measures in Incentive Contracting." *American Economic Review* 90 (2000): 415–20.

BAKER, G. P. "Distortion and Risk in Optimal Incentive Contracts." *The Journal of Human Resources* 37 (2002): 728–51.

BAKER, G. P.; R. GIBBONS; AND K. J. MURPHY. "Subjective Performance Measures in Optimal Incentive Contracts." *Quarterly Journal of Economics* 109 (1994): 1125–56.

BAKER, G. P.; R. GIBBONS; AND K. J. MURPHY. "Relational Contracts and the Theory of the Firm." *Quarterly Journal of Economics* 117 (2002): 39–83.

BANKER, R. D., AND S. M. DATAR. "Sensitivity, Precision, and Linear Aggregation of Signals for Performance Evaluation." *Journal of Accounting Research* 27 (1989): 21–39.

BANKER, R. D.; G. POTTER; AND D. SRINIVASAN. "An Empirical Investigation of an Incentive Plan that Includes Nonfinancial Performance Measures." *The Accounting Review* 75 (2000): 65–92.

BERCK, P., AND K. SYDSAETER. *Economists' Mathematical Manual*, Second edition. Berlin: Springer, 1993.

BULL, C. "The Existence of Self-Enforcing Implicit Contracts." *Quarterly Journal of Economics* 102 (1987): 147–59.

BUSHMAN, R. M.; R. J. INDJEJIKIAN; AND A. SMITH. "CEO Compensation: The Role of Individual Performance Evaluation." *Journal of Accounting & Economics* 21 (1996): 161–93.

DATAR, S. M.; S. KULP; AND R. A. LAMBERT. "Balancing Performance Measures." *Journal of Accounting Research* 39 (2001): 75–92.

DEMOUGIN, D., AND C. FLUET. "Mechansim Sufficient Statistic in the Risk-Neutral Agency Problem." *Journal of Institutional and Theoretical Economics* 154 (1998): 622–39.

DIKOLLI, S. "Agent Employment Horizons and Contracting Demand for Forward-Looking Performance Measures." *Journal of Accounting Research* 39 (2001): 481–94.

DIKOLLI, S., AND S. KULP. "Interrelated Performance Measures, Interactive Effort, and Optimal Incentives." Harvard NOM Research Paper No. 03–17, Harvard Business School, 2002.

FELTHAM, G. A., AND J. XIE. "Performance Measure Congruity and Diversity in Multi-Task Principal/Agent Relations." *The Accounting Review* 69 (1994): 429–53.

FELTHAM, G. A., AND M. G. H. WU. "Public Reports, Information Acquisition by Investors, and Management Incentives." *Review of Accounting Studies* 5 (2000): 155–90.

GIBBS, M.; K. A. MERCHANT; W. A. VAN DER STEDE; AND M. E. VARGUS. "Determinants and Effects of Subjectivity in Incentives." *The Accounting Review* 79 (2004): 409–36.

HAYES, R. M., AND S. SCHAEFER. "Implicit Contracts and the Explanatory Power of Top Executive Compensation for Future Performance." *RAND Journal of Economics* 31 (2000): 273–93.

HOLMSTRÖM, B., AND P. MILGROM. "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design." *Journal of Law, Economics and Organization* 7 (1991): 24–52.

ITOH, H. "Incentives to Help in Multi-Agent Situations." *Econometrica* 59 (1991): 611–36.

ITTNER, C. D.; D. F. LARCKER; AND M. W. MEYER. "Subjectivity and the Weighting of Performance Measures: Evidence from the Balanced Scorecard." *The Accounting Review* 78 (2003): 725–58.

JOHNSON, L. W.; R. D. RIESS; AND J. T. ARNOLD. *Introduction to Linear Algebra*, Third edition. Boston, MA: Addison-Wesley, 1993.

KAPLAN, R. S., AND D. P. NORTON. "The Balanced Scorecard—Measures that Drive Performance." *Harvard Business Review* (1992): 71–9.

KAPLAN, R. S., AND D. P. NORTON. "Using the Balanced Scorecard as a Strategic Management System." *Harvard Business Review* (1996): 75–87.

KAPLAN, R. S., AND D. P. NORTON. *The Strategy-Focussed Organization*. Boston, MA: Harvard Business School Publishing, 2001.

KERR, S. "On the Folly of Rewarding for A While Hoping for B." *Academy of Management Journal* 18 (1975): 769–83.

MACLEOD, B. "Optimal Contracting with Subjective Evaluation." *American Economic Review* 93 (2003): 216–40.

PEARCE, D. G., AND E. STACCHETTI. "The Interaction of Implicit and Explicit Contracts in Repeated Agency." *Games and Economic Behavior* 23 (1998): 75–96.

RAJAN, M. V., AND S. REICHELSTEIN. "Subjective Performance Indicators and Discretionary Bonus Pools." *Journal of Accounting Research* 44 (2006): 585–618.

SMITH, M. J. "Gaming Nonfinancial Performance Measures." *Journal of Management Accounting Research* 14 (2002): 119–33.