# When is it Foolish to Reward for A while Benefiting from B?

Wendelin Schnedler*
Department of Economics
University of Heidelberg

July 25, 2008

**Abstract**

A performance measure may or may not reflect the relative importance of different tasks for the production of benefit: it can be *aligned* or *unaligned*. Here, we examine when using an aligned measure generates a larger surplus in a principal-agent relationship than using an unaligned but otherwise identical measure. We find that (i) the agent's effort costs matter for the optimal way of measuring performance and (ii) the optimal measure is not aligned but tilted towards tasks that the agent finds easy. Failing to recognize these insights may lead to false predictions about the use of incentives.

JEL-Codes: J33, D86, M52, M55
Keywords: hidden action, moral-hazard, agency model, multitasking, incentives, signal-to-noise ratio, congruity, congruence, distortion

*I thank Uschi Backes-Gellner, George Baker, Jörg Budde, Ingolf Dittmann, Oliver Fabel, Paul Grout, Matthias Kräkel, Ernst Maug, Jörg Oechssler, Anne Pommier, Dieter Sadowski, Dirk Sliwka, Uwe Sunde, and Alfred Wagenhofer for their help. The paper also benefited from discussions with participants at the CMPO seminar in Bristol, the Kolloquium für Personalökonomik in Zürich 2003, the Symposium zur ökonomischen Analyse der Unternehmung in Frankfurt 2003 and the Annual Meeting of the Verein für Socialpolitik 2004. Contact: wendelin.schnedler@awi.uni-heidelberg.de

# 1 Introduction

Incentives are used to motivate an agent to engage in an activity that is costly to him but beneficial to a principal. At the heart of any incentive scheme is some overall assessment of the agent's performance, which is based on information about the agent's activity such as the shareholder value of the employing firm, the number of working hours, profits of his department, etc.[1] A particular way to aggregate this information is to mimic the benefit with the performance measure. Performance measures that attach exactly the same relative importance to tasks (or aspects of the activity) as the benefit are *aligned* with the benefit. Such aligned performance measures seem appealing: by rewarding the agent according to the benefit, the agent not only incurs the costs of exerting effort at different tasks but also internalizes (at least partially) the created benefit. Alignment between performance measure and benefit has been advocated by practitioners and theorists alike. Kerr (1975) provides various real-life examples that show why it is "foolish to reward for A while hoping for B." Baker, Gibbons, and Murphy (1994) argue[2] that "basing pay on an employee's contribution to firm value would have prevented seemingly dysfunctional behavior." This suggests that a performance measure that is aligned with the benefit generates a higher surplus than one that is not. The conditions under which aligned measures yield a higher surplus, however, have not yet been examined. Here, we fill this gap and identify when alignment is valuable.

Apart from the well-known difficulty of providing incentives when the agent dislikes risk and his performance depends on factors beyond his control, there is often the problem that the principal is more interested in certain aspects of the work than the agent. For example, many academics have a stronger preference for research as opposed to teaching than their employing institution. An important finding of this article is that the value of performance measures is affected by the agent's preferences. More specifically, there is a force that favors unaligned performance measures that emphasize aspects of the work that the agent finds relatively attractive. In the example, this force tilts incentives for academics towards research rather than teaching.[3] In order to tease out this force, consider a stripped-down version of the relationship between university and academic and two different performance measures. The first reflects the relative importance of research and teaching to the university; it is aligned with the university's benefit. The second emphasizes research rather than teaching; it is not aligned. Otherwise both performance measures are identical; in particular, the measurement error is the same. Which of the two measures should the university use?

There is a sizable literature in personnel economics and accounting that quantifies the similarity between benefit and performance measure (see e.g., Feltham and Xie 1994; Baker 2000 and 2002; Feltham and Wu 2000; Datar, Cohen Kulp and Lambert 2001). This literature is of little help in deciding which measure to select because it does not link this similarity to the surplus that can be generated with a measure. However, we can use the prevailing idea in this literature that optimal incentives are the result of a trade-off involving alignment. For example,

[1] Jensen, Murphy, and Wruck (2004, p.17) forcefully argue that performance information has to be aggregated to a one-dimensional measure in order to tie remuneration to it.

[2] In their example, the benefit of the activity is the firm value.

[3] Indeed incentives seem to be geared in this way—see, e.g., the report of the National Committee of Inquiry into Higher Education (1997) commissioned by the UK government (chapter 14, section 6).

the trade-off between distortion and risk proposed by Baker (2000, 2002) suggests that less distorted measures generate more surplus when risk is held constant and that the university values the aligned measure more. In Section 2, we analyze a formal version of this simple example and find that the opposite is true: the university has a higher willingness to pay for the measure that emphasizes research. The intuition is straightforward: letting the academic carry out research is relatively cheap and produces a gain that more than offsets the loss from focusing too little on teaching.

In order to examine this finding in a more general context, we use the famous linear multitasking model by Holmström and Milgrom (1991). Following Feltham and Xie (1994), we assume that costs are quadratic in order to obtain a tractable version with closed form solutions. The model, which we describe in Section 3, particularly suits the purpose of studying the effect of alignment on the value of a performance measure. In this model, benefit and performance measure are a linear function of efforts and inspired by the work of Baker (2000, 2002), we can thus represent the performance measure using three simple parameters. The first parameter describes how different tasks are weighted in the performance measure. Mathematically, this parameter is the direction of the marginal effect vector. If this direction is the same as the direction of the marginal effect vector of the benefit, the measure is aligned. The second parameter is the variation of the performance measure that is controlled by the agent (the length of the marginal effect vector). Finally, the third parameter is the variation beyond his control (the standard deviation of the measurement error). This parameterization enables us to examine the effect of each parameter while holding the others fixed. A first simple result of this type is the following: once the weighting of tasks is fixed, the surplus generated by a performance measure gets larger if and only if the ratio of controlled and uncontrolled variation, the signal-to-noise ratio, increases (Lemma 1). A high signal-to-noise ratio is thus a valuable property of a performance measure. The finding extends a result by Kim and Suh (1991) to multitasking models. Since it is only the ratio of controlled and uncontrolled variation that matters, two rather than three parameters suffice to characterize a performance measure with respect to the generated surplus: the relative weighting of tasks and the signal-to-noise ratio. Hence, we can study whether alignment is valuable by comparing aligned and unaligned measures with the same signal-to-noise ratio. The risk parameter suggested by Baker (2000, 2002) is equivalent to the signal-to-noise ratio and the relative weighting is related to distortion. Baker's trade-off between distortion and risk implicitly assumes that a higher signal-to-noise ratio is desirable. The result provides a formal underpinning for this assumption. In addition, it confirms the intuitive prediction that the principal prefers a performance measure over which the agent has more control if it faces two performance measures with equal weights. Notice that a higher signal-to-noise ratio is valuable irrespective of the specific preferences of the agent.

The above-mentioned literature, which quantifies the similarity between benefit and performance measure, assumes that the same level of effort on different tasks leads to the same costs and that the costs of effort at different tasks do not interact. In Section 4, we examine this setting and find that aligned performance measures generate more surplus than otherwise identical measures (Proposition 1). This provides a first justification for why alignment is valuable. It generalizes a respective result by Datar, Cohen Kulp, and Lambert (2001) to a world where the agent is risk-averse and performance is measured with error. Since alignment is equivalent to no distortion in the sense of Baker (2000, 2002), the result implies that both quantities from

Baker's trade-off are related to surplus: less distortion (more alignment) and lower risk (higher signal-to-noise ratio) are both desirable and trading them off against each other makes sense. Taking this result to our simple example, we predict the university prefers an aligned to an unaligned performance measure over which the academic has the same degree of control *if* he finds research and teaching equally difficult. However, the assumption that the agent is indifferent with respect to different aspects seems somewhat restrictive.

In Section 5, we lift the restriction that effort costs have to be identical and independent across tasks. If we compare an aligned measure with an otherwise identical measure, we find that aligned performance measures are only more valuable if costs are identical and independent (Proposition 2). When costs differ across tasks, we cannot maintain that alignment is a valuable property of performance measures. So, the idea that more alignment is desirable is not robust. Moreover, the trade-off between the two parameters suggested by Baker (2000, 2002), distortion and risk, is no longer meaningful once costs differ: a more distorted performance measure may well be preferred to a less distorted but equally risky one. The result shows that the parameters of a performance measure are not sufficient to determine its value. The preferences of the agent matter and ignoring them leads to false predictions. In our thought experiment, we would predict that the university selects the aligned performance measure although it prefers to emphasize research.

It is well known from the single-task agency literature pioneered by Holmström (1979, 1982) and Shavell (1979) that insurance properties of a performance measure influence the surplus that can be generated with this measure. In their seminal article on multitasking, Holmström and Milgrom (1991) demonstrate that the way performance is measured may have dramatic consequences on how effort is allocated across tasks. The value of a measure is hence determined by insurance as well as allocative properties and it is tempting to conclude that the signal-to-noise ratio relates to insurance properties while alignment reflects allocative properties. We explore this idea in Section 6. We hold the signal-to-noise ratio constant and find the weighting of tasks that maximizes surplus. This optimal weighting emphasizes tasks with low costs more than the aligned measure (Proposition 3). This result informs us how effort costs affect the optimal weighting. It also sheds light on the question of why optimal measures are not aligned. All else being equal, the most valuable performance measure from the university's perspective emphasizes the task that the agent finds easier: research. Independent of the measure that is employed, the agent fully internalizes costs. With an aligned measure, there is hence no need to provide additional incentives for the agent to engage in research. At the same time, emphasizing research means that the agent exerts less effort to prepare his teaching. The focus on research and the ensuing detrimental effect weakens if the agent becomes less risk-averse and disappears entirely if the agent is risk-neutral. Insurance is thus critical to understand why the unaligned measure is preferred to the aligned one. This also explains, why the trade-off between effort allocation and insurance cannot be reflected using the distortion and risk parameters proposed by Baker (2000, 2002): Fixing risk in the sense of Baker does not fully determine insurance properties; some insurance aspects are reflected by distortion, i.e. the relative weight of tasks.

In the final section we summarize and discuss the results, examine their implications, and suggest future avenues for research.

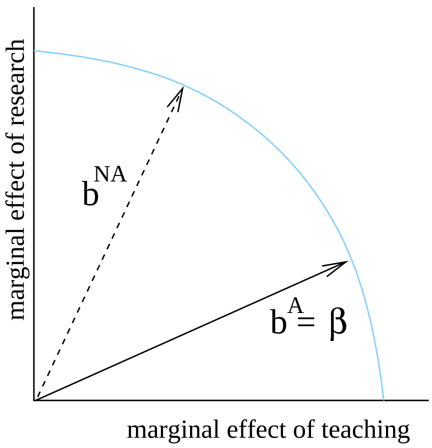
## 2 An example

This section provides a very basic version of the university example to show that an unaligned performance measure can lead to a higher surplus than an aligned measure. The only difference between the two measures is the weight they assign to the two tasks, research and teaching.

Denote the effort exerted by the agent for teaching and research by $e = (e_{\mathrm{T}}, e_{\mathrm{R}})$ and let $C(e)$ be the effort costs. For any given level of effort, the opportunity costs for teaching are higher than those for research: $C(e) = \frac{10}{2} \cdot e_{\mathrm{T}}^2 + \frac{1}{2} \cdot e_{\mathrm{R}}^2$. The effort creates a benefit $B(e)$ for the university, who values teaching more: $B(e) = 2 \cdot e_{\mathrm{T}} + e_{\mathrm{R}}$. Neither efforts $e_{\mathrm{T}}$ and $e_{\mathrm{R}}$ nor the benefit are directly contractible.

There are two proposals to measure performance in order to provide incentives. We abstract from the origins of these proposals and the possibility that the two ways to measure performance may be combined. Instead, we focus on the question: Which of the two proposals does the university want to implement? The first proposal for measuring performance reflects the university's benefit rather well: $P^{\mathrm{A}}(e, \epsilon^{\mathrm{A}}) = 2 \cdot e_{\mathrm{T}} + e_{\mathrm{R}} + \epsilon^{\mathrm{A}}$. The only difference between this performance measure and the benefit is the term $\epsilon^{\mathrm{A}}$, which represents errors when measuring performance. The alternative proposal puts more emphasis on research: $P^{\mathrm{NA}}(e, \epsilon^{\mathrm{NA}}) = e_{\mathrm{T}} + 2 \cdot e_{\mathrm{R}} + \epsilon^{\mathrm{NA}}$. We want to assume that the measurement errors have the same distribution: $\epsilon^{\mathrm{A}} \sim N(0,1)$ and $\epsilon^{\mathrm{NA}} \sim N(0,1)$. Consequently, the two performance measures differ only by the weights they assign to the two tasks and we can completely characterize them by the marginal effects of the agent's effort on each task. These marginal effects are $b^{\mathrm{A}} := (2,1)'$ for $P^{\mathrm{A}}$ and $b^{\mathrm{NA}} := (1,2)'$ for $P^{\mathrm{NA}}$. Similarly, it is possible to represent the marginal effects of effort on the benefit by a vector $\beta^{\mathrm{A}} = (2,1)'$. As pointed out by Baker (2002), the direction of these marginal effect vectors indicates how much emphasis the performance measure assigns to each task. Figure 1 depicts the two measures. The performance measure $P^{\mathrm{A}}$ weighs the two tasks in the same way as the

Figure 1: The marginal effects of teaching and research on the aligned measure, $b^{\mathrm{A}}$, and on the benefit, $\beta$, are the same; the unaligned measure, $b^{\mathrm{NA}}$, emphasizes research (the task with the lower marginal benefit).




benefit; it is *aligned* with the benefit. The other performance measure $P^{\mathrm{NA}}$ is not aligned and emphasizes research, which leads to a lower benefit and lower costs. In order to focus attention on the comparison of the two performance measures, the university is only allowed to buy one performance measure. Which of the two ways to aggregate information is more valuable: the aligned $P^{\mathrm{A}}$ or the unaligned $P^{\mathrm{NA}}$?

The value of a performance measure lies in the surplus that can be generated by an incentive scheme based on this measure. In order to compute this surplus, we impose some additional assumptions. These assumptions have been used by Holmström and Milgrom (1991) and are now standard in models that formally analyze the similarity between measure and benefit.[4] It is assumed that the university is risk-neutral, acts as a mechanism designer, and can condition the agent's wage $W$ on the performance measure $P$ in the following way: $W(P) = w_0 + w_1 P$, where $w_0$ is a base wage and $w_1$ a performance dependent piece-rate.[5] The agent's utility is modeled using a negative exponential form: $U(v) = 1 - e^{(-vr)}$, where $v := W(P) - C$ is the income of the agent and $r$ is the Arrow-Pratt measure for risk aversion; for the example, we set $r = 0.5$.

With these assumptions in place, it is possible to compute the optimal base wage $w_0$ and piece-rate $w_1$ given that a specific performance measure is used. In this context, optimality means that the surplus of the university is maximized while ensuring that the agent participates (participation constraint) and has no incentive to carry out a different activity (incentive constraint). If we measure performance by using $P^{\mathrm{A}}$, the optimal choice of $w_0$ and $w_1$ generates a surplus of $\frac{49}{95} \approx 0.52$. If, on the other hand, we base incentives on the unaligned measure $P^{\mathrm{NA}}$ instead, the optimal contract generates a surplus of $\frac{121}{230} \approx 0.53$; the surplus computations can be found in Appendix B. Although the two performance measures differ only in terms of the emphasis that they place on the two tasks, the university has a higher willingness to pay for the measure that is not aligned with the benefit. This example illustrates that alignment is not desirable under all circumstances and that greater similarity between measure and benefit does not always lead to a higher surplus.

Observe that there is a clear disadvantage of using the unaligned measure: the academic internalizes the relative benefit of teaching to a lower degree than with the aligned measure. The unaligned measure nevertheless generates a higher surplus because efforts for teaching and research are below efficient levels (due to the insurance problem). In this situation, increasing effort is beneficial and the cheapest way to do so is to boost effort on the task that the academic likes. The respective gain more than outweighs the loss due to the wrong focus.[6] A more in-depth argument how insurance issues affect the optimality of the unaligned measures follows in Section 6.

In the next section, we introduce a more general model in order to study the extend of the

[4]see, e.g., Feltham and Xie 1994; Baker 2000 and 2002; Feltham and Wu 2000; Banker and Thevaranjan 2000; Datar, Cohen Kulp, and Lambert 2001.

[5]As Mirrlees (1999) points out, linearity of the contract is not an innocuous assumption. Nevertheless, we follow Holmström and Milgrom (1991) and maintain this assumption so that our results are comparable with the existing multitasking literature. In the examined setting, linearity can also be justified by appealing to Holmström and Milgrom (1987)— see the discussion of the linear multitasking model by Salanié (1998).

[6]An alternative reason that has been given for the use of unaligned measures are commitment problems (Fershtman and Judd 1987; Sklivas 1987): if, for example, an employer faces Cournot-competition, he may want to commit to an aggressive selling strategy by rewarding the agent for sales rather than profits even though his benefit is defined by profits rather than sales. Notice that this explanation does not apply here because commitment problems of this type do not feature in the example.

findings from this example.

## 3 The model

In the previous section, we analyzed a simple two-task example. In this section, we extend this example to the general $n$-task case. The model, which we are going to use throughout the article, is based on the linear principal-agent model by Holmström and Milgrom (1991). As in the example, the principal is risk-neutral, the agent has negative exponential utility and contracts are linear in performance measures. By adding the assumptions that costs are quadratic, Feltham and Xie (1994) obtain a particularly tractable version that allows the computation of closed form solutions and has been employed in personnel economics (Baker 2000, 2002) as well as in accounting (see, e.g., Feltham and Wu 2000 or Datar, Cohen Kulp and Lambert 2001). Here, we generalize this assumption and allow costs to differ across tasks and be interrelated:

$$C(e) = e'Ce,$$

where the matrix $C = (c_{ij})$ is assumed to be symmetric and positive definite, so that costs are always positive: $e'Ce > 0$ if $e \neq (0, \ldots, 0)'$. The cost function considered by Feltham and Xie (1994) and the ensuing literature (see Baker 2000 and 2002; Feltham and Wu 2000; Banker and Theveranjan 2000; Datar, Cohen Kulp, and Lambert 2001) is a special case of the cost function considered here and can be obtained by setting $C = \frac{1}{2}I$, where $I$ is the identity matrix.

Denote the benefit of the principal in the $n$-task case by

$$B(e, \eta) = \beta' e + \eta,$$

where $\beta' = (\beta_1, \ldots, \beta_n)$ is the vector of marginal effects and $\eta$ is a noise term with mean zero, which reflects influences on the benefit beyond the control of the agent. Performance measures in this setting can be written as

$$P(e, \epsilon) = b'e + \epsilon,$$

where $b' = (b_1, \ldots, b_n)$ is the vector of marginal effects and $\epsilon$ is a normally distributed noise term with expected value zero and variance $\sigma^2$. Again, the noise term captures any effects on the performance measure that cannot be influenced by the agent. The noise that influences the benefit may stand in an arbitrary relation to the noise that influences the performance measure: it may be independent, correlated or identical ($\epsilon = \eta$). Hence the description of the performance measure encompasses the special case that it is identical to the benefit ($b = \beta$ and $\epsilon = \eta$).

Any performance measure is uniquely characterized by the marginal effect vector $b$ and the variance of the error term $\sigma^2$. Equivalently, performance measures can be represented using three parameters. The first parameter is the direction of the marginal effect vector $\hat{b} := \frac{b}{||b||} = \frac{b}{\sqrt{b'b}}$, which determines how much weight is attached to each task and is standardized to length one. The second parameter is the length of the marginal effect vector $||b|| = \sqrt{b'b}$, which represents the variation of the performance measure under control of the agent. The third parameter is the standard deviation $\sigma$, which stands for the variation that is not controlled by the agent.

Next, we formally link the marginal effect vector of the performance measure to that of the benefit.

**Definition 1** (Alignment)**.** *Performance measure and benefit are aligned if their marginal effect vectors have the same direction:* $b = \gamma\beta$ *for some* $\gamma$ *different from zero.*

It is natural to think of a positive factor $\gamma$. However, $\gamma$ can also be negative. Then, a larger measurement means that the benefit is lower and the measure is an indicator of "bad" rather than "good" performance. Rewarding for less "bad performance" is perfectly equivalent to rewarding for more "good performance." Thus, it does not matter for practical purposes whether $\gamma$ is positive or negative. The definition relates to various concepts that describe the similarity between benefit and performance measure: aligned performance measures minimize the discongruity of Feltham and Xie (1994), the non-congruency by Feltham and Wu (2000), the distortion by Baker (2000) (for positive $\gamma$) and the incongruity of Datar, Cohen Kulp, and Lambert (2001).[7]

In addition to alignment, we define the following concept.

**Definition 2** (Signal-to-noise ratio)**.** *The signal-to-noise ratio is the ratio of the variation under the control of the agent* $||b||$ *and the variation beyond his control* $\sigma$*:*

$$\rho := \frac{\sqrt{\left(\frac{\partial E_\epsilon(P(e,\epsilon))}{\partial e}\right)'\left(\frac{\partial E_\epsilon(P(e,\epsilon))}{\partial e}\right)}}{\sigma} = \frac{||b||}{\sigma}.$$

The squared signal-to-noise ratio $\rho^2$ is identical to a respective concept defined by Baker (2002), which in turn generalizes the single-task concept by Kim and Suh (1991) and is equivalent to the product of sensitivity and precision considered by Banker and Datar (1989) in the single-task setting.[8] Finally, the squared signal-to-noise ratio amounts to the sum of the squared (single-task) signal-to-noise ratios at the different tasks: $\rho^2 = \frac{b_1^2}{\sigma^2} + \ldots + \frac{b_n^2}{\sigma^2}$. In summary, the signal-noise ratio can in many ways be regarded as a multiple-task extension of respective concepts for the single-task model. Moreover, just as in the single task case, a higher signal-to-noise ratio is valuable.

**Lemma 1.** *A performance measure generates a higher value than another performance measure with the same direction* $\hat{b}$ *if and only if it has a higher signal-to-noise ratio.*

*Proof.* The surplus generated by a measure $(b, \sigma^2)$ is computed in Lemma 2 in Appendix A to be:

$$\phi(b, \sigma^2) = \frac{1}{4}\frac{b'C^{-1}\beta\beta'C^{-1}b}{b'C^{-1}b + 2r\sigma^2}.$$

Using that $\sigma^2 = \frac{||b||^2}{\rho^2}$ and replacing $b = \hat{b} \cdot ||b||$, we get the following surplus formula:

$$\phi\left(\hat{b} \cdot ||b||, \frac{||b||^2}{\rho^2}\right) = \frac{1}{4}\frac{\hat{b}'C^{-1}\beta\beta'C^{-1}\hat{b}}{\hat{b}'\left(C^{-1} + 2\frac{r}{\rho^2}I\right)\hat{b}}. \quad (1)$$

[7] The former two concepts are sums of squared deviations and reduce to zero if we plug in a marginal effect vector for the performance measure, which is a multiple of the marginal effect vector $\beta$ of the benefit. The latter depends on the cosine of the angle between $b$ and $\beta$, which attains its maximum value of one if $b = \gamma\beta$ with positive $\gamma$.

[8] Banker and Datar (1989) also suggest a signal-to-noise ratio for multiple signals. This suggestion does not concern us here, as we are supposing that the information of various signals is already embodied in the performance measure.

The result then follows from the observation that this quantity is monotonically increasing in $\rho$. □

Hence, a high signal-to-noise ratio is desirable. This generalizes a respective finding by Kim and Suh (1991) to the multi-task setting. The intuition is the same as in the single-task case (Kim and Suh, 1991): a higher signal-to-noise ratio means that the same incentives can be provided while imposing less uncertainty on the agent. The result also implies that instead of having separate parameters for controlled and uncontrolled variation, it suffices to use the signal-to-noise ratio when describing the surplus generated by a performance measure. It seems reasonable to require from any notion of risk that a performance measure with less risk is not ranked worse than a performance measure with more risk (otherwise we could simply add noise to the former). Since the signal-to-noise ratio is inversely related to Baker's notion of risk (2002), the result asserts that his notion fulfills this requirement. It also confirms the very intuitive prediction that the university prefers performance measures over which the agent has relatively more control. Notice that a larger signal-to-noise ratio is more valuable independent of costs $C$, the degree of risk aversion $r$, or other parameters in the model. The prediction that the university prefers a higher signal-to-noise ratio is hence robust to changes in these parameters. In the next section, we encounter specific conditions under which greater alignment is more valuable.

# 4 Identical and independent costs across tasks

In the above example, we have already seen that aligned measures may not be more valuable than otherwise identical unaligned measures. In this section, we limit our attention to the case of costs that are identical across tasks and independent. This case deserves particular attention because several models that examine the relationship between benefit and performance measure are based on this assumption.[9]

For the case of equal and independent costs, $C = \frac{1}{2}I$, the surplus that can be generated by a performance measure becomes:[10]

$$\frac{1}{2} \cdot \frac{\hat{b}'\beta\beta'\hat{b}}{\hat{b}'\hat{b}\left(1 + \frac{r}{\rho^2}\right)}.$$

Using the definition of the vector dot product, we can rewrite $\hat{b}'\beta = ||\hat{b}|| \cdot ||\beta|| \cdot \cos\theta = ||\beta|| \cdot \cos\theta$, where $\theta$ is the angle between $\beta$ and $\hat{b}$. Rewriting the surplus accordingly yields:

$$\frac{1}{2} \cdot \frac{(\cos(\theta))^2\beta'\beta}{1 + \frac{r}{\rho^2}}. \tag{2}$$

For an aligned measure, the angle is either zero or 180°, the cosine is 1 or -1, and the surplus attains:

$$\frac{1}{2} \cdot \frac{\beta'\beta}{1 + \frac{r}{\rho_A^2}},$$

[9] see, e.g., Feltham and Xie 1994; Baker 2000 and 2002; Feltham and Wu 2000; Datar, Cohen Kulp, and Lambert 2001.

[10] This follows if we replace $C^{-1} = 2 \cdot I$ in equation (1).

where $\rho_A$ is the signal-to-noise ratio of the aligned measure. Comparing this surplus with (2), we observe that a performance measure can only create a higher surplus if it has a higher signal-to-noise ratio, $\rho > \rho_A$, but not if controlled and uncontrolled variation are constant, i.e. $\rho = \rho_A$. We summarize these considerations in the following proposition:

**Proposition 1.** *Suppose costs are independent and identical: $C = \frac{1}{2}I$. Then, an aligned performance measure maximizes the generated surplus if compared to otherwise identical unaligned performance measures and even if compared to the larger class of measures with the same signal-to-noise ratio.*

This simple result provides a first formal argument as to why alignment may be desirable in a world where insurance matters. The result extends a respective finding by Datar, Cohen Kulp, and Lambert (2001) to the case that the agent is risk-averse and performance is measured with error. In addition, it links the second parameter in Baker's trade-off (2000, 2002) to surplus and hence completes the justification of the implicit assumption that lower risk and lower distortion are desirable: Holding the signal-to-noise ratio constant, a less distorted measure generates more surplus (see Equation (2)) and the maximal surplus is achieved by an aligned measure (Proposition 1).[11] This link between quantities that are traded-off and surplus is not evident. As an example, consider the concepts of non-congruency component and risk-minimization component by Feltham and Wu (2000). Once the risk minimization component is fixed, the relationship between the directions of the marginal effect vector of benefit and performance measure has no effect on surplus. In particular, a more aligned measure does not generate a higher surplus.[12]

Most importantly, the result confirms the received wisdom that alignment is desirable and offers the prediction that the university selects the aligned rather than an otherwise identical measure that is not aligned. Still, the university chooses the performance measure that emphasized research rather than the aligned measure. This seems to be a contradiction. Proposition 1, however, rests on the assumption that costs are identical and independent across tasks while the agent preferred research to teaching. In the next section, we examine more closely the importance of identical costs for the value of aligned measures.

[11] Strictly speaking, less distortion does not necessarily imply a higher surplus when using the definition by Baker (2000, 2002). Baker defines distortion to be larger, the lower $\cos(\theta)$, where $\theta$ is the angle between the marginal effect vectors. This means, however, that a measure $b^I$ that points in the opposite direction of the benefit, so that the angle is $180^o$ has the largest distortion, while the distortion of a measure $b^{II}$ that nearly points in the same direction as the benefit, say with an angle of $10^o$, is lower: $\cos(180^o) = -1 < 0.98 < \cos(10^o)$. Still, $b^I$ generates a larger surplus than $b^{II}$—see equation (2). This problem can easily be rectified by re-defining distortion to be lower, the larger $(\cos\theta)^2$ or the larger $|\cos\theta|$.

[12] A referee of an earlier draft conjectured that aligned measures generate more surplus than unaligned measures when the risk-minimization component is held constant. Appendix E refutes this conjecture.

## 5 Differing and interdependent costs

If we allow costs to differ and to be interrelated, the surplus that is generated from optimally using a performance measure with alignment $\hat{b}$ and signal-to-noise ratio $\rho$ amounts to:[13]

$$\frac{1}{4} \cdot \frac{\hat{b}'C^{-1}\beta\beta'C^{-1}\hat{b}}{\hat{b}'(C^{-1} + 2\frac{r}{\rho^2}I)\hat{b}}.$$

Maximizing this surplus in the alignment $\hat{b}$ is no longer simple because the cost matrix $C$ enters the relationship between $\hat{b}$ and $\beta$ and we cannot use the convenient representation based on the angle between $\hat{b}$ and $\beta$. The following proposition formalizes the effect of costs—the proof is given in Appendix C.

**Proposition 2.** *An aligned performance measure maximizes the surplus amongst otherwise identical measures if and only if effort costs are identical for all tasks and independent:* $C = I \cdot c$*, with* $c$ *being some strictly positive real number.*

An important insight from this result is that greater alignment between benefit and measure is generally not desirable. The context in which alignment is desirable is specific in many ways. As an example take the quadratic form of effort costs or the linearity of the benefit function. Proposition 2 informs us that even within this restricted setting, the optimality of alignment is not robust. Notice that the linearity of the benefit function implies that efforts do not interact in the production of benefit. The units of effort are arbitrary and can possibly be chosen such that benefit becomes a linear function in effort. This, however, induces an interaction of efforts in the cost function. Since there is often some form of interaction in reality (either on the benefit or the cost side), the optimality of aligned performance measures is likely to be rare.

Another implication of the finding is that the trade-off between distortion and risk in the form suggested by Baker (2002) is not robust either. There is no reason to trade low risk for low distortion if low distortion is not desirable. Finally, the result reveals that costs are important to compare the value of performance measures and hence to make predictions about the choice of performance measures.[14] Ignoring that the agent likes research more than teaching would lead to the false prediction that the university prefers the aligned measure.

In the single-task model, it is possible to compare the value of measures by looking at its parameters alone (Kim and Suh, 1991). The analysis of Feltham and Xie (1994) and Baker (2000, 2002) indicates that comparisons under multitasking also require information on how benefit is created. In the general case, which we have considered here for the first time, even that is not sufficient: agent's costs also matter. In multitasking models, the value of a performance measure in relation to another measure is thus highly context specific: it depends on the principal (in form of $\beta$) *and* on the agent (in form of $C$).

In the next section, we examine in more detail why greater alignment does not necessary lead to a higher surplus.

[13]See the re-parameterized surplus formula from equation (1).

[14]Re-parameterizing the model in such a way that effort costs are identical and independent does not eliminate this problem. Effort costs then enter the signal-to-noise ratio (see Appendix F)

## 6 Alignment, insurance, and allocation

In the previous section, we have seen that alignment is not necessarily a desirable property of performance measures. In order to shed more light on this phenomenon, we compute the relative weighting of tasks such that the respective performance measure generates the largest surplus amongst all measures with the same controlled and uncontrolled variation—the computation can be found in Appendix D.

**Proposition 3.** *The performance measure that maximizes the surplus amongst all measures with controlled variation $\xi$ and uncontrolled variation $\sigma$ has the marginal effect vector:*

$$b^* = k\left(I + \frac{2r}{\rho^2}C\right)^{-1}\beta, \tag{3}$$

*where $k > 0$ is a standardization factor,*[15] *which ensures that the vector has the controlled variation $\xi$ and $\rho = \frac{\xi}{\sigma}$.*

The proposition characterizes the optimal measure, where optimal means that the surplus is maximized holding controlled and uncontrolled variation constant. It also gives us two interesting insights into the relationship between optimal and aligned measures. These insights allow us to identify why the intuition that alignment is desirable fails, why the university prefers the unaligned measure, and why the trade-off between distortion and risk in the form suggested by Baker (2000, 2002) only works in the case of equal and identical costs.

First, notice that the agent always fully internalizes the relative marginal costs of tasks. With an aligned measure, he also internalizes the relative marginal benefits. The optimal measure weighs tasks differently. For example in the case of independent costs ($c_{ij} = 0$), it emphasizes those tasks that have lower marginal costs (given the same effort level)—just as the performance measure in the university example.[16] Consequently, the cost side is given more weight than the benefit side and the agent's allocation of effort across tasks will be skewed towards tasks with lower marginal costs. In this sense, effort is "misallocated across tasks" if the optimal rather than the aligned measure is used.

Second, the difference in emphasis between the optimal and the aligned measure is reduced if the signal-to-noise ratio increases. Likewise, it drops if the agent's risk-aversion falls; the optimal measure is aligned if the agent is risk-neutral. So, the weighting of the optimal measure approaches that of an aligned measure if the effect of uncertainty gets smaller, either because there is less uncertainty or because the agent cares less about uncertainty.

In addition to these two insights, observe that by focusing effort on a specific task, this task is measured more accurately; the *task-wise* signal-to-noise ratio increases. Taken together, these observations imply that insurance plays a crucial role in explaining why aligned measures do not maximize surplus. It also means that the signal-to-noise ratio, the parameter suggested by Baker (2002) to measure risk, does not fully capture all aspects related to risk. So, the parameters proposed by Baker do not describe a trade-off between effort allocation and insurance if costs

[15]The standardization factor $k$ takes the value $k = \pm\frac{\xi}{\sqrt{\beta'(2\frac{r}{\rho^2}C+I)^{-2}\beta}}$.

[16]Since costs are quadratic, marginal costs depend on the effort level. Holding this effort at the same level for two tasks $i$ and $j$ allows us to compare their marginal costs by comparing the cost coefficients $c_{ii}$ and $c_{jj}$.

differ. In a more general sense, however, such a trade-off seems to operate:[17] if the agent is risk-neutral, then insurance does not matter, mis-allocation considerations dominate and the aligned measure is optimal. As the agent gets more risk-averse, insurance becomes more important. Then, insurance considerations outweigh the worse allocation of effort across tasks and it pays for the principal to focus the effort of the agent using the performance measure. Tasks with lower costs are emphasized in the performance measure because this is the least-costly way to focus the agent's effort.

The overall message that emerges from this section is that the way a performance measure affects insurance does not only depend on the signal-to-noise ratio. The weighting of tasks in the performance measure is also important.

# 7 Conclusion

We have analyzed under which conditions measures that are aligned with the benefit are more valuable than measures that are not aligned. While it is known that alignment is desirable if agents are risk neutral or have full control over the performance measure, we find that aligned measures also generate a larger surplus than otherwise identical measures in the more interesting case that insurance matters. This, however, holds true only if effort costs are identical across tasks and independent.

While equal and independent costs ensure that aligned measures generate a higher surplus than otherwise identical measures, they do not guarantee that a greater similarity between measure and benefit increases surplus. Such a claim is specific to the concept that we use to quantify similarity and the literature offers different respective concepts: the discongruity by Feltham and Xie (1994), the non-congruency by Feltham and Wu (2000), the distortion by Baker (2000, 2002) and incongruity by Datar, Cohen Kulp, and Lambert (2001). Examining the relationship between these concepts and the generated surplus has not been the aim of this article and is left for future analysis. In all of these concepts, however, alignment plays a prominent role, and here we supply a condition that is necessary for alignment to be valuable. The present article thus provides a starting point for examining when a greater similarity between measure and benefit is desirable.

More importantly, the article shows that the trade-off between effort allocation and insurance involved in incentive design can be reflected by the two parameters of performance measures proposed by Baker (2000, 2002), distortion and risk, if effort costs are identical and independent across tasks. The beauty of distortion and risk is that they are easy to grasp and possibly even observable in practice. While distortion and risk may be helpful pedagogical devices to illustrate the trade-off between effort allocation and insurance, they are, however, not enough to fully understand this trade-off.

We have seen that more distorted measures can generate more surplus than aligned measures with the same risk, where risk is defined in the sense of Baker (2002). They do so despite the fact that they attach more emphasis to tasks with low marginal benefit than unbiased measures and channel the agent's effort towards the "wrong" tasks. The higher the risk-aversion of the

[17]Schnedler (2008) decomposes agency costs into effort allocation loss and insurance loss in order to formalize the quantities that are traded-off.

agent, the more pronounced is this "misallocation" effect. So, the attitude of the agent towards risk matters although the risk characteristic proposed by Baker (2002), the signal-to-noise ratio, is fixed. This finding implies that the signal-to-noise ratio alone does not capture all aspects of the performance measure that are relevant to insurance. Conversely, the relative weighting of tasks also affects the risk imposed on the agent. This observation calls for a new concept that could describe the insurance properties of a performance measure and disentangle them from its allocative properties. Future research could address this lack.

An important insight from this article is that parameters of the performance measure alone cannot reflect its value relative to other measures. Predictions about the use of performance measures have to take into account how difficult the agent finds which task. Here, we assumed that the employer knows effort costs but in a more general setting, the employer might have to elicit them. For obvious reasons, it is not possible to let the worker, who usually knows his costs, have complete freedom in choosing the performance measure. An interesting question for follow-up research may be whether it is possible to elicit these costs by offering the worker a menu of performance measures from which he can select one.

## References


Baker, George. 2000. The use of performance measures in incentive contracting. *American Economic Review* 90, no. 2:415–420.

Baker, George. 2002. Distortion and risk in optimal incentive contracts. *Journal of Human Resources* 37, no. 4:728–751. Special Issue on Designing Incentives to Promote Human Capital.

Baker, George, Robert Gibbons, and Kevin J. Murphy. 1994. Subjective performance measures in optimal incentive contracts. *Quarterly Journal of Economics* 109, no. 4:1125–1156.

Banker, Rajiv D. and Srikant Datar. 1989. Sensitivity, precision, and linear aggregation of signals for perfomance evaluation. *Journal of Accounting Research* 27, no. 1:21–39.

Banker, Rajiv D. and Alex Thevaranjan. 2000. Goal congruence and evaluation of performance measures. Unpublished Manuscript. School of Management. University of Texas at Dallas.

Datar, Srikant, Susan Cohen Kulp, and Richard A. Lambert. 2001. Balancing performance measures. *Journal of Accounting Research* 39, no. 1:75–92.

Feltham, Gerald .A. and Martin G. H. Wu. 2000. Public reports, information acquisition by investors, and management incentives. *Review of Accounting Studies* 5, no. 2:155–190.

Feltham, Gerald A. and Jim Xie. 1994. Performance measure congruity and diversity in multi-task Principal/Agent relations. *The Accounting Review* 69, no. 3:429–453.

Fershtman, Chaim and Kenneth L. Judd. 1987. Equilibrium incentives in oligopoly. *The American Economic Review* 77, no. 5:927–940.

Holmström, Bengt. 1979. Moral hazard and observability. *Bell Journal of Economics* 10, no. 1:74–91.

Holmström, Bengt. 1982. Moral hazard in teams. *Bell Journal of Economics* 13, no. 2:324–340.

Holmström, Bengt and Paul Milgrom. 1987. Aggregation and linearity in the provision of intertemporal incentives. *Econometrica* 55, no. 2:303–328.

Holmström, Bengt and Paul Milgrom. 1991. Multitask principal-agent-analysis: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics, and Organization* 7:24–52. Special issue.

Jensen, Michael C., Kevin J. Murphy, and Eric Wruck. 2004. Remuneration: Where we've been, how we got to here, what are the problems, and how to fix them. Finance Working Paper 44, ECGI.

Kerr, Steven. 1975. On the folly of rewarding A while hoping for B. *Academy of Management Journal* 18, no. 4:769–783.

Kim, Son Ku and Yoon S. Suh. 1991. Ranking of accounting information systems for management control. *Journal of Accounting Research* 29, no. 2:386–396.

Kräkel, Matthias. 1999. *Organisation und Management*. Tübingen: Mohr Siebeck. Chapter B: Hidden-action im LEN-Modell, pp. 62–77.

Mirrlees, James A. 1999. The theory of moral hazard and unobservable behavior (part 1). *Review of Economic Studies* 66:3–21. Reprint of a manuscript from 1975.

Salanié, Bernard. 1998. *Economics of contracts*. Massachussets: MIT Press.

Schnedler, Wendelin. 2008. Incentives and the level and allocation of effort. Unpublished Paper. University of Heidelberg.

Shavell, Steven. 1979. Risk-sharing and incentives in the principal-agent relationship. *Bell Journal of Economics* 10, no. 1:55–73.

Sklivas, Steven D. 1987. The strategic choice of managerial incentives. *The RAND Journal of Economics* 18, no. 3:452–458.

The National Committee of Inquiry Into Higher Education (chaired by Sir Ron Dearing). 1997. Higher education in the learning society. Website: www.leeds.ac.uk/educol/ncihe/.

# A Lemmata

At various instances, we need the maximal surplus that can be generated by a specific performance measure $(b, \sigma^2)$. This surplus is computed in the following lemma.

**Lemma 2.** *Given a performance measure with marginal effects $b$ and variance $\sigma^2$, the optimal rate of performance pay is:*

$$w_1^* = \frac{\beta' C^{-1} b}{b' C^{-1} b + 2r\sigma^2}. \tag{4}$$

*Under this rate, the agent will exert the effort:*

$$e^* = \frac{w_1^*}{2} C^{-1} b, \tag{5}$$

*which yields a surplus of:*

$$\phi(b, \sigma^2) = \frac{1}{4} \cdot \frac{b' C^{-1} \beta \beta' C^{-1} b}{b' C^{-1} b + 2 r \sigma^2}. \tag{6}$$

*Proof.* The best linear scheme that is based on performance measure $P$ maximizes the surplus of the principal while ensuring that the agent has an incentive to select a particular activity (incentive constraint) and is accepting the contract (participation constraint):

$$\max_{w_0, w_1} \mathrm{E}\left(B(e) - w_0 - w_1 P(e)\right) \tag{7}$$

$$\text{such that} \quad e \in \mathrm{argmax}_{\check{e}} \mathrm{E}\left(U\left(w_0 + w_1 P(\check{e}) - C(\check{e})\right)\right) \tag{8}$$

$$\text{and} \quad \mathrm{E}\left(U\left(w_0 + w_1 P(e) - C(e)\right)\right) \geq 0, \tag{9}$$

where the outside option of the agent has been standardized to zero. Using the linearity of the performance measure, the normality distribution of the noise, and the shape of the utility function, the base wage can be chosen to compensate the agent for any harmful effects of uncertainty ($w_0 = C(e) - w_1 b' e + w_1^2 r \frac{b'b}{\rho^2}$), so that the participation constraint can be eliminated from the program (see e.g. Salanié 1998 or Kräkel 1999). Recalling the definitions of performance measure, benefit, and costs, as well as the risk neutrality of the principal, the program then simplifies to:

$$\begin{gathered} \max_{w_1} \beta' e - e' C e - \frac{w_1^2}{2} r \sigma^2 \\ \text{such that } e \in \mathrm{argmax}_{\check{e}} \ \mathrm{E}\left(U\left(w_0 + w_1 b' \check{e} - \check{e}' C \check{e}\right)\right). \end{gathered} \tag{10}$$

Consider the problem that the agent faces for a given incentive scheme $(w_0, w_1)$, i.e. the side constraint of the principal's problem. Because the agent's utility function $U$ is monotonic, the agent chooses effort so as to maximize the difference between received wage and effort costs:

$$\max_{e} w_1 b' e - e' C e.$$

The objective function is concave because the second derivative is a symmetric, negative definite matrix $(-C)$. Thus, the maximizer can be determined by the first-order condition. Solving for $e$ yields:

$$e = \frac{w_1}{2} C^{-1} b. \tag{11}$$

Note, that $C^{-1}$ exists because $C$ is positive definite. Replacing $e$ in equation (10) by the optimal effort from equation (11) yields the following expression for the objective function:

$$\max_{w_1} \frac{w_1}{2} \beta' C^{-1} b - \frac{w_1^2}{2^2} b' C^{-1} b - \frac{w_1^2}{2} r \sigma^2. \tag{12}$$

Again, the objective function is concave and solving the first order condition gives the maximizer:

$$w_1 = \frac{\beta' C^{-1} b}{b' C^{-1} b + 2r\sigma^2}. \tag{13}$$

Using the optimal wage rate in (12) finally results in the surplus from an optimal incentive scheme based on the performance measure with marginal effect vector $b$ and variance $\sigma^2$:

$$\phi(b, \sigma^2) = \frac{1}{4} \cdot \frac{b' C^{-1} \beta \beta' C^{-1} b}{b' C^{-1} b + 2r\sigma^2}.$$

□

The next lemma helps us to find the set of maximizing marginal effect vectors for a specific quadratic form. It is needed to prove Lemma 4.

**Lemma 3.** *The set of maximizers of the expression*

$$\frac{b' C^{-1} \beta \beta' C^{-1} b}{b' H b},$$

*where $H$ is a symmetric and positive definite matrix, is*

$$\left\{ b^* \middle| b^* = k \cdot (CH)^{-1} \beta, \text{ with } k \in \mathbb{R}, k > 0 \right\}.$$

*Proof.* Because $H$ is symmetric and positive definite, we can decompose it: $H = P \Lambda P'$ and define $H^{\frac{1}{2}} = P \Lambda^{\frac{1}{2}} P'$. Now, define $\check{b} := H^{\frac{1}{2}} b$, so that $b = H^{-\frac{1}{2}} \check{b}$ and consider the transformed problem:

$$\max_{\check{b}} \frac{\check{b}' H^{-\frac{1}{2}} C^{-1} \beta \beta' C^{-1} H^{-\frac{1}{2}} \check{b}}{\check{b}' \check{b}}.$$

To advance on this problem, we fix the length of $\check{b}$ to some arbitrary value $k^2$: $\check{b}'\check{b} = k^2$. By varying $k$, we will later obtain the set of all possible solutions. The respective Lagrangian for a given $k$ is:

$$L(\check{b}, k) = \check{b}' H^{-\frac{1}{2}} C^{-1} \beta \beta' C^{-1} H^{-\frac{1}{2}} \check{b} - \lambda \left( \check{b}'\check{b} - k^2 \right).$$

The corresponding first-order conditions are:

$$\left( (H^{-\frac{1}{2}})' C^{-1} \beta \beta' C^{-1} H^{-\frac{1}{2}} - \lambda I \right) \check{b}^* = 0 \text{ and } (\check{b}^*)' \check{b}^* = k^2. \tag{14}$$

The first condition is an eigenvalue problem; to obtain $\check{b}^*$, we have to find the eigenvalues $\lambda$ of the matrix $(H^{-\frac{1}{2}})' C^{-1} \beta \beta' C^{-1} H^{-\frac{1}{2}}$. By defining $x := (H^{-\frac{1}{2}})' C^{-1} \beta$, the matrix can be re-written as $xx'$ and it becomes apparent that the matrix is symmetric and of rank one. Due to the latter, there can be only one non-zero eigenvalue. Hence, this value is identical to the trace of the matrix $xx'$:

$$\text{EV}(xx') = \text{tr}(xx') = \text{tr}(x'x) = x'x,$$

where $\text{EV}(\cdot)$ denotes the eigenvalue and $\text{tr}(\cdot)$ is the trace-operator. Replacing the eigenvalue in the eigenvalue problem, we get:

$$(xx' - x'xI)\check{b}^* = 0.$$

Obviously, $\check{b}^* = kx$ is a solution to this problem. Again due to the rank of $xx'$, it is also the only solution. To recover the solution in the original problem, we have to reverse the transformation:

$$b^* = H^{-\frac{1}{2}}\check{b}^* = kH^{-\frac{1}{2}}x = kH^{-\frac{1}{2}}H^{-\frac{1}{2}}C^{-1}\beta = kH^{-1}C^{-1}\beta = k(CH)^{-1}\beta$$

The set of all solutions is then obtained by letting $k$ vary. □

The following lemma helps to identify the marginal effect vectors that maximize the surplus generated by a performance measure. We use it to prove Proposition 2.

**Lemma 4.** *A performance measure that maximizes the surplus*

$$\phi\left(b, \frac{||b||^2}{\rho^2}\right) = \frac{1}{4}\frac{b'C^{-1}\beta\beta'C^{-1}b}{b'(C^{-1} + 2\frac{r}{\rho^2}I)b} \tag{15}$$

*has a marginal effect vector from the set*

$$\left\{ b \,\middle|\, b = k\left(I + \frac{2r}{\rho^2}C\right)^{-1}\beta \text{ with } k \in I\!R \right\}.$$

*Proof.* Using the definition of the signal-to-noise ratio, we can replace $\sigma^2$ in the surplus formula (6) by $\frac{b'b}{\rho^2}$. Hence, the surplus amounts to: $\frac{1}{4}\frac{b'C^{-1}\beta\beta'C^{-1}b}{b'(C^{-1}+2\frac{r}{\rho^2}I)b}$. The matrix $\left(C^{-1} + \frac{2r}{\rho^2}I\right)$ is symmetric and positive definite. Thus, we can apply Lemma 3 with $H = \left(C^{-1} + \frac{2r}{\rho^2}I\right)$ and find that the set of maximizers to be: $\left\{ b \,\middle|\, b = k\left(I + \frac{2r}{\rho^2}C\right)^{-1}\beta \text{ with } k \in I\!R \right\}$. □

# B Computations for the example

Equation (6) from Lemma 2 directly supplies us with a general formula to compute the surplus that can be generated using a performance measure $(b, \sigma^2)$. We set the Arrow-Pratt measure of risk aversion and the variance of the error term to the values in the example, $\sigma^2 = 1$ and $r = 0.5$. If we then plug in the marginal effects of the aligned measure $P^{\mathrm{A}}$, $b^{\mathrm{A}} = (2,1)'$, we obtain the following surplus:

$$\phi^{\mathrm{A}} = \frac{1}{4}\frac{\left((2,1)\begin{pmatrix}\frac{2}{10} & 0\\ 0 & 2\end{pmatrix}\begin{pmatrix}2\\1\end{pmatrix}\right)^2}{(2,1)\begin{pmatrix}\frac{2}{10} & 0\\ 0 & 2\end{pmatrix}\begin{pmatrix}2\\1\end{pmatrix}+1} = \frac{\left(\frac{14}{10}\right)^2}{\frac{28}{10}+1} = \frac{49}{95} \approx 0.52 \tag{16}$$

If we use the marginal effects of the unaligned performance measure $P^{\mathrm{NA}}$, $b^{\mathrm{NA}} = (1,2)'$, we get:

$$\phi^{\mathrm{NA}} = \frac{1}{4}\frac{\left((1,2)\begin{pmatrix}\frac{2}{10} & 0\\ 0 & 2\end{pmatrix}\begin{pmatrix}2\\1\end{pmatrix}\right)^2}{(1,2)\begin{pmatrix}\frac{2}{10} & 0\\ 0 & 2\end{pmatrix}\begin{pmatrix}1\\2\end{pmatrix}+1} = \frac{\left(\frac{2}{10}+\frac{20}{10}\right)^2}{\frac{2}{10}+\frac{80}{10}+1} = \frac{121}{230} \approx 0.53. \tag{17}$$

## C Proof of Proposition 2

By Lemma 4, the surplus is maximized by a performance measure if the marginal effect vector is in the set $\left\{ b \,\middle|\, b = k\left(I + \frac{2r}{\rho^2}C\right)^{-1}\beta \text{ with } k \in I\!R \right\}$. For an arbitrary $\beta$, alignment requires that $b = \gamma\beta$. So, aligned performance measures maximize the surplus if and only if $k\left(I + \frac{2r}{\rho^2}C\right)^{-1} = \gamma I$ or $(2\frac{r}{\rho^2}C + I) = I\kappa$, where $\kappa$ is any real number. This is equivalent to $C = I\frac{\kappa-1}{2}\frac{\rho^2}{r}$ for any $\kappa$. In other words, costs must be a multiple of the identity matrix: $C = cI$, where $c$ has to be positive because $C$ has to be positive definite.

## D Proof of Proposition 3

Ignoring for a moment the restriction that performance measures should have the same controlled and uncontrolled variation, we can use Lemma 4 to obtain the set of marginal effect vectors for performance measures that maximize the surplus:

$$\left\{ b \,\middle|\, b = k\left(I + \frac{2r}{\rho^2}C\right)^{-1}\beta \text{ with } k \in I\!R \right\}.$$

Let $\xi$ be the controlled variation under consideration, then a subset of this set can be obtained by setting $k = \pm\frac{\xi}{\sqrt{\beta'(2\frac{r}{\rho^2}C+I)^{-2}\beta}}$. The elements of this subset still maximize the surplus but they also have controlled variation $\xi$. Thus, a performance measure from the set with a respective $k$ and uncontrolled variation $\sigma^2$ maximizes surplus in the class of all performance measures with controlled variation $\xi$ and uncontrolled variation $\sigma^2$.

## E Risk minimization component

Feltham and Wu (2000) suggest reflecting the uncertainty of a performance measure by a *risk minimization component* (defined on page 164 below equation (9) in their article):

$$R := r \cdot S_l e^*,$$

where $e^*$ is the effort vector in the first-best situation, i.e., when effort $e$ can be legally enforced, and $S_l$ is a sensitivity measure defined in Equation (6) by Feltham and Wu. Notice that their Equation (6) reduces to $S_l = \hat{\mu}_l$ in the case considered here, which is nothing but the (transposed) marginal effect vector standardized by the uncontrolled variation: $\hat{\mu}_l = \frac{b'}{\sigma}$. The first best effort vector equals the benefit marginal effect vector $e^* = \beta$ (see middle of page 160 in Feltham and Wu 2000). Overall the risk minimization component is:

$$R = \frac{r}{\sigma} \cdot b'\beta. \tag{18}$$

By equation (6), the surplus generated by a performance measure in the case considered by Feltham and Wu ($C = \frac{1}{2}I$) is:

$$\Phi(b, \sigma^2) = \frac{1}{2}\frac{b'\beta\beta'b}{b'b(1 + r\sigma^2)}.$$

Using the definition of $R$, this is equivalent to

$$\Phi(b,\sigma^2) = \frac{1}{2}\frac{R^2}{b'b(1+r\sigma^2)} \cdot \frac{\sigma^2}{r^2} = \frac{1}{2} \cdot \frac{1}{(\frac{r^2}{\sigma^2}+r^3)} \cdot \frac{R^2}{b'b}.$$

Holding $R$ constant, the surplus no longer depends on $\beta$. Consequently, the generated surplus is independent from the relationship between $\beta$ and $b$ once the risk-minimization component $R$ is fixed.

# F Re-parametrization

In the model, performance measures, costs and benefit are stable and common knowledge. The same model, however, can be parameterized differently by re-defining effort since the units in which we express effort in the model are to some degree arbitrary. In fact, it is possible to re-parameterize the model by rescaling effort such that costs become independent and identical. Take effort $e$ and transform it to obtain $\tilde{e} = C^{\frac{1}{2}}e$, where $C^{\frac{1}{2}}$ is a decomposition of the positive definite matrix $C$ such that $(C^{\frac{1}{2}})'C^{\frac{1}{2}} = C$. After re-parameterization, costs are $\tilde{C}(\tilde{e}) = \tilde{e}'\tilde{e} = e'Ce = C(e)$ and hence independent and identical. Expressing the marginal effect vector for benefit and performance measure in the rescaled units, we get: $\tilde{b} = (C^{-\frac{1}{2}})'b$ and $\tilde{\beta} = (C^{-\frac{1}{2}})'\beta$. Since costs under re-parameterization are independent and identical, aligned performance measures are more valuable than unaligned measures with the same signal-to-noise ratio, $\tilde{\rho} = \frac{||\tilde{b}||}{\sigma}$ by Proposition 1.

The re-parameterization does not affect whether a performance measure is aligned: a measure that is aligned before re-parameterization, $b = \gamma\beta$, is also aligned after re-parameterization, $\tilde{b} = (C^{-\frac{1}{2}})'b = \gamma(C^{-\frac{1}{2}})'\beta = \gamma\tilde{\beta}$. On the other hand, the controlled variation (and hence the signal-to-noise ratio) is not the same after re-parameterization:

$$||b|| = \sqrt{b'b} \neq \sqrt{b'C^{-1}b} = ||\tilde{b}|| = ||b||_{C^{-1}},$$

where $||\cdot||_{C^{-1}}$ is an inverse-cost norm that is used instead of the Euclidean norm. This explains why alignment is valuable in the re-parameterized model but not in the original model: aligned measures are compared to different measures in the two models. The set of measures considered in the original model depends only on the parameters of the measure (the variation under control and beyond control of the agent) while the set of measures in the re-parameterized model also depends on the relative costs. So, either more alignment is not necessarily preferable (original model) or costs are taken into account in the signal-to-noise ratio (re-parameterized model). Overall, re-parameterization cannot alter the structure of the underlying model and does not alter the central insight of Proposition 2: if risk is defined in the sense of Baker (2002) as the signal-to-noise ratio that is not adjusted for effort costs and if this signal-to-noise ratio is held constant, more distortion may increase surplus.