

Distorted Performance Measures and Dynamic Incentives

Oddvar M. Kaarbøe* and Trond E. Olsen†

December 10, 2004

Abstract

Incentive contracts must typically be based on performance measures that do not exactly match agents' true contribution to principals' objectives. Such misalignment may impose difficulties for effective incentive design. We analyze to what extent implicit dynamic incentives such as career concerns and ratchet effects alleviate or aggravate these problems. Our analysis demonstrates that the interplay between distorted performance measures and implicit incentives implies that career and ratchet effects have real effects, that career and monetary incentives may be complements, and that stronger ratchet effects or more distortion may increase optimal monetary incentives.

*Programme for Health Economics (HEB), Department of Economics, University of Bergen. Fosswinkelsg. 6, N-5007 Bergen. Norway. Email: oddvar.kaarboe@econ.uib.no
URL <http://www.econ.uib.no/stab/oddvar.html>

†Department of Finance and Management Science, Norwegian School of Economics and Business Administration, and Programme for Health Economics (HEB). Email: trond.olsen@nhh.no

1 Introduction

A general problem for designing incentive schemes is that available performance measures seldom capture precisely agents' true contributions to principals' objectives. Performance measures are typically influenced by stochastic factors that agents can't control, and they often do not reflect all aspects that principals care about. For instance, quantitative performance measures often neglect important qualitative (soft) aspects of an agent's performance. Such measures are distorted from, or 'not well aligned with', the principal's true objectives. As is well known, such misalignments may impose severe difficulties for effective incentive design. (Holmstrom and Milgrom 1991; Baker 1992; Feltham and Xie 1994; and Baker 2002.)

Baker (2002) argues that an understanding of how distorted performance measures affect the design of incentive contracts may explain several issues and puzzles in the literature; including (i) why high-signal-to-noise ratio performance measures may receive low weight in an incentive scheme, (ii) how the distinction between paying for "inputs" versus paying for "outputs" can be interpreted, and (iii) why seemingly informative performance measures degrade, (Baker 2002, pp. 738-40). The latter issue is illustrated by a school system that administers standardized tests to its students, but does not use the scores to motivate teachers. A reason for not including these seemingly informative test scores as a performance measure in an incentive system, is that teachers will then have incentives to "teach to the test", and may thus engage in dysfunctional behavior that increases the performance measure without increasing the school's real objective.

Here we want to point out that, while it certainly is true that incentives to "teach to the test" are affected by direct monetary rewards, it may nevertheless well be the case that teachers face incentives to engage in this kind of behavior even if such direct monetary rewards are absent. Good test scores may for example give the school administration a signal that the teacher is valuable, and result in future salary increases. Or, test scores may be used as a criterion to allocate teachers to different classes. A complete understanding of how distorted performance measures affect overall incentive design requires that implicit incentives are also taken into account.

In this paper we analyse the interplay between implicit dynamic incentives and explicit incentives based on distorted performance measures. We examine to what extent such implicit incentives alleviate or aggravate problems related to distorted measures. Implicit incentives of this form arise when explicit contracts can be renegotiated as time unfolds. Hence, implicit incentives reflect the fact that future periods' pay depends on today's performance. If today's performance improves the agent's position in the labor market, career concerns are present, (Fama 1980; Holmstrom 1982; and Gibbons and Murphy 1992). Ratchet effects are present if better performance today implies a tougher performance standard tomorrow, (Weitzman 1976).

The analysis is based on a dynamic version of the model developed in Baker (2002). First we consider the case where the principal can provide incentives on

a verifiable, but distorted, performance measure (z). In addition some information about the agent's performance is provided to the principal (and the market) through the non-verifiable value measure (y) that reflects how the agent's performance contributes to the principal's true objective. In this case implicit incentives are related both to the distorted and the undistorted performance measures (and hence the degree of misalignments between them).

By using this model we show that both career and ratchet effects do have real effects; neither can costlessly be neutralized by monetary incentives. Furthermore we find, contrary to what is found models with non-distorted performance measures (e.g. Gibbons 1987; Meyer, Olsen, and Torsvik 1996; and Meyer and Vickers 1997), that stronger ratchet effects may increase optimal monetary incentives. The intuition behind this result is that the ratchet effect works through both the true value measure and through the verifiable performance measure. If it is the case that stronger ratchet effects reduce net implicit incentives on the true value measure, then the principal should increase monetary incentives on the verifiable measure to compensate for the former effects. Finally we notice that this dynamic model reproduces some of the results of the static version (Baker 2002); that better alignment between the performance measures increases monetary incentives, and that the better aligned the performance measure is with the true value, the higher is the (total) surplus among the principal and the agent. The first of these two results is however not trivial since better alignments strengthens the ratchet elements. This effect tends to lower monetary incentives. We can however show that this latter effect will never dominate the direct effect of better alignments, and hence that monetary incentives do increase with better alignments.

It is often the case that in addition to verifiable (and distorted) performance measures, there are other non-verifiable measures that may yield valuable information about the agent's performance. A typical case is one where quantity aspects are verifiable but quality aspects are not, yet some information about these quality aspects is observable for the relevant parties. Such information may be hard or impossible to verify in a court, but may be used by principals and peers to assess agents' abilities and performance, and hence induce implicit incentives for agents to exert effort.

We also consider such a setting, and show that some new issues arise.¹ In particular, we point out the following features. First we show that career and monetary incentives may be complements rather than substitutes, second that incentives may increase with more distortion in the (verifiable) performance measure, and finally that it may well be advantageous (in terms of total surplus) that the verifiable performance measure is distorted relative to the measure of true value.

The intuition for the first result is that in a setting where efforts on quality aspects are rewarded by strong implicit career incentives, and where efforts on quality

¹In addition to measures y and z , there is now a third non-verifiable measure q . We may think of y reflecting the true mix of quality and quantity aspects that the principal cares about, z being a verifiable measure of quantity aspects, and q being a non-verifiable measure of quality aspects.

and quantity aspects are substitutes for the agent, the principal may have to match strong implicit incentives on 'quality' with strong explicit incentives on 'quantity'.

We also note that this mechanism behind the complementarity result is different from the one obtained in Dewatripont, Jewitt, and Tirole (1999). In their analysis a complementarity effect between monetary and career incentives may arise when there is a technological complementarity between effort and talent in the way they affect performance.

To see the intuition behind the second result (that incentives may increase with more distortion), note that career concerns and monetary incentives are always substitutes when only one verifiable (and non-distorted) performance measure is available, and effort affects this measure additively (Gibbons and Murphy 1992). In such a case the principal never needs to match strong career incentives with strong monetary incentives. But if it is the case that additional non-verifiable performance measures are available, and career concerns and monetary incentives are complements, more distortion may imply that the principal has to increase monetary incentives to maintain the appropriate balance of the agent's effort among the tasks. Hence in this case, and opposed to what is suggested in Kerr (1975), it may be appropriate to "reward for A, while hoping for B".²

An intuition for the third result (that distortion may be advantageous) goes as follows. If some non-verifiable measure of quality aspects is not aligned with the true value, and implicit incentives on this measure induces the agent to focus on these quality aspects, then it may be advantageous that explicit incentives can be used to induce efforts on quantity aspects rather than on a balanced mix of both aspects. This is just to say that it may be advantageous that the verifiable measure is not perfectly aligned with the measure of true value.

The paper is organized as follows. In section 2 the basic model is outlined, while the optimal contracts are derived in section 3. In section 4 we consider the case where an additional non-verifiable measure provides some information about the agent's performance. First we analyze under which conditions career concerns and monetary incentives are complements (section 4.1). Then, in section 4.2 we show that incentives may increase with more distortion. Finally, in section 4.3 we show that it may be advantageous (in terms of total surplus) that the verifiable performance measure is distorted relative to the measure of true value. Section 5 provides some concluding remarks.

2 The Model

There is one agent, n tasks, and two periods. The model is a dynamic variant of the framework used by Feltham and Xie (1994) and Baker (2002) to analyze distorted performance measures. In each period the agent privately supplies his choice of effort $\mathbf{a}_t = (a_{t1}, a_{t2}, \dots, a_{tn})$ on the n tasks. The agent's choice of efforts determines

²A similar result is shown in Schnedler (2003) in the case where the principal does not know the set of actions the agent can choose.

the agent's *total* contribution to the principal, denoted by y_t . That is, y_t reflects everything the principal cares about, except for wages, in period t . We assume that no contract on y can be enforced in court because it is prohibitively costly to specify this outcome ex ante in such a way that it can be verified by a third party ex post. We do however assume that all parties—insiders as well as outsiders—observe the y -signal ex post, and favorable realizations of this signal improves the agent's standing on the job market. Hence, some incentives are provided for y through career concerns.

Let

$$y_t = h'\eta + \mathbf{f}\mathbf{a}_t + \varepsilon_t,$$

where $\mathbf{f} = \{f_1, f_2, \dots, f_n\}$ is an n -dimensional vector of marginal products of the agent's efforts, $\mathbf{f}\mathbf{a}_t = f_1 a_{t1} + \dots + f_n a_{tn}$ denotes the scalar product, and $\varepsilon_t \sim N(0, \sigma_y^2)$ represents random effects. η is the agent's unknown ability. The ability η is drawn at the beginning of the first period from an independent normal distribution with mean m_0 and variance σ_η^2 . The agent's ability has productivity h' for the principal.

While the agent's total contribution is not verifiable, there is a performance measure z that is verifiable, so monetary incentives can be provided through this signal. Incentives on this signal serves as a means to increase the agent's total contribution for the principal. Let

$$z_t = \eta + \mathbf{g}\mathbf{a}_t + \zeta_t,$$

where $\mathbf{g} = \{g_1, g_2, \dots, g_n\}$ is an n -dimensional vector of the marginal products of actions on the verifiable performance measure and $\zeta_t \sim N(0, \sigma_z^2)$ is the effect of uncontrollable events. Let ζ be independent of ε and of η . Since z is verifiable all parties observe it. We summarize the information given to the principal by the signals in period t by $\mathbf{x}_t = (y_t, z_t)$. Furthermore, we assume that the principal offers the agent linear payments $w_t = A_t + \alpha_t z_t$.³

The agent which is risk-averse privately chooses actions a_{ti} , $i = 1, \dots, n$. The private cost of effort in monetary units is denoted $c(\mathbf{a}_t)$, and is (for simplicity) assumed to be a quadratic expression. For most of the analysis (section 4 is the exception) we assume that effort costs are independent across tasks, and hence given by $c(\mathbf{a}_t) = \sum_{i=1}^n \frac{a_{ti}^2}{2}$.

The agent's utility function is exponential, and there is no discounting:

$$u = -\exp\left\{-r \sum_{t=1}^2 [w_t - c(\mathbf{a}_t)]\right\},$$

where the coefficient $r \geq 0$ measures the agent's risk aversion. With linear compensation, exponential utility, and normal random variables, the agent's certainty equivalent is $CE = \sum_t [\mathbf{E}w_t - c(\mathbf{a}_t)] - \frac{r}{2} \text{var}(w_1 + w_2)$, where \mathbf{E} is the expectation operator.

³The focus on linear contracts can be justified by appeal to a richer dynamic model in which linear payments are optimal Holmstrom and Milgrom (1987).

The principal is risk neutral, has net benefit in period t given by $y_t - w_t$, and can observe neither the actions taken by the agent nor his ability. She only observes the signals $\mathbf{x}_t = (y_t, z_t)$ and may use it in every period to update her beliefs about the agent's ability.

The parties cannot commit not to renegotiate contracts. The second-period contract will therefore be efficient, given the information available at that time.

3 Optimal Contracts

We first characterize the optimal contract in the second, and last period. Note that there are no career incentives in this period, and hence the optimal incentives in period 2 correspond to the optimal bonus in the one-period model.

Second period. The true expected value for the principal is $\mathbf{E}(y_2 \mid \mathbf{x}_1) = h' \mathbf{E}(\eta \mid \mathbf{x}_1) + \mathbf{f}\mathbf{a}_2$, where the expectation is conditional on the signals $\mathbf{x}_1 = (y_1, z_1)$ observed in period 1. The expected value of the verifiable measure is $\mathbf{E}(z_2 \mid \mathbf{x}_1) = \mathbf{E}(\eta \mid \mathbf{x}_1) + \mathbf{g}\mathbf{a}_2$, where $\mathbf{E}(\eta \mid \mathbf{x}_1)$ reflects the updated belief about the agent's ability, and is given by

$$\mathbf{E}(\eta \mid \mathbf{x}_1) = \mathbf{E}\eta + \rho_z(z_1 - \mathbf{E}z_1) + \rho_y(y_1 - \mathbf{E}y_1). \quad (1)$$

The exact expressions for the regression coefficients $\rho_i = \frac{\partial}{\partial i} \mathbf{E}(\eta \mid \mathbf{x}_1)$, $i = y, z$ are contained in Appendix A. Here we simply note that $\rho_i \in [0, 1]$ and depends on the noise terms σ_i^2 , $i = \eta, y, z$, as well as the productivity parameter of ability h' . Furthermore we note that if the z -signal is more noisy than the y -signal (i.e. $\sigma_z^2 > \sigma_y^2$), more weight is put on y relative to z in estimating the agent's ability.

The certainty equivalent for the agent in period 2 is

$$CE_2 = \mathbf{E}w_2 - c(\mathbf{a}_2) - \frac{r}{2} \text{var}(w_2 \mid \mathbf{x}_1) = A_2 + \alpha_2 \mathbf{g}\mathbf{a}_2 - \sum_i \frac{a_{i2}^2}{2} - \frac{r}{2} \alpha_2^2 \sigma_{2c}^2;$$

where $\sigma_{2c}^2 := \text{var}(z_2 \mid \mathbf{x}_1) = \text{var}(\eta \mid \mathbf{x}_1) + \text{var}(\zeta)$. (Again we refer to Appendix A for the exact expression of the conditional variance.) The agent chooses effort \mathbf{a}_2 to maximize this certainty equivalent, and this yields

$$\mathbf{a}_2 = \alpha_2 \mathbf{g}.$$

Total expected surplus in period 2 is

$$TCE_2 = \mathbf{E}(y_2 \mid \mathbf{x}_1) - c(\mathbf{a}_2) - \frac{r}{2} \alpha_2^2 \sigma_{2c}^2 = h' \mathbf{E}(\eta \mid \mathbf{x}_1) + \mathbf{f}\mathbf{a}_2 - \sum_i \frac{a_{i2}^2}{2} - \frac{r}{2} \alpha_2^2 \sigma_{2c}^2 \quad (2)$$

By maximizing this w.r.t. α_2 , and taking into account the agent's response, we obtain the optimal incentive for period 2. It is given by

$$\alpha_2^* = \frac{\mathbf{f}\mathbf{g}}{|\mathbf{g}|^2 + r\sigma_{2c}^2}. \quad (3)$$

In this expression $\mathbf{fg} = f_1g_1 + f_2g_2 + \dots + f_n g_n$ is the scalar product of the \mathbf{f} and \mathbf{g} vectors, and $|\mathbf{g}| = \sqrt{\sum_{i=1}^n g_i^2}$ is the length of the vector of marginal products on the performance measure. Note that if $\mathbf{f} = \mathbf{g}$ (i.e., the performance measure and the principal's valuation of the marginal products are perfectly aligned) then $\alpha_2^* = \frac{|\mathbf{f}|}{|\mathbf{f}| + r\sigma_{2c}^2}$. In this case the optimal incentive is increasing in the "length" of the vector \mathbf{f} . This follows since the length is a measure of the contribution of the agent's action to the principal's value relative to the contribution of noise in the production function. Furthermore, the optimal incentive is decreasing in the agent's risk aversion (r) and in the variance of outcome (σ_{2c}^2). When $\mathbf{f} \neq \mathbf{g}$, the optimal incentive is reduced, relative to the case where $\mathbf{f} = \mathbf{g}$, since paying on z is less valuable for increasing y . If e.g. \mathbf{f} and \mathbf{g} are orthogonal, i.e. $\mathbf{fg} = \mathbf{0}$, then $\alpha_2^* = 0$, since the incentives created by paying on z are useless for increasing y . Finally we note that if $\mathbf{fg} < 0$, i.e., \mathbf{g} points 'opposite' to \mathbf{f} , the optimal bonus is negative, which may not be feasible. We will assume $\mathbf{fg} \geq \mathbf{0}$.

The sharing of the total surplus TCE_2 will be determined by the parties' bargaining strength (and the terms specified in the initial contract). We assume that the agent has some bargaining power and hence can obtain some share of the surplus. The agent's bargaining strength may for instance be due to outside principals competing for his services in period 2. If the agent can negotiate for himself some share s of the expected surplus TCE_2 (at the start of period 2), then the fixed wage component A_2 will be adjusted to reflect the information (\mathbf{x}_1) revealed in period 1 about the agent's ability as follows:

$$A_2 = (h - \alpha_2^*)\mathbf{E}(\eta | \mathbf{x}_1) + const$$

where $h = sh'$ and the constant represent terms that do not depend on \mathbf{x}_1 . The formula follows from the fact that the agent's expected equilibrium payoff in period 2 must equal the share s of that period's total surplus, and hence that $sTCE_2 = A_2 + \alpha_2^*\mathbf{E}(z_2 | \mathbf{x}_1) + k$, where k does not depend on \mathbf{x}_1 . Substituting from the expression for TCE_2 above, we see that the stated formula for A_2 must hold.

The second-period wage contract offered to the agent is thus:

$$w_2(\mathbf{x}_1) = (h - \alpha_2^*)\mathbf{E}(\eta | \mathbf{x}_1) + \alpha_2^*z_2 + const$$

where the updated expected ability $\mathbf{E}(\eta | \mathbf{x}_1)$ for the agent is given by (1).

First period. After characterizing the second-period wage contract we turn to period one. First of all we notice that since the second period compensation depends on the first period signals, $\mathbf{x}_1 = (y_1, z_1)$, the agent has incentives to exert effort in the first period to increase his market value. The agent thus chooses effort according to

$$\begin{aligned} & \max_{\mathbf{a}_1} \{ \alpha_1 \mathbf{g} \mathbf{a}_1 - c(\mathbf{a}_1) + (h - \alpha_2^*)\mathbf{E}(\eta | \mathbf{x}_1) + const \} \\ \Rightarrow & \mathbf{a}_1 = (\alpha_1 + \beta_z)\mathbf{g} + \beta_y\mathbf{f}, \text{ where } \beta_z = (h - \alpha_2^*)\rho_z, \quad \beta_y = (h - \alpha_2^*)\rho_y. \end{aligned}$$

In the last expression β_i is the implicit incentive on signal $i = y, z$. We see that this consists of a positive career element ($h\rho_i$) and a negative ratchet element ($\alpha_2^*\rho_i$).

The net implicit incentive β_i may be positive or negative, depending on the sign of $h - \alpha_2^*$.

To characterize optimal first-period incentives consider the total intertemporal surplus, and note that the variance of total wages may be written as

$$\begin{aligned} \text{var}(w_1 + w_2) &= \text{var}(\alpha_1 z_1 + \beta_y y_1 + \beta_z z_1 + \alpha_2^* z_2) \\ &= \text{var}((\tilde{\alpha}_1 + \alpha_2^* \rho_z) z_1 + h \rho_y y_1 \\ &\quad + \alpha_2^* [z_2 - \rho_y y_1 - \rho_z z_1]) \end{aligned}$$

where $\tilde{\alpha}_1 = \alpha_1 + \beta_z$ is the *effective* incentive on the z -variable. The stochastic variables in the two last lines are uncorrelated, and the variance of the latter (in square brackets) is $\sigma_{2c}^2 = \text{var}(z_2 \mid y_1, z_1)$. The total intertemporal surplus may therefore be written as $TCE = TCE_1 + TCE_2^*$, where TCE_2^* is the (equilibrium) second-period surplus and TCE_1 is given by

$$TCE_1 = E y_1 - c(\mathbf{a}_1) - \frac{r}{2} ((\tilde{\alpha}_1 + \alpha_2^* \rho_z)^2 \sigma_{1z}^2 + (h \rho_y)^2 \sigma_{1y}^2 + 2(\tilde{\alpha}_1 + \alpha_2^* \rho_z) h \rho_y \sigma_{1yz}) \quad (4)$$

where $\sigma_{1z}^2 = \text{var}(z_1)$, $\sigma_{1y}^2 = \text{var}(y_1)$ and $\sigma_{1yz} = \text{cov}(y_1, z_1)$.⁴

Maximizing this expression, taking account of the agent's effort choice $\mathbf{a}_1 = \tilde{\alpha}_1 \mathbf{g} + \beta_y \mathbf{f}$, we see that the optimal effective incentive in period 1 is given by

$$\tilde{\alpha}_1^* = \alpha_1^* + \beta_z = \frac{(1 - \beta_y) \mathbf{f} \mathbf{g} - r [\alpha_2^* \rho_z \sigma_{1z}^2 + h \rho_y \sigma_{1yz}]}{|\mathbf{g}|^2 + r \sigma_{1z}^2}. \quad (5)$$

A number of observations follow directly from this.

1. The *career element* ($h \rho_z$) in the z -variable does not appear in the formula for the optimal effective incentive and hence has no real effects. The career element in this variable can be adjusted by monetary incentives and has no implications for effective incentives and for welfare. This parallels the observation in Meyer and Vickers (1997) that career incentives have no welfare effects in a setting where the principal can contract directly on the true value measure (i.e. y in this model)
2. The *ratchet element* ($\alpha_2^* \rho_z$) in the z -variable does have real effects. The first term in the square bracket accounts for this ratchet element. This term lowers effective incentives. It is costly—in terms of risk costs—to compensate with monetary incentives for the ratchet element, and hence effective incentives are optimally reduced. This is in line with results for settings where one can contract directly on the measure of true value (e.g. Gibbons 1987, Meyer, Olsen, and Torsvik 1996).
3. Both the career element and the ratchet element in the y -variable do have real effects (recall that $\beta_y = (h - \alpha_2^*) \rho_y$). A higher career element in y reduces

⁴We have $\sigma_{1z}^2 = \sigma_z^2 + \sigma_\eta^2$, $\sigma_{1y}^2 = \sigma_y^2 + \sigma_\eta^2$ and $\sigma_{1yz} = \sigma_\eta^2$ in the given specification.

effective incentives on z . *Career incentives (on y as well as on z) and monetary incentives are thus substitutes.* The latter property is in line with findings for the setting where one can contract directly on the measure of true value (Gibbons and Murphy 1992), but the fact that career elements have real effects is at variance with results from that setting (Meyer and Vickers 1997). In the present model the career element in the y -variable works via two channels. First it increases net incentives β_y on y , and when \mathbf{f} and \mathbf{g} are to some extent aligned ($\mathbf{fg} > 0$), incentives on z can then be reduced. Second, a strong career element on y will increase the variance of payments, and thus the risk costs. When the y, z - variables are (positively) correlated, this variance can be reduced by reducing the explicit incentive on the z - variable

4. A stronger ratchet element in the y -variable will (all else equal) increase optimal effective (and monetary) incentives on z . The reason is that the stronger ratchet element reduces net incentives β_y on y , and when \mathbf{f} and \mathbf{g} are to some extent aligned ($\mathbf{fg} > 0$), incentives on z should be increased to compensate for the reduced incentives on y . The fact that stronger ratchet effects may increase monetary incentives is quite opposite to what one finds in settings where contracting on a non-distorted measure is feasible (e.g. Gibbons 1987, Meyer and Vickers 1997).
5. Comparing effective incentives over time ($\tilde{\alpha}_1^*$ and α_2^*), we see that these are highest in period 2 when net implicit incentives on y are non-negative ($\beta_y \geq 0$). Monetary incentives are then also highest in period 2, since net implicit incentives on z are non-negative as well ($\beta_z \geq 0$).

We summarize these results in the following Proposition.

Proposition 1 *For independent efforts we have:*

- (i) *While the career element ($h\rho_z$) in the performance measure (z) can be costlessly neutralized by monetary incentives and has thus no effects on total surplus, the ratchet element ($\alpha_2^*\rho_z$) in this measure lowers effective incentives and thus does have real effects.*
- (ii) *Both the career element and the ratchet element in the value measure (y) do have real effects. A higher career element in y reduces effective incentives on z . Career incentives (on y as well as on z) and monetary incentives are thus substitutes.*
- (iii) *A stronger ratchet element in the y -variable will (all else equal) increase optimal effective (and monetary) incentives on z .*
- (iv) *Given non-negative net implicit incentives ($\beta_y, \beta_z \geq 0$), then optimal monetary incentives are increasing over time, i.e., $\alpha_1^* < \alpha_2^*$.*

Following Baker (2002) we can use the angle θ between vectors \mathbf{f} and \mathbf{g} , defined by $\cos \theta = \frac{\mathbf{fg}}{\|\mathbf{f}\|\|\mathbf{g}\|}$, as a measure of how distorted or misaligned the performance measure z is from the measure of true value y . In a static setting –and here for period 2– one sees

that incentives on the performance measure z are stronger the better aligned are the two measures. In a dynamic setting the relationship is more complicated. From the formula (5) for the effective first-period incentive $\tilde{\alpha}_1^*$, we see that there are both direct and indirect effects associated with better alignment. First there is a direct positive effect in that \mathbf{fg} gets larger. (We keep $|\mathbf{f}||\mathbf{g}|$ fixed and consider only parameters that yield non-negative incentives, so in particular $\beta_y < 1$.) But second, there are indirect effects working via the ratchet elements. This is so because better alignment increases second-period incentives (α_2^*), and this in turn strengthens the ratchet elements. There are two ratchet elements; one associated with the z - measure ($\alpha_2^*\rho_z$) and one associated with the y - measure ($\alpha_2^*\rho_y$, which enters through $\beta_y = (h - \alpha_2^*)\rho_y$). The latter increases the first-period optimal effective incentive $\tilde{\alpha}_1^*$ (by reducing the net implicit incentive β_y on y , which is optimally compensated by a stronger monetary incentive on z). But the stronger ratchet element associated with z reduces the effective incentive $\tilde{\alpha}_1^*$. Hence we see that better alignment induces opposing effects on the first-period effective incentive.

Although there are opposing effects generated by dynamic implicit incentives, it turns out that better alignment does in fact increase effective incentives also in the first period, at least for all parameters that yield non-negative effective incentives in this model. We verify this in Appendix A. Note that this implies that monetary incentives must also increase (and by even more) since the implicit incentive β_z is reduced.

In Appendix A we also verify the intuitively reasonable result that the equilibrium surplus (TCE) is also higher the better aligned is the performance measure with the measure of true value. Thus we have:

Proposition 2 *For independent efforts and parameters that yield non-negative effective incentives we have: As the performance measure z gets better aligned with (less distorted from) the measure of value y , (i) optimal effective and monetary incentives in both periods increase, and (ii) the total surplus increases.*

These results show that to the extent that design of performance measures is feasible, it is (all else equal) optimal to construct or choose a measure that is least distorted relative to the measure of true value. As we shall see in the next section, this is however generally true only when such performance measures are verifiable.

4 Additional Non-Verifiable Measures

It is often the case that in addition to verifiable (and distorted) performance measures, there are other non-verifiable measures that may yield valuable information about the agent's performance. A typical case is one where quantity aspects are verifiable but quality aspects are not, yet some information about these quality aspects is observable for the relevant parties. Such information may be hard or impossible to verify in a court, but may be used by principals and peers to assess agents' abilities and performance, and hence induce implicit incentives for agents to exert effort.

We now consider such a setting, and show that some new issues arise. In particular, we point out three new features: (i) that career and monetary incentives may be complements rather than substitutes, (ii) that incentives may increase with more distortion in the (verifiable) performance measure, and (iii) that it may well be advantageous (in terms of total surplus) that the verifiable performance measure (z) is distorted relative to the measure of true value (y). The intuition for (i) is that in a setting where efforts on quality aspects are rewarded by strong implicit career incentives, and where efforts on quality and quantity aspects are substitutes for the agent, the principal may have to match strong implicit incentives on 'quality' with strong explicit incentives on 'quantity'.

To see the intuition behind the second result, note that career concerns and monetary incentives always are substitutes when only one verifiable (but non-distorted) performance measure is available. In such a case the principal never needs to match strong career incentives with strong monetary incentives. The result (ii) follows when more distortion implies that the principal has to increase monetary incentives to maintain the appropriate balance of the agent's effort among the tasks.

An intuition for result (iii) is that if some non-verifiable measure (say q) of quality aspects is not aligned with the true value (y), and implicit incentives on this measure induces the agent to focus on these quality aspects, then it may be advantageous that explicit incentives can be used to induce efforts on quantity aspects rather than on a balanced mix of both aspects. This is just to say that it may be advantageous that the verifiable measure (z) is not aligned with the measure of true value (y).

To model these issues, suppose now there is an additional non-contractible 'quality' variable

$$q_t = \mathbf{k}\mathbf{a}_t + \eta + \kappa_t$$

where k is an n -dimensional vector, and κ_t is a noise term. To simplify the exposition in this section, we henceforth consider *risk-neutral* agents. On the other hand we allow for non-independence among efforts in the agent's cost function. Effort costs are

$$c(\mathbf{a}) = \frac{1}{2} \mathbf{a}' C \mathbf{a} = \frac{1}{2} \sum_{i,j} c_{ij} a_i a_j$$

where C is a symmetric matrix. (The prime denotes transpose.) We consider the three issues in turn.

4.1 Career and Monetary Incentives as Complements

In the second period the agent chooses effort \mathbf{a}_2 to maximize $\alpha_2 \mathbf{g} \mathbf{a}_2 - c(\mathbf{a}_2)$, which implies $C \mathbf{a}_2 = \alpha_2 \mathbf{g}$ (marginal cost equals marginal income for each effort component). Given this effort choice, the principal chooses the bonus α_2 on z_2 to maximize $E y_2 - c(\mathbf{a}_2) = \alpha_2 \mathbf{f}' C^{-1} \mathbf{g} - \alpha_2^2 \frac{1}{2} \mathbf{g}' C^{-1} \mathbf{g}$. The second-period (and static) optimal incentive on z is thus

$$\alpha_2^* = \frac{\mathbf{f}' C^{-1} \mathbf{g}}{\mathbf{g}' C^{-1} \mathbf{g}} > 0$$

We assume that both terms entering the fraction on the right are positive.

In period 1 the agent faces implicit incentives on the three measures y, z, q . These incentives are given by

$$\beta_i = (h - \alpha_2^*)\rho_i, \quad i = y, z, q$$

where the ρ_i 's are the regression coefficients for the conditional expectation of ability (η), given first-period observations (y_1, z_1, q_1) .⁵ The agent optimally chooses efforts such the vector of marginal costs (the cost gradient $C\mathbf{a}_1$) equals the vector of marginal benefits, i.e.

$$C\mathbf{a}_1 = \tilde{\alpha}_1\mathbf{g} + \beta_y\mathbf{f} + \beta_q\mathbf{k}, \quad \tilde{\alpha}_1 = \alpha_1 + \beta_z \quad (6)$$

Here α_1 and $\tilde{\alpha}_1$ denote as before the explicit (monetary) and the effective incentives, respectively, on the verifiable performance measure z . Maximization of the first-period surplus $h'\eta + \mathbf{f}\mathbf{a}_1 - c(\mathbf{a}_1)$ with \mathbf{a}_1 given by (6), yields the optimal first-period effective incentive

$$\tilde{\alpha}_1^* = \alpha_1^* + \beta_z = (1 - \beta_y)\frac{\mathbf{f}'C^{-1}\mathbf{g}}{\mathbf{g}'C^{-1}\mathbf{g}} - \beta_q\frac{\mathbf{k}'C^{-1}\mathbf{g}}{\mathbf{g}'C^{-1}\mathbf{g}} \quad (7)$$

We see that explicit incentives α_1^* are decreasing in β_y and in β_z , so explicit incentives and the implicit incentives on measures z and y are still substitutes.

Whether the implicit incentive β_q on the non-verifiable 'quality' measure q also is a substitute to the explicit incentive α_1^* , depends on the sign of $\mathbf{k}'C^{-1}\mathbf{g}$. The point is now that this expression may well be negative; an illustrative case is given below. In such cases we see that α_1^* and β_q are complements; a stronger implicit incentive on q implies a stronger optimal explicit incentive on z .

Rewriting the above expression (7) in terms of career and ratchet elements we have

$$\tilde{\alpha}_1^* = \alpha_1^* + (h - \alpha_2^*)\rho_z = (1 - (h - \alpha_2^*)\rho_y)\frac{\mathbf{f}'C^{-1}\mathbf{g}}{\mathbf{g}'C^{-1}\mathbf{g}} - (h - \alpha_2^*)\rho_q\frac{\mathbf{k}'C^{-1}\mathbf{g}}{\mathbf{g}'C^{-1}\mathbf{g}}$$

We may then ask whether a stronger career element (e.g. larger h) may increase effective and explicit incentives. The former is obviously the case if $-\rho_q\mathbf{k}'C^{-1}\mathbf{g} > \rho_y\mathbf{f}'C^{-1}\mathbf{g}$, and the latter holds if $-\rho_q\mathbf{k}'C^{-1}\mathbf{g} > \rho_y\mathbf{f}'C^{-1}\mathbf{g} + \rho_z\mathbf{g}'C^{-1}\mathbf{g}$. Any of these inequalities may hold (see below). Thus, *stronger career effects may well increase effective as well as explicit incentives; career incentives and explicit incentives may thus be complements.*

Example 1. The following 2-dimensional case illustrates these results. Suppose tasks 1 and 2 promote quantity and quality, respectively, and that the true respective marginal values are given by (f_1, f_2) . Suppose further that the verifiable performance measure z only reflects quantity aspects, and the non-verifiable measure q only

⁵Note that the regression coefficients are different now compared to those outlined in the former sections. In Appendix A we give the exact formulas.

reflects quality aspects. Suppose also that efforts on the two tasks are substitutes for the agent. The relevant parameters then have the following form

$$\mathbf{f}' = (f_1, f_2), \quad \mathbf{g}' = (g_1, 0), \quad \mathbf{k}' = (0, k_2), \quad C = \begin{bmatrix} 1 & \gamma \\ \gamma & 1 \end{bmatrix}, \quad 0 < \gamma < 1$$

Here we find

$$\mathbf{f}'C^{-1}\mathbf{g} = \frac{(f_1 - f_2\gamma)g_1}{1 - \gamma^2}, \quad \mathbf{k}'C^{-1}\mathbf{g} = \frac{-k_2\gamma g_1}{1 - \gamma^2}$$

The optimal first- and second-period bonuses are

$$\begin{aligned} \alpha_2^* &= \frac{\mathbf{f}'C^{-1}\mathbf{g}}{\mathbf{g}'C^{-1}\mathbf{g}} = \frac{(f_1 - f_2\gamma)g_1}{g_1^2} = \frac{f_1 - f_2\gamma}{g_1} \\ \tilde{\alpha}_1^* &= \alpha_1^* + (h - \alpha_2^*)\rho_z = (1 - (h - \alpha_2^*)\rho_y) \frac{(f_1 - f_2\gamma)}{(1 - \gamma^2)g_1} + (h - \alpha_2^*)\rho_q \frac{k_2\gamma}{(1 - \gamma^2)g_1} \end{aligned}$$

Here we see that the effective incentive $\tilde{\alpha}_1^*$ is increasing in the career incentive parameter h if $\rho_q k_2 \gamma > \rho_y (f_1 - f_2 \gamma)$. The monetary incentive α_1^* will be increasing in this parameter if $\rho_q k_2 \gamma > \rho_y (f_1 - f_2 \gamma) + \rho_z$. There are obviously parameter combinations (say $f_1 - f_2 \gamma$ small, k_2 large and g_1 small) that satisfy these inequalities.

The analysis in this section is summed up in the following Proposition.

Proposition 3 *When efforts are substitutes and there is an additional non-verifiable performance measure (q) we have: stronger career effects may well increase effective as well as explicit incentives; career incentives and explicit incentives may thus be complements.*

In particular, under risk neutrality career incentives and explicit incentives are complements when $-\rho_q \mathbf{k}'C^{-1}\mathbf{g} > \rho_y \mathbf{f}'C^{-1}\mathbf{g} + \rho_z \mathbf{g}'C^{-1}\mathbf{g}$.

4.2 Distortion and Incentives

We now consider the relationship between distortion and incentive strength in the extended model. We want to make the point that in this setting it may well be the case that *first-period incentives become stronger when the performance measure becomes more distorted*. While it is still the case that second-period (and static) incentives are maximal when the performance measure is non-distorted (i.e. when $\mathbf{g} = \mathbf{f}$), first-period incentives may well be much lower for the non-distorted compared to a distorted measure.

To see this consider the optimal first-period incentives as given by (7). Let superscript P refer to the case of a 'perfect' (non-distorted) measure ($\mathbf{g} = \mathbf{f}$), and note that the second-period optimal incentive in this case is $\alpha_2^{*P} = 1$. Comparing incentives for this non-distorted measure to some other measure ($\mathbf{g} \neq \mathbf{f}$) we have

$$\tilde{\alpha}_1^{*P} = \alpha_1^{*P} + \beta_z^P = (1 - \beta_y^P)1 - \beta_q^P \frac{\mathbf{k}'C^{-1}\mathbf{f}}{\mathbf{f}'C^{-1}\mathbf{f}}$$

$$\tilde{\alpha}_1^* = \alpha_1^* + \beta_z = (1 - \beta_y)\alpha_2^* - \beta_q \frac{\mathbf{k}'C^{-1}\mathbf{g}}{\mathbf{g}'C^{-1}\mathbf{g}}$$

and $\beta_i^P = (h - 1)\rho_i$, $\beta_i = (h - \alpha_2^*)\rho_i$. Comparing these two we see that, although $\alpha_2^* < 1$ and hence a lower net implicit incentive on y ($\beta_y < \beta_y^P$) tend to make the effective first-period incentive lower in the distorted case, the term associated with implicit incentives on the non-verifiable measure q may compensate for that. In particular, we may have $\mathbf{k}'C^{-1}\mathbf{f} > 0$ and $\mathbf{k}'C^{-1}\mathbf{g} < 0$. In such cases one can easily verify that effective and monetary incentives may be higher for the distorted measure than for the non-distorted one. We illustrate this with an example.

Example 2. Consider the example in the previous section. Mainly to simplify notation we specialize and consider some numerical values. Let

$$\mathbf{f}' = (1, 1), \quad \mathbf{g}' = (2, 0), \quad \mathbf{k}' = (0, k_2), \quad \gamma = \frac{1}{2}$$

Then we find $\frac{\mathbf{k}'C^{-1}\mathbf{f}}{\mathbf{f}'C^{-1}\mathbf{f}} = \frac{f_1 k_2}{2} = \frac{k_2}{2}$, $\frac{\mathbf{k}'C^{-1}\mathbf{g}}{\mathbf{g}'C^{-1}\mathbf{g}} = -\frac{k_2 \gamma}{g_1} = -\frac{k_2}{4}$, $\alpha_2^* = \frac{f_1(1-\gamma)}{g_1} = \frac{1}{4}$, and hence

$$\tilde{\alpha}_1^{*P} = \alpha_1^{*P} + \beta_z^P = (1 - \beta_y^P) - \beta_q^P \frac{k_2}{2}$$

$$\tilde{\alpha}_1^* = \alpha_1^* + \beta_z = (1 - \beta_y) \frac{1}{4} + \beta_q \frac{k_2}{4}$$

with $\beta_i^P = (h - 1)\rho_i$, $\beta_i = (h - \frac{1}{4})\rho_i$. There is clearly a range of values for the remaining parameters that yields higher incentives for the distorted measure ($\tilde{\alpha}_1^* > \tilde{\alpha}_1^{*P}$ and $\alpha_1^* > \alpha_1^{*P}$).

Proposition 4 *When efforts are substitutes and there is an additional non-verifiable measure (q) we have: While second-period (and static) incentives are maximal when the verifiable performance measure (z) is non-distorted (i.e. when $\mathbf{g} = \mathbf{f}$), first-period incentives may well be much lower for a non-distorted compared to a distorted measure*

4.3 Non-Distorted Performance Measure Is Not Optimal

In this section we consider the relationship between distortion and total surplus. We consider variations in the performance measure (z), and in particular variations in its degree of distortion from the true value, as measured by the angle θ between vectors \mathbf{g} and \mathbf{f} . To simplify we consider risk neutral agents. In a static case it will be optimal to have a verifiable performance measure (z) that is completely aligned with (non-distorted from) the true value (y), i.e. such that $\theta = 0$, or equivalently $\mathbf{g} = \gamma\mathbf{f}$, $\gamma > 0$. The first-best can then be achieved under risk neutrality (by setting $\alpha = \gamma^{-1}$).

For the dynamic case we want to point out that, unless the non-verifiable 'quality measure' q is completely aligned with the true value y , it will not be optimal to have

\mathbf{g} completely aligned with \mathbf{f} . Thus, *in the presence of dynamic implicit incentives it will in most cases not be optimal to have a 'perfect' verifiable performance measure.*

The intuition is fairly simple; when there are (say) career incentives on q , the agent's attention is drawn in the direction defined by vector \mathbf{k} . Ideally the agent's efforts should be aligned with \mathbf{f} . When \mathbf{k} and \mathbf{f} are not aligned, monetary incentives on \mathbf{g} should ideally draw the agent's attention towards \mathbf{f} , and this will generally not be least costly to do when \mathbf{g} is perfectly aligned with \mathbf{f} .

For notational simplicity consider independent efforts ($c(\mathbf{a}) = \frac{1}{2}\mathbf{a}'\mathbf{a}$). For given performance measures the optimal surplus can be written as⁶

$$\begin{aligned} TCE^* &= TCE_1^* + TCE_2^* \\ &= \max_{\tilde{\alpha}_1} [\mathbf{f}'\mathbf{a}_1(\tilde{\alpha}_1) - c(\mathbf{a}_1(\tilde{\alpha}_1))] + \max_{\alpha_2} [\mathbf{f}'\mathbf{a}_2(\alpha_2) - c(\mathbf{a}_2(\alpha_2))] \end{aligned}$$

where

$$\mathbf{a}_2(\alpha_2) = \alpha_2\mathbf{g}, \quad \mathbf{a}_1(\tilde{\alpha}_1) = \tilde{\alpha}_1\mathbf{g} + \beta_y\mathbf{f} + \beta_q\mathbf{k}$$

Consider a marginal variation in the component g_i ; this yields

$$\begin{aligned} \frac{\partial TCE^*}{\partial g_i} &= (f_i - a_{1i})\tilde{\alpha}_1^* + (f_i - a_{2i})\alpha_2^* \\ &= ((1 - \beta_y)f_i - \tilde{\alpha}_1^*g_i - \beta_qk_i)\tilde{\alpha}_1^* + (f_i - \alpha_2^*g_i)\alpha_2^* \end{aligned}$$

We see that for $\mathbf{g} = \gamma\mathbf{f}$ (perfect alignment) the second term in the above expression vanishes, but the first term does not, and hence such perfect alignment will not be optimal.

In fact, the formula shows that some linear combination, say $\mathbf{g} = \mathbf{f} + \lambda\mathbf{k}$, will be optimal. In practice it will hardly be possible to fine-tune performance measures to find the optimal balance, so to characterize the optimum may not be so interesting. The point we want to make is that some 'distorted' performance measure may in these cases well be better than a non-distorted one.

Proposition 5 *When there is a distorted non-verifiable performance measure (q) that generates implicit incentives, it is not optimal that the verifiable performance measure (z) is non-distorted.*

5 Conclusion

A general problem for designing incentive schemes is that available performance measures seldom capture precisely agents' true contributions to principals' objectives. In this paper we have analyzed to what extent implicit dynamic incentives such as career concerns and ratchet effects may alleviate or aggravate these problems.

First we considered the case where the principal provides incentives on a verifiable, but distorted, performance measure, and in addition some information about

⁶For notational simplicity we drop the term $h'\eta$.

the agent's performance is provided to the principal (and the market) through a non-distorted but non-verifiable measure of true value. In this case implicit incentives are related both to the distorted and the undistorted performance measures (and hence the degree of misalignments between them).

The analysis demonstrated that implicit dynamic incentives have important real effects in such settings, and that these effects are in several respects different from the corresponding effects in settings where a non-distorted performance measure is available. In particular, we found that both career and ratchet elements have real effects; neither can costlessly be neutralized by monetary incentives, and that stronger ratchet effects may increase optimal monetary incentives. The findings that career elements may have real effects and that stronger ratchet effects may increase monetary incentives are quite opposite to what one finds in settings where contracting on a non-distorted measure is feasible (e.g. Gibbons 1987; Meyer and Vickers 1997).

The second model we present captures the fact that in addition to a verifiable (and distorted) performance measure, there are often other non-verifiable measures that may yield valuable information about the agent's performance. A typical case is one where quantity aspects are verifiable but quality aspects are not, yet some information about these quality aspects is observable for the relevant parties. In this setting we show that some new issues arise. Most notably, we show that career and monetary incentives may be complements rather than substitutes, and that explicit incentives may increase with more distortion in the (verifiable) performance measure. The latter effect occurs when more distortion induces the principal to increase monetary incentives in order to maintain an appropriate balance of the agent's effort among tasks. Hence in this case, and opposed to what is suggested in Kerr (1975), it may be appropriate to "reward for A, while hoping for B".

References

- Baker, G. (1992). Incentives contracts and performance measurement. *Journal of Political Economy* 100(3), 598–614.
- Baker, G. (2002). Distortion and risk in optimal incentive contracts. *Journal of Human Resources* 37(3), 728–751.
- DeGroot, M. H. (1970). *Optimal Statistical Decisions*. New York: McGraw-Hill.
- Dewatripont, M., I. Jewitt, and J. Tirole (1999). The economics of career concerns, Part II: Application to missions and accountability of government agencies. *Review of Economic Studies* 66, 199–217.
- Fama, E. F. (1980). Agency problems and the theory of the firm. *Journal of Political Economy* 88, 288–307.
- Feltham, G. and J. Xie (1994). Performance measure congruity and diversity in multitask/agent relations. *The Accounting Review* 69.
- Gibbons, R. (1987). Piece-rate incentives schemes. *Journal of Labor Economics* 5, 413–29.
- Gibbons, R. and K. J. Murphy (1992). Optimal incentives in the presence of career concerns: Theory and evidence. *Journal of Political Economy* 100, 468–505.
- Holmstrom, B. (1982). Managerial incentive schemes—a dynamic perspective. In *Essays in economics and management in honour of Lars Wahlbeck*. Svenska Handelshogskolan, Helsinki.
- Holmstrom, B. and P. Milgrom (1987). Aggregation and linearity in the provision of intertemporal incentives. *Econometrica* 55, 303–328.
- Holmstrom, B. and P. Milgrom (1991). Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics and Organization* 7, 24–52.
- Kerr, S. (1975). On the folly of rewarding for A while hoping for B. *Academy of Management Journal* 18(4), 769–783.
- Meyer, M. A., T. E. Olsen, and G. Torsvik (1996). Limited intertemporal commitment and job design. *Journal of Economic Behavior and Organization* 31, 401–417.
- Meyer, M. A. and J. Vickers (1997). Performance comparisons and dynamic incentives. *Journal of Political Economy* 105, 547–581.
- Schnedler, W. (2003). On the prudence of rewarding A, while hoping for B: Performance measures and effort allocation between tasks. Discussion Paper Bgse-7, Bonn University.
- Weitzman, M. (1976). The ratchet principle and performance incentives. *Bell Journal Of Economics* 11, 302–308.

Appendices

A Technicalities

In this appendix we provide more details regarding some of the calculations in this paper.

A.1 Regression Coefficients

We first consider the case outlined in section 2 and 3. In this case the information signals are

$$\begin{aligned} y_t &= h' \eta + \mathbf{f} \mathbf{a}_t + \varepsilon_t \\ z_t &= \eta + \mathbf{g} \mathbf{a}_t + \zeta_t. \end{aligned}$$

We seek $\mathbf{E}(z_2 | z_1, y_1)$ and $\mathbf{E}(\eta | z_1, y_1)$. The covariance matrixes (z_2, z_1, y_1) and (η, z_1, y_1) are

$$\begin{aligned} &\begin{bmatrix} \sigma_\eta^2 & \sigma_\eta^2 & h' \sigma_\eta^2 \\ \sigma_\eta^2 & \sigma_\eta^2 + \sigma_z^2 & h' \sigma_\eta^2 \\ h' \sigma_\eta^2 & h' \sigma_\eta^2 & (h')^2 \sigma_\eta^2 + \sigma_y^2 \end{bmatrix}, \\ &\begin{bmatrix} \sigma_\eta^2 + \sigma_z^2 & \sigma_\eta^2 & h' \sigma_\eta^2 \\ \sigma_\eta^2 & \sigma_\eta^2 + \sigma_z^2 & h' \sigma_\eta^2 \\ h' \sigma_\eta^2 & h' \sigma_\eta^2 & (h')^2 \sigma_\eta^2 + \sigma_y^2 \end{bmatrix}. \end{aligned}$$

By inverting and applying well-known formulas (e.g. DeGroot 1970) we obtain:

$$\begin{aligned} \sigma_{2c}^2 &= \text{Var}(z_2 | z_1, y_1) = \frac{\sigma_\eta^2 \sigma_y^2 \sigma_z^2}{\sigma_\eta^2 \sigma_y^2 + \sigma_z^2 h'^2 \sigma_\eta^2 + \sigma_z^2 \sigma_y^2} + \sigma_z^2 \\ \rho_z &= \frac{\partial}{\partial z} \mathbf{E}(\eta | \mathbf{x}_1) = \left[\frac{\sigma_\eta^2}{\sigma_\eta^2 \sigma_y^2 + \sigma_z^2 h'^2 \sigma_\eta^2 + \sigma_z^2 \sigma_y^2} \right] \sigma_y^2 \\ \rho_y &= \frac{\partial}{\partial y} \mathbf{E}(\eta | \mathbf{x}_1) = \left[\frac{h' \sigma_\eta^2}{\sigma_\eta^2 \sigma_y^2 + \sigma_z^2 h'^2 \sigma_\eta^2 + \sigma_z^2 \sigma_y^2} \right] \sigma_z^2. \end{aligned}$$

Consider the case outlined in Section 4. In this case the information signals are

$$\begin{aligned} y_t &= h' \eta + \mathbf{f} \mathbf{a}_t + \varepsilon_t \\ z_t &= \eta + \mathbf{g} \mathbf{a}_t + \zeta_t \\ q_t &= \eta + \mathbf{k} \mathbf{a}_t + \kappa_t. \end{aligned}$$

The covariance matrix for (η, y_1, z_1, q_1) is now

$$\begin{bmatrix} \sigma_\eta^2 & h'\sigma_\eta^2 & \sigma_\eta^2 & \sigma_\eta^2 \\ h'\sigma_\eta^2 & h'^2\sigma_\eta^2 + \sigma_y^2 & h'\sigma_\eta^2 & h'\sigma_\eta^2 \\ \sigma_\eta^2 & h'\sigma_\eta^2 & \sigma_\eta^2 + \sigma_z^2 & \sigma_\eta^2 \\ \sigma_\eta^2 & h'\sigma_\eta^2 & \sigma_\eta^2 & \sigma_\eta^2 + \sigma_q^2 \end{bmatrix}$$

By inverting and applying well-known formulas we get

$$\begin{aligned} \rho_y &= \frac{\partial}{\partial y} \mathbf{E}(\eta \mid y_1, z_1, q_1) = \frac{h'\sigma_\eta^2\sigma_z^2\sigma_q^2}{h'^2\sigma_\eta^2\sigma_z^2\sigma_q^2 + \sigma_y^2\sigma_\eta^2\sigma_q^2 + \sigma_y^2\sigma_z^2\sigma_\eta^2 + \sigma_y^2\sigma_z^2\sigma_q^2} \\ \rho_z &= \frac{\partial}{\partial z} \mathbf{E}(\eta \mid y_1, z_1, q_1) = \frac{\sigma_\eta^2\sigma_y^2\sigma_q^2}{h'^2\sigma_\eta^2\sigma_z^2\sigma_q^2 + \sigma_y^2\sigma_\eta^2\sigma_q^2 + \sigma_y^2\sigma_z^2\sigma_\eta^2 + \sigma_y^2\sigma_z^2\sigma_q^2} \\ \rho_q &= \frac{\partial}{\partial q} \mathbf{E}(\eta \mid y_1, z_1, q_1) = \frac{\sigma_\eta^2\sigma_y^2\sigma_z^2}{h'^2\sigma_\eta^2\sigma_z^2\sigma_q^2 + \sigma_y^2\sigma_\eta^2\sigma_q^2 + \sigma_y^2\sigma_z^2\sigma_\eta^2 + \sigma_y^2\sigma_z^2\sigma_q^2}. \end{aligned}$$

A.2 Proof of Proposition 2

Recall that first- and section period incentives are given by

$$\tilde{\alpha}_1^* = \alpha_1^* + \beta_z = (1 - \beta_y) \frac{\mathbf{f}'C^{-1}\mathbf{g}}{\mathbf{g}'C^{-1}\mathbf{g}} - \beta_q \frac{\mathbf{k}'C^{-1}\mathbf{g}}{\mathbf{g}'C^{-1}\mathbf{g}}, \quad (5)$$

$$\alpha_2^* = \frac{\mathbf{f}\mathbf{g}}{|\mathbf{g}|^2 + r\sigma_{2c}^2}. \quad (3)$$

From (3) we have $\mathbf{f}\mathbf{g} = \alpha_2^*K'$, where $K' = |g|^2 + r\sigma_{2c}^2$. Better alignment will thus increase α_2^* . Defining $K = |g|^2 + r\sigma_{1z}^2$, we have moreover from (5)

$$\begin{aligned} \tilde{\alpha}_1^*K &= (1 - \beta_y)\alpha_2^*K' - r\alpha_2^*\rho_z\sigma_{1z}^2 - rh\rho_y\sigma_{1yz} \\ &= [(1 - (h - \alpha_2^*)\rho_y)K' - r\rho_z\sigma_{1z}^2] \alpha_2^* - rh\rho_y\sigma_{1yz} \end{aligned}$$

where in particular the square bracket must be positive. Differentiation yields

$$\frac{\partial \tilde{\alpha}_1^*}{\partial \alpha_2^*} K = [(1 - (h - \alpha_2^*)\rho_y)K' - r\rho_z\sigma_{1z}^2] + \alpha_2^*\rho_y K'$$

This is positive for all parameter values that yield non-negative $\tilde{\alpha}_1^*$. This proves the first part of the proposition.

Consider next the equilibrium total surplus $TCE_1^* + TCE_2^*$. From the envelope property, and taking into account that the equilibrium first-period effort is $\mathbf{a}_1^* = \tilde{\alpha}_1^*g + \beta_y f$, we obtain from (4):

$$\begin{aligned} \frac{\partial TCE_1^*}{\partial g_i} &= f_i \tilde{\alpha}_1^* - a_{1i}^* \tilde{\alpha}_1^* - r((\tilde{\alpha}_1^* + \alpha_2^*\rho_z)\rho_z\sigma_{1z}^2 + \rho_z h\rho_y\sigma_{1yz}) \frac{\partial \alpha_2^*}{\partial g_i} \\ &= ((1 - \beta_y)f_i - \tilde{\alpha}_1^*g_i)\tilde{\alpha}_1^* - r((\tilde{\alpha}_1^* + \alpha_2^*\rho_z)\sigma_{1z}^2 + h\rho_y\sigma_{1yz})\rho_z \frac{\partial \alpha_2^*}{\partial g_i} \end{aligned}$$

where

$$\frac{\partial \alpha_2^*}{\partial g_i} = \frac{\partial}{\partial g_i} \frac{\Sigma f_i g_i}{\Sigma g_i^2 + r \sigma_{2c}^2} = \frac{f_i}{(\Sigma g_i^2 + r \sigma_{2c}^2)} - \frac{2 \alpha_2^* g_i}{(\Sigma g_i^2 + r \sigma_{2c}^2)}$$

In a similar way we obtain from (2):

$$\frac{\partial TCE_2^*}{\partial g_i} = f_i \alpha_2^* - a_{2i}^* \alpha_2^* = (f_i - g_i \alpha_2^*) \alpha_2^*$$

All in all we thus have, for the equilibrium total surplus

$$\frac{\partial TCE^*}{\partial g_i} = A f_i - B g_i$$

where A, B are independent of i . The formula shows that the surplus is maximal when the vector g is perfectly aligned with the vector \mathbf{f} . QED.