Performance Comparisons and Dynamic Incentives
Author(s): Margaret A. Meyer and John Vickers
Source: *The Journal of Political Economy*, Vol. 105, No. 3 (Jun., 1997), pp. 547-581
Published by: The University of Chicago Press
Stable URL: http://www.jstor.org/stable/2138851
Accessed: 02/04/2010 19:12

Performance Comparisons and Dynamic Incentives

Margaret A. Meyer

Nuffield College, Oxford University

John Vickers

Institute of Economics and Statistics, Oxford University

It is well known that comparative performance information (CPI) can enhance efficiency in static principal-agent relationships by improving the trade-off between insurance and incentives in the design of explicit contracts. In dynamic settings, however, there may be implicit as well as explicit incentives, for example, managerial career concerns and the ratchet effect in regulation. We show that the dynamic effects of CPI on implicit incentives can either reinforce or oppose the familiar (static) insurance effect and in either case can be more important for efficiency. The overall welfare effects of CPI are thus ambiguous and can be characterized in terms of the underlying information structure.

I. Introduction

Incentive theory has shown that comparative performance information can improve incentives and efficiency in principal-agent rela-

We are grateful to Bob Gibbons, Shinsuke Kambe, Paul Klemperer, Paul Milgrom, Steve Nickell, Trond Olsen, Patrick Rey, and Jeff Zwiebel for helpful discussions; to seminar participants, Sherwin Rosen, and the referee for valuable comments; to Susan Athey and Kate Dewhurst for excellent research assistance; and to the Economic and Social Research Council for support under the Contracts and Competition Initiative (grant L114251038). Meyer thanks the Stanford Graduate School of Business and the Yale School of Management, where parts of this research were carried out.

[*Journal of Political Economy*, 1997, vol. 105, no. 3]

tionships.[1] The ability to compare the performance of one agent with that of others increases the precision with which efforts can be estimated and thereby improves the terms of the trade-off between insurance and efficiency in designing incentives for a risk-averse agent. Insofar as competition yields comparative performance information, this *insurance effect* is a way in which competition can promote productive efficiency. In sum, competition can "provide a richer information base on which to write contracts" (Holmstrom and Tirole 1989, p. 96).

But not all incentives are created by design. Managers in firms, for example, may be motivated not only by explicit contractual links between pay and performance, but also because good performance will enhance the managerial labor market's perception of their productivity and hence improve their future earnings. This *reputation effect* is a type of implicit incentive, as distinct from explicit incentives in the form of pay contractually determined by current performance. Another well-known type of implicit incentive is the *ratchet effect.* The incentives of, say, a regulated firm to cut costs will be weakened if the firm anticipates that the regulator will respond by bringing down price. In this setting, the implicit incentives discourage effort.

In dynamic incentive problems, unless the designer of incentives has full powers of precommitment, in which case a dynamic incentive problem becomes essentially static, implicit incentives such as the ratchet effect are potentially important even when explicit incentives can be provided. The reason is that current performance affects not only the current reward but also the terms of the future explicit incentive contract. In such settings, overall *effective incentives* are the sum of explicit and implicit incentives.[2]

The question arises, then, of how comparative performance information influences effective incentives and welfare in dynamic principal-agent relationships. The aim of this paper is to answer this question in a very simple framework and, in particular, to explore how comparative performance information influences the ratchet effect (which turns out to be more fundamental than the reputation effect, as will be explained below). Our general conclusion is that the dy-

[1] Holmstrom (1982*b*), Mookherjee (1984), and Meyer and Mookherjee (1987) analyze general forms of comparative performance evaluation; Lazear and Rosen (1981), Green and Stokey (1983), Nalebuff and Stiglitz (1983), and Rosen (1986) focus exclusively on tournaments.

[2] Implicit incentives of the type we analyze arise when a principal has some ex post discretion as to how to respond to an agent's performance *and* learns from current performance about future performance. A broader notion of implicit incentives would include all incentives enforced through repeated interactions (as in, e.g., Baker, Gibbons, and Murphy [1994]).

namic effect of comparative performance information can either reinforce or oppose the familiar (static) insurance effect and in either case can be more important for efficiency. Whereas comparative performance information is always beneficial in static settings and when full precommitment is possible, we show here that it has ambiguous welfare consequences, which depend on the underlying information structure, in dynamic settings when precommitment powers are limited. Most agency relationships in practice fall into this latter category.

Since the availability of comparative performance information is just one example of an improvement in information, our analysis illustrates a still more general lesson about the value of better information in dynamic agency relationships. When incentives are provided wholly or partly by implicit mechanisms, better information does not necessarily increase welfare because better information can weaken implicit incentives.[3]

The paper is organized as follows. In Section II, we present two examples, which abstract from risk aversion and the design of explicit incentives, each of which shows that the incentive effects of comparative performance information are ambiguous when implicit incentives are present. We provide a simple characterization of these effects in terms of the correlation between agents' time-invariant features (e.g., their intrinsic productivities) relative to that between time-independent shocks (e.g., performance measurement errors). Depending on the relative sizes of these correlations, the introduction of comparative performance information will either raise or lower the weight attached to an agent's current performance in determining his future payoff. The first example illustrates how comparative performance information influences the reputation effect in a model of managerial career concerns. The second shows how it influences the ratchet effect in a model of yardstick regulation. Comparative performance information has very different—indeed opposite—effects in the two examples. This contrast can nevertheless be explained using a more general framework that allows for differences in agents' bargaining power and for differences in their rate of direct current gain from improvements in performance.

In Section III, we analyze a still more general model that allows for the design of explicit incentives within periods but limited precommitment between periods, risk aversion on the part of agents,

[3] Using different types of incentive models, Besley and Case (1995), Cremer (1995), and Zwiebel (1995) also examine the impact of improvements in information when precommitment powers are limited and find, as we do, that welfare may decline.

and arbitrary degrees of bargaining power for agents. First, we show that, in contrast to Section II, the efficiency consequences of comparative performance information are independent of agents' bargaining power and hence independent of the strength of the reputation effect. The reason for this result, put loosely, is that it is costless, and therefore optimal, to design explicit incentives exactly to offset differences in bargaining power. By contrast, the efficiency consequences of comparative performance information do depend on its impact on the ratchet effect: whereas explicit incentives can be designed to offset any intensification of the ratchet effect, such an adjustment lowers welfare because it imposes more risk on the agents. We then analyze in detail the impact of comparative performance information on effort incentives, risk, and overall efficiency. We decompose its overall welfare impact into a static effect (the insurance gain) and a dynamic effect (the ratchet gain/loss) and examine how these effects depend on the information structure. We show that the dynamic effect can be larger in magnitude than the static effect, whether it reinforces or opposes it.

The analysis in Section III can apply to a principal who must decide, prior to contracting with an agent, whether or not to acquire comparative performance information. For example, a firm may be deciding whether to institute a policy of benchmarking, which would involve choosing another company with a related technology and comparing its own employees' performance with that at the benchmark company. Or division managers may be deciding whether to adopt a policy of sharing information about their subordinates' performances. In other settings, a principal may not be able to prevent herself from observing and using comparative performance information, but she may have some control over its information content. Consider, for example, a manager with several subordinates. Through choices about how to monitor, as well as about hiring and training policies (whether to hire individuals with similar or different education and experience, and whether they should receive similar or different training), she can affect the correlations between performance measurement errors and between employees' time-invariant features. The analysis in Section III shows how welfare depends on these correlations.

Section IV discusses applications and extensions of the analysis, and Section V presents a conclusion.

II. Comparative Performance Information and Implicit Incentives

In this section, we present two contrasting examples of how comparative performance information (CPI) affects implicit incentives under moral hazard. The first is a model of managerial career concerns based on the analysis of Holmstrom (1982*a*), but with the addition of relative performance evaluation. The second is a model of yardstick regulation and the ratchet effect.

A. *Relative Performance Evaluation and Managerial Career Concerns*

Fama (1980) and Holmstrom (1982*a*) have argued that, even in the absence of explicit incentive contracts, managers will have some motivation to work hard because doing so improves their reputation in the labor market and hence their future earnings. We begin by describing a simple version of Holmstrom's one-agent model. There are two time periods, and there is no discounting. A risk-neutral manager, M, has utility function

$$U = w_1 - C(e_1) + w_2 - C(e_2),$$

where w_t is M's wage in period t, e_t is his effort level, and $C(\cdot)$ is a strictly convex cost-of-effort function with $C'(0) = 0$. The output of M in period t, at any firm, is $x_t = e_t + a + u_t$, where a is a time-invariant characteristic of M (his ability, say) and u_t is a transient shock (measurement error, say). The information structure is as follows: e_t is privately chosen by M in period t; x_t is publicly observed at the end of that period; a, u_1, and u_2, which are initially unknown to all, are normally distributed with mean zero and uncorrelated with one another; and u_1 and u_2 have equal variances. A useful measure of the relative variances of a and u_t is

$$\tau = \frac{\mathrm{var}(a)}{\mathrm{var}(a) + \mathrm{var}(u_t)},$$

which is larger, the larger the signal-to-noise ratio in observations of x_t about a. Explicit incentive contracts linking w_t to x_t are assumed to be impossible. Thus M has no incentive to exert effort in period 2, and so $e_2 = 0$.

In period 1, however, M has an implicit incentive to exert effort, insofar as doing so increases M's reputation for ability and hence w_2. With perfect competition between risk-neutral employers to hire M in period 2, it will be the case that $w_2 = E(x_2 | x_1)$, which equals $E(a | x_1)$ given that $e_2 = 0$. Under our assumptions, it is a standard result that $E(a | x_1)$ is a weighted average of (i) the prior mean of ability (here taken to be zero) and (ii) period 1 output (adjusted to account for the contribution of M's period 1 effort), the weight on the latter term being τ. Thus $w_2 = \tau(x_1 - \hat{e}_1)$, where $\hat{e}_1$ is the labor market's conjecture about period 1 effort. The first-order condition for M's choice of e_1 is simply $C'(e_1) = \tau \leq 1$. Thus the larger τ is, the stronger the "reputation effect" on M's incentives. Note that this first-order condition is independent of $\hat{e}_1$. With rational expectations in the managerial labor market, $\hat{e}_1$ will satisfy this equation.

Since the social marginal benefit of effort is one, the efficient effort level satisfies $C'(e_1^*) = 1$, so the equilibrium effort level is always lower than optimal. The larger the variance of prior beliefs about ability or the smaller the variance of measurement error, the easier it is for M, by increasing his effort, to influence the market's estimate of his ability, and hence the stronger the implicit effort incentives.

How does CPI affect incentives in this setting? Suppose now that there is a second manager with symmetric preferences and technology, and that firms can base their wage offers to each manager on both managers' previous outputs. Suppose also that managers choose their efforts noncooperatively. Using i and j subscripts to differentiate between the managers, let $\eta = \text{corr}(a_i, a_j)$ denote the correlation between their time-invariant characteristics and let $\rho = \text{corr}(u_{ti}, u_{tj})$ denote the correlation between the transient shocks to their measured performance in each period. There is zero correlation between each pair of a and u terms and zero temporal correlation between the u's. Henceforth, "CPI" refers to the case in which at least one of η and ρ is nonzero. (If $\eta = \rho = 0$, observations on manager j are uninformative about manager i.)

It remains true that $e_{2i} = 0$ and that w_{2i} equals the conditional expectation of a_i, but that expectation is now conditional on x_{1j} as well as x_{1i}. Given our assumptions, the variables a_i, x_{1i}, and x_{1j} have a multivariate normal distribution with covariance matrix proportional to

$$\begin{pmatrix} \tau & \tau & \eta\tau \\ \tau & 1 & \kappa \\ \eta\tau & \kappa & 1 \end{pmatrix},$$

where $\kappa = (1 - \tau)\rho + \tau\eta$ is the "cross-section" correlation between x_{1i} and x_{1j}. Bayesian updating (which is analogous, in this setting, to standard ordinary least squares methods; see, e.g., Theil [1971, chap. 4]) yields

$$w_{2i} = E(a_i | x_{1i}, x_{1j}) = \frac{\tau}{1 - \kappa^2}[(1 - \eta\kappa)(x_{1i} - \hat{e}_{1i}) + (\eta - \kappa)(x_{1j} - \hat{e}_{1j})].$$

By the same reasoning as before, manager i's first-period effort level when CPI is available solves

$$C'(e_{1i}) = \tau\left(\frac{1 - \eta\kappa}{1 - \kappa^2}\right) \equiv \psi \leq 1.$$

Comparing this with the single-manager case, we have the following proposition.

PROPOSITION 1. In the managerial career concerns model, effort incentives and efficiency are greater with performance comparisons than without if and only if $\kappa(\rho - \eta) > 0$.

If the managers' performances are positively correlated (i.e., if $\kappa > 0$), the condition is simply whether or not $\rho > \eta$. To gain some intuition for proposition 1, consider first the special case $\eta = 0$ and $\rho \neq 0$. Comparative performance information improves effort incentives because the observation of x_{1j} effectively reduces the variance of the "noise" u_{1i} and so increases the weight on x_{1i} in estimating a_i. This is reminiscent of the insurance effect of CPI in a static principal-agent model with explicit incentives (see Sec. IIIA below): there, observation of j's output increases the precision with which i's effort is estimated, leading the principal to impose stronger incentives. On the other hand, if $\eta \neq 0$ and $\rho = 0$, the observation of x_{1j} effectively reduces the "prior" variance of ability a_i and so reduces the weight on x_{1i} in estimating a_i. (This effect has no counterpart in a static model with explicit incentives.)

Manager i's period 2 wage w_{2i} is decreasing (increasing) in manager j's period 1 output x_{1j}, according as $\rho > (<) \eta$. If $\rho > \eta$, there is a negative externality between managers and some rivalry between them. Again, this is qualitatively similar to the optimal contract with CPI in a static model. If $\eta > \rho$, there is a positive externality, and each manager free-rides to some degree on the efforts of the other to enhance reputation. This free-riding helps explain why CPI worsens incentives if $\kappa(\eta - \rho) > 0$ and has no counterpart in a static setting with explicit incentives.

Thus, when incentives are provided by career concerns, the use of CPI in determining wages is sensitive to precisely how managers' performances are related and can differ from its use when all incentives are explicit. These findings may have implications for the interpretation of empirical work on relative performance evaluation (Antle and Smith 1986; Gibbons and Murphy 1990).

When incentives are provided entirely by explicit contracts, efficiency can never be reduced by CPI because the principal could always commit to ignoring the extra information. In contrast, when incentives are provided by career concerns, efficiency depends not on how much information is available to firms but on how much weight they place on an agent's output in estimating his ability.

B. *Yardstick Regulation and the Ratchet Effect*

A major problem in the economics of monopoly regulation is that a regulated firm's incentive to reduce costs is blunted if the price that it is allowed to charge is reduced in line with its cost level. Yardstick regulation—the linking of one monopolist's permitted price level to the cost levels of firms in similar conditions (e.g., monopolists in other regions)—seeks to overcome this problem (see Shleifer 1985). In static models in which the regulator can explicitly condition price on cost levels, yardstick regulation is beneficial because, by improving the regulator's information, it lessens the cost of providing incentives for firms (in a manner analogous to the insurance effect described earlier). In practice, however, CPI is sometimes used not in explicit price controls but only at times of regulatory review when price levels are set for the next period ahead (see Armstrong, Cowan, and Vickers 1994, chap. 6). In such cases, yardstick regulation is implicit rather than explicit, as it were. It is not our purpose to explain why yardstick regulation sometimes takes this form, but rather to analyze its consequences in a simple model. Of particular interest is the way in which CPI influences the "ratchet effect": the ratchet effect describes the dampening of a firm's incentives to reduce current costs because of its anticipation that future price reductions will result.

Consider first a single risk-neutral regulated firm that faces inelastic demand for one unit. Its production cost level in period $t = 1, 2$ is

$$c_t = k - x_t = k - (a + e_t + u_t),$$

where k is a known constant, and x_t and its components are as in the model in subsection A. We shall call a the firm's "intrinsic efficiency" rather than "ability" in this context.[4] The cost-of-effort function $C(e_t)$ is also as above. The firm's objective is to maximize the undiscounted sum of revenues minus total costs (i.e., production costs plus effort costs).

We shall assume that the regulator's policy is to set price (and hence revenue) p_t in period t equal to the firm's expected total costs

[4] Since the regulator and firm begin with identical beliefs about a, our analysis of the ratchet effect involves moral hazard but not adverse selection. In contrast, Freixas, Guesnerie, and Tirole (1985) and Laffont and Tirole (1993) and the references therein study the ratchet effect in models with adverse selection.

in that period, conditional on past information.[5] In particular,

$$p_2 = k - E(a|x_1) - \hat{e}_2 + C(\hat{e}_2),$$

where $\hat{e}_2$ is the regulator's conjecture about the firm's period 2 cost-reduction effort. Since nothing further happens after period 2, the firm will choose e_2 to maximize $x_2 - C(e_2)$, so effort in period 2 will be efficient. In period 1, however, the first-order condition for cost-reduction effort is

$$C'(e_1) = 1 - \tau \leq 1.$$

Effort in the first period has a marginal benefit of one to the firm in that period because price p_1 is fixed and unit production cost c_1 falls, but a marginal cost of τ in period 2 because of the unfavorable impact on p_2. This latter effect is precisely the ratchet effect. In contrast to the managerial career concerns model, low τ is good for incentives and efficiency in this regulation model. Thus, for example, more precise prior information about intrinsic efficiency is good because less weight is then given to x_1 in setting p_2.

Suppose now that the regulator can observe the performance of a second firm in symmetric circumstances and that firms choose efforts noncooperatively. As before, let η be the correlation between intrinsic efficiency levels and let ρ be the correlation between transient shocks. Using the same reasoning as before, we get

$$C'(e_{1i}) = 1 - \tau\left(\frac{1 - \eta\kappa}{1 - \kappa^2}\right) = 1 - \psi \leq 1$$

when yardstick regulation is used. As to the effect of CPI, we have the following proposition.

PROPOSITION 2. In the yardstick regulation model, effort incentives and efficiency are greater with performance comparisons than without if and only if $\kappa(\rho - \eta) < 0$.

This is precisely the opposite of the condition in proposition 1. In the yardstick regulation model, CPI is harmful if it leads to an increase in the weight on first-period performance in forming expec-

[5] If the firm had the option of refusing to produce in period 2, then given the pricing policy described, it would actually prefer, in period 1, to follow a "take-the-money-and-run" strategy: working hard (at the socially efficient level) in the first period, in the knowledge that it would shut down in the second. However, in practice, the period 2 price set by the regulator is likely to cover not just period 2 costs but also sunk costs incurred earlier. As long as this "deferred compensation" is large enough, it will ensure that the firm prefers to produce in both periods, choosing efforts as described below, rather than adopt the take-the-money-and-run strategy.

tations in period 2, whereas in the managerial career concerns model, such a change makes CPI desirable.

C. *A Simple Framework Encompassing the Reputation and Ratchet Effects*

The contrast between propositions 1 and 2 can be traced to two differences that are of more general importance. The first is that the manager in subsection *A* and the regulated firm in subsection *B* are in very different positions in terms of their period 2 bargaining power. Because of competition between employers for his services, the manager gains (or loses) *all* of the upward (or downward) revision of his expected ability level—in short, of his reputation—that occurs at the end of period 1. In the regulation model, on the other hand, all of this expected gain (or loss) is borne by the regulator (or rather consumers) and *none* by the firm since the regulator sets price so that the firm's expected profit is zero. The larger the agent's period 2 bargaining power, the stronger the reputation effect on implicit incentives.

The second key difference is that the regulated firm makes a direct current profit gain one-for-one as unit cost decreases. The manager, on the other hand, makes zero direct current wage gain as output increases.

We can summarize and generalize these differences as follows. Let b denote agent i's bargaining power—the share the agent receives of changes in his reputation $E(a_i|x_{1i}, x_{1j})$—and let d_t denote agent i's rate of direct current gain from changes in performance x_{ti}. Then agent i's period 2 compensation depends on the variables x_{1i}, x_{1j}, and x_{2i} through the terms $bE(a_i|x_{1i}, x_{1j}) + d_2[x_{2i} - E(x_{2i}|x_{1i}, x_{1j})]$. Thus the overall marginal benefit to agent i of first-period effort—the *effective* first-period incentive—is $d_1 + (b - d_2)\psi$. The first term is the direct effect in period 1, and the second term, $(b - d_2)\psi$, is the implicit incentive representing the effect of period 1 effort on period 2 payoffs. The reputation effect, which raises incentives, is measured by $b\psi$, and the ratchet effect, which dampens incentives, is measured by $d_2\psi$. For the manager, $b = 1$ and $d_t = 0$: the reputation effect is present whereas the ratchet effect is absent, and the marginal benefit of period 1 effort is ψ. For the regulated firm, $b = 0$ and $d_t = 1$. So only the ratchet effect is present, and the marginal benefit of first-period effort is $1 - \psi$.

The introduction of CPI alters the weight, ψ, that the principal places on the agent's period 1 performance in estimating the agent's intrinsic productivity, so CPI causes the reputation and ratchet effects either to increase or to decrease together. However, since these effects push incentives in opposite directions, CPI has opposite impacts on incentives in the two models described above. More generally, this framework suggests that whether CPI strengthens or weakens implicit incentives depends not only on whether it raises or lowers ψ, but also on the relative sizes of the reputation and ratchet effects.

While this framework integrates the models of the previous two subsections, it nevertheless has serious limitations. It has no role for the design of explicit contracts. Also, by assuming risk neutrality, it assumes away the insurance value of CPI. More broadly, it ignores the value of CPI for improving ex post decisions. The analysis that follows addresses these issues and shows that they have a significant impact on the efficiency consequences of CPI.

III. A Model of Implicit and Explicit Incentives

Agent A_i works for the firm of principal P, who can also observe the performance of another agent, A_j. Agents A_i and A_j behave noncooperatively. The setting of optimal incentives for A_i, which is the focus of our analysis, is unaffected by whether A_j is also employed by P. The technology and information structure are very similar to those of the model of Section II*A*.

The output of A_k, $k = i, j$, in period $t = 1, 2$, at any firm, is

$$x_{tk} = e_{tk} + a_k + u_{tk}, \tag{1}$$

where e_{tk} is A_k's privately chosen effort level in period t, a_k is a time-invariant characteristic (of A_k or of the job he performs in both periods), and u_{tk} is a transient shock to measured output. In each period, the cost of effort e to each agent is $C(e) = \frac{1}{2}e^2$. The random terms a_k and u_{tk} are initially unknown to all and have the following distributions:

$$a_k \sim N(0, \tau\sigma^2), \quad u_{tk} \sim N(0, (1-\tau)\sigma^2), \quad k = i, j,$$
$$\operatorname{corr}(a_i, a_j) = \eta, \quad \operatorname{corr}(u_{ti}, u_{tj}) = \rho. \tag{2}$$

There is zero correlation between a_i and u_{tk} and between a_j and u_{tk} $(k = i, j)$, and there is zero temporal correlation between the u_{tk} terms. The expressions below for variances follow from these assumptions:[6]

$$\operatorname{var}(x_{tk}) = \sigma^2, \quad k = i, j, \tag{3a}$$

$$\operatorname{var}(x_{ti} \mid x_{tj}) = (1 - \kappa^2)\sigma^2 \equiv v_1\sigma^2, \tag{3b}$$

[6] The Appendix contains the derivations of equations (3d), (4), (20), and (21), as well as proofs of the main propositions in the paper. The fact that the conditional variances in (3b)–(3d) are independent of the realizations of the conditioning variables is a feature of the multivariate normal distribution.

where

$$\kappa = \eta\tau + \rho(1 - \tau),$$

$$\text{var}(x_{2i} | x_{1i}) = (1 - \tau^2)\sigma^2, \quad (3c)$$

$$\text{var}(x_{2i} | x_{1i}, x_{1j}, x_{2j}) = (1 + \gamma)(1 - \tau)(1 - \rho^2)\sigma^2 \equiv v_2\sigma^2, \quad (3d)$$

where $\gamma = \text{cov}(x_{2i}, x_{1i} | x_{1j}, x_{2j})/\text{var}(x_{1i} | x_{1j}, x_{2j})$ is the coefficient on x_{1i} in the expectation $E(x_{2i} | x_{1i}, x_{1j}, x_{2j})$. The expression for this coefficient, which will have central importance in the analysis, is

$$\gamma = \tau\left[1 + 2(1 - \tau)\frac{N}{D}\right] \in [0, 1], \quad (4)$$

where

$$N = (\rho - \eta)(\rho + \eta\tau),$$

$$D = (1 - \rho^2)(1 + \tau) + 2\tau N \geq 0.$$

All firms are risk-neutral and, in their dealings with A_i, maximize

$$E(x_{1i}) - E(w_{1i}) + E(x_{2i}) - E(w_{2i}) = e_{1i} - E(w_{1i}) + e_{2i} - E(w_{2i}), \quad (5)$$

where w_{ti} is the wage paid to A_i in period t. Agent A_i is risk-averse and has the constant absolute risk aversion utility function

$$U_i = -\exp\{-r[w_{1i} - \tfrac{1}{2}(e_{1i})^2 + w_{2i} - \tfrac{1}{2}(e_{2i})^2]\}. \quad (6)$$

All firms can observe x_{ti} and x_{tj} at the end of period t.

We make two assumptions about contracting possibilities. First, only one-period contracts are enforceable.[7] Second, one-period contracts take the linear form[8]

$$w_{ti} = \alpha_t + \beta_t x_{ti} + \epsilon_t x_{tj}. \quad (7)$$

Given the normality assumptions in (2) and the linearity of contracts in (7), it follows from (6) that A_i's expected utility has the certainty equivalent

$$E(w_{1i}) - \tfrac{1}{2}(e_{1i})^2 + E(w_{2i}) - \tfrac{1}{2}(e_{2i})^2 - \tfrac{1}{2}r\,\text{var}(w_{1i} + w_{2i}) \equiv ACE. \quad (8)$$

The timing of events in the relationship between P and A_i is as follows. First, the period 1 contract (i.e., α_1, β_1, and ϵ_1) is set by P. Then e_{1i} is chosen privately by A_i and e_{1j} by A_j. Then x_{1i}, x_{1j}, and w_{1i} are determined. Next, the period 2 contract (i.e., α_2, β_2, and ϵ_2) is

[7] When firms have identical technologies, the optimal sequence of one-period contracts yields exactly the same efforts and welfare as the optimal renegotiation-proof long-term contract (see Gibbons and Murphy 1992).

[8] See Holmstrom and Milgrom (1987, 1991) or Milgrom and Roberts (1992, chap. 7) for a justification of the restriction to linear contracts in similar models.

set by P. Then e_{2i} is chosen by A_i and e_{2j} by A_j, and finally x_{2i}, x_{2j}, and w_{2i} are determined.

In setting the contracts, P must observe three types of constraints. First, there are the incentive compatibility constraints: A_i will choose e_{1i} and e_{2i} to maximize his expected utility. Thus $C'(e_{2i}) = e_{2i} = \beta_2$, and e_{1i} will be chosen with a view to its effect on w_{2i} as well as on w_{1i} (see below). The second type of constraint on P is the time-consistency constraint that, at the start of period 2, she will set the contract terms for that period to maximize her expected payoff using the information gained from period 1.

Third, there are participation constraints for A_i: at the start of each period, entering or continuing the relationship with P must offer A_i expected utility at least as high as his reservation level. Agent A_i's reservation level at the start of period 1 is exogenous, so the period 1 participation constraint can be written as $ACE \geq \bar{u}$, where ACE is given by (8) and $\bar{u}$ is a constant. Agent A_i's period 2 reservation utility, however, is influenced by his (and by A_j's) performance in the first period. Good performance by A_i improves his outside options, as in the model of managerial career concerns in Section IIA. In particular, competition in contracts among the identical firms for A_i's services in period 2 implies that for every unit increase in the total period 2 certainty equivalent, A_i's period 2 reservation utility, expressed in certainty equivalent terms, increases by one unit. (Period 2 certainty equivalents are evaluated conditional on period 1 outputs x_{1i} and x_{1j}.) Thus A_i in effect has all the period 2 bargaining power in this setting. It is illuminating, however, to consider a more general form of the period 2 participation constraint: that for every unit increase in the total period 2 certainty equivalent, A_i's period 2 reservation certainty equivalent increases by $b \in [0, 1]$. The parameter b can be seen as a measure of A_i's period 2 bargaining power.[9]

[9] A reservation utility constraint of this form with $b < 1$ might be implied by (i) a bargaining game between P and A_i in period 2, in contrast to the scenario of perfect competition among firms described above, or (ii) period 2 competition among firms, but with the sensitivity of A_i's period 2 output to his intrinsic productivity being lower in the event that he changes firms. If one were to develop a model along the lines of point ii, one would need to recognize the possibility that if A_i's expected productivity after period 1 were extremely low, it would be efficient ex post for him to change firms. However, that possibility would be of negligible importance as long as the gap between his ex ante expected outputs at P's firm and at other firms (or the fixed cost of changing firms) were large in relation to σ^2. In modeling point ii, one would also have to be mindful of the take-the-money-and-run problem mentioned in n. 5 above. That problem would not exist if A_i received a sufficiently large lump-sum payment in period 2 for remaining with his period 1 firm, e.g., a deferred payment, which would be part of a renegotiation-proof two-period contract, as discussed by Laffont and Tirole (1990). In fact, the outcome derived in subsection B below is the same as the outcome of the optimal renegotiation-proof two-period contract.

Principal P's problem of maximizing (5) subject to the incentive compatibility, time-consistency, and participation constraints can be simplified as follows. Using the (binding) first-period participation constraint to substitute for $E(w_{1i}) + E(w_{2i})$ in (5) converts the maximand into

$$e_{1i} - \tfrac{1}{2}(e_{1i})^2 + e_{2i} - \tfrac{1}{2}(e_{2i})^2 - \tfrac{1}{2}r\,\mathrm{var}(w_{1i} + w_{2i}) - \bar{u} \equiv W - \bar{u}. \quad (9)$$

Hence the optimal sequence of contracts maximizes W, the sum of P's and A_i's certainty equivalents, and we can take W as our measure of total welfare. Note that $\bar{u}$ will affect only P's choice of lump-sum transfer α_1; it will have no impact on the optimal coefficients β_t and ϵ_t, on the optimal effort levels, or on total welfare.

In a first-best world, in which P could perfectly observe e_{ti}, P would induce A_i to choose $e^*_{ti} = 1$, and A_i would face no wage risk. It follows that the first-best welfare level $W^* = 1$. The welfare loss relative to the first-best can be expressed as

$$L = 1 - W = \tfrac{1}{2}[(1 - e_{1i})^2 + (1 - e_{2i})^2 + r\,\mathrm{var}(w_{1i} + w_{2i})]. \quad (10)$$

It is often more convenient to work with L than with W.

A. *Optimal Incentives in the One-Period Version of the Model*

The two-period model must be solved backward. Insight into the final period can be obtained from considering the one-period version of the model, which is of interest anyway as a benchmark. In the one-period model, A_i's reservation utility is exogenous and (like $\bar{u}$ above) has no effect on the optimal values of β and ϵ in $w_i = \alpha + \beta x_i + \epsilon x_j$ or on the optimal value of e_i. These values minimize the one-period welfare loss l relative to the first-best (the one-period analogue of L in [10]), subject to the incentive compatibility constraint that A_i chooses e_i to satisfy $C'(e_i) = e_i = \beta$. The problem facing P is thus equivalent to choosing β and ϵ to minimize

$$\begin{aligned} l &= \tfrac{1}{2}[(1 - \beta)^2 + r\,\mathrm{var}(\beta x_i + \epsilon x_j)] \\ &= \tfrac{1}{2}[(1 - \beta)^2 + r\sigma^2(\beta^2 + \epsilon^2 + 2\beta\epsilon\kappa)]. \end{aligned} \quad (11)$$

For any value of β, the optimal ϵ minimizes the variance of w_i; $\epsilon = -\beta[\mathrm{cov}(x_i, x_j)/\mathrm{var}(x_j)] = -\beta\kappa$. Thus the optimal contract has the form w_i = constant + $\beta[x_i - E(x_i|x_j)]$, and $\mathrm{var}(w_i) = \beta^2\,\mathrm{var}(x_i|x_j) = \beta^2 v_1\sigma^2$ (see [3b]). The optimal β is given by

$$\beta = \frac{1}{1 + r\sigma^2 v_1}, \tag{12}$$

and the minimized value of l is

$$l = \tfrac{1}{2}(1 - \beta) = \frac{\tfrac{1}{2}r\sigma^2 v_1}{1 + r\sigma^2 v_1}.$$

Define the strictly increasing loss function

$$\lambda(v) = \frac{\tfrac{1}{2}r\sigma^2 v}{1 + r\sigma^2 v},$$

which satisfies $\lambda(0) = 0$ and $\lambda(\infty) = \tfrac{1}{2}$. Then we have

$$l = \lambda(v_1). \tag{13}$$

The value of CPI is evident: for any given β, basing i's wage on x_j as well as on x_i reduces its variance to $v_1 = 1 - \kappa^2$ times its original level. As a result, the welfare loss relative to the first-best falls from $\lambda(1)$ to $\lambda(v_1)$. This is precisely the insurance effect at work: with CPI, P can filter out some of the uncertainty, either about a_i or about u_i, that prevents accurate estimation of e_i and can thereby lower the risk cost of providing incentives to A_i. In the one-period model, CPI always improves welfare and results in effort closer to its first-best level.

B. Sequentially Optimal Incentives in the Two-Period Model

In the two-period model, P chooses contracts period by period subject to providing A_i with a period 2 certainty equivalent that rises with the total period 2 certainty equivalent at rate b; A_i works for P in both periods. The final stage is the same as the one-period model just analyzed, with the addition of information from period 1. Write the conditional expectation $E(x_{2i}|x_{1i}, x_{1j}, x_{2j})$ in the form

$$E(x_{2i}|x_{1i}, x_{1j}, x_{2j}) = \hat{e}_{2i} + \gamma(x_{1i} - \hat{e}_{1i}) + \delta_1(x_{1j} - \hat{e}_{1j}) + \delta_2(x_{2j} - \hat{e}_{2j}), \tag{14}$$

where $\hat{e}_{tk}$ denotes the conjecture (which is correct in equilibrium) about k's effort in period t. In period 2, given any values of α_2 and

β_2, P chooses ϵ_2 to minimize $\text{var}(w_{2i}|x_{1i}, x_{1j}) = \text{var}(\beta_2 x_{2i} + \epsilon_2 x_{2j}|x_{1i}, x_{1j})$, so the optimal

$$\epsilon_2 = -\beta_2 \left[\frac{\text{cov}(x_{2i}, x_{2j}|x_{1i}, x_{1j})}{\text{var}(x_{2j}|x_{1i}, x_{1j})} \right] = -\beta_2 \delta_2.$$

This yields

$$\text{var}(w_{2i}|x_{1i}, x_{1j}) = (\beta_2)^2 \text{var}(x_{2i}|x_{1i}, x_{1j}, x_{2j}) = (\beta_2)^2 v_2 \sigma^2$$

(see [3d]). Hence optimal effort incentives in period 2 are given by

$$e_{2i} = \beta_2 = \frac{1}{1 + r\sigma^2 v_2}, \tag{15}$$

obtained by replacing v_1 by v_2 in (12). Thus the period 2 loss relative to the first-best is $\lambda(v_2)$ (see [13]).

At this point it might be useful to anticipate some general features of the analysis to follow. We have just seen that, as in the one-period model, CPI reduces the period 2 loss in the dynamic model by lowering the risk cost of providing period 2 incentives. This is of course an instance of the insurance effect. However, equation (15) for optimal period 2 incentives is also a *constraint* on P's period 1 optimization problem, which, as was shown earlier (see [9]), is equivalent to the maximization of W. In general, CPI will have implications for this constraint—and, through this route, for optimal efforts and welfare—as well as for risk reduction. The analysis below reveals that the ratchet effect is the key to understanding the implications of this constraint. The bargaining power/reputation effect, on the other hand, does not influence optimal efforts and welfare (see proposition 3 below), essentially because the solution to the problem of maximizing W does not depend on how the second-period surplus is shared.

To explore the implications of CPI for effort incentives and welfare given (15), consider A_i's effort incentive in period 1. In addition to the explicit incentive β_1, there are two types of implicit incentive, both arising from the adjustment of the α_2 term, given the observed x_{1i} and x_{1j}, to meet A_i's period 2 participation constraint. This binding constraint has the form $ACE_2 = s + b(TCE_2)$, where s is an arbitrary constant,

$$ACE_2 = \alpha_2 + \beta_2 E(x_{2i} - \delta_2 x_{2j}|x_{1i}, x_{1j}) - \tfrac{1}{2}(\hat{e}_{2i})^2 - \tfrac{1}{2} r(\beta_2)^2 v_2 \sigma^2,$$

and

$$TCE_2 = E(a_i | x_{1i}, x_{1j}) + \hat{e}_{2i} - \tfrac{1}{2}(\hat{e}_{2i})^2 - \tfrac{1}{2}r(\beta_2)^2 v_2 \sigma^2.$$

Solving for α_2 and substituting into the wage contract (7) yields

$$\begin{aligned} w_{2i} &= \text{constant} + bE(a_i | x_{1i}, x_{1j}) \\ &\quad + \beta_2 [x_{2i} - \delta_2 x_{2j} - E(x_{2i} - \delta_2 x_{2j} | x_{1i}, x_{1j})] \\ &= \text{constant} + bE(a_i | x_{1i}, x_{1j}) + \beta_2 [x_{2i} - E(x_{2i} | x_{1i}, x_{1j}, x_{2j})], \end{aligned} \tag{16}$$

where constant $= s + b\hat{e}_{2i} + (1 - b)[\tfrac{1}{2}(\hat{e}_{2i})^2 + \tfrac{1}{2}r(\beta_2)^2 v_2 \sigma^2]$, which is independent of all output levels. The overall *effective* incentive coefficient in period 1, $\tilde{\beta}_1$, which is the coefficient on x_{1i} in A_i's *overall* wage $w_{1i} + w_{2i}$, is therefore given by

$$\begin{aligned} \tilde{\beta}_1 &= \beta_1 + b \frac{\partial}{\partial x_{1i}} E(a_i | x_{1i}, x_{1j}) - \beta_2 \frac{\partial}{\partial x_{1i}} E(x_{2i} | x_{1i}, x_{1j}, x_{2j}) \\ &= \beta_1 + b\psi - \beta_2 \gamma, \end{aligned} \tag{17}$$

and A_i's first-period effort satisfies

$$e_{1i} = \tilde{\beta}_1. \tag{18}$$

The first of the two types of implicit incentives in (17) is the bargaining power/reputation effect that A_i gains fraction b of each unit increase in his expected ability $E(a_i | x_{1i}, x_{1j})$. The coefficient on x_{1i} in this conditional expectation is $\psi = \tau[(1 - \eta\kappa)/(1 - \kappa^2)]$ as in Section II, so the size of this component of implicit incentives is $b\psi$. The second type of implicit incentive is the ratchet effect, which reduces w_{2i} by $\beta_2\gamma$ for each unit increase in x_{1i}, where γ as given by (4) is the coefficient on x_{1i} in $E(x_{2i} | x_{1i}, x_{1j}, x_{2j})$. To understand the source of the ratchet effect, suppose that b is zero, so A_i's period 2 reservation utility is independent of first-period outcomes. Then satisfying A_i's binding period 2 participation constraint requires basing A_i's period 2 wage not on x_{2i} alone but on $x_{2i} - E(x_{2i} | x_{1i}, x_{1j}, x_{2j})$, that is, on performance relative to expectations held at the start of period 2. The larger β_2 is, the more A_i is penalized for high expectations about period 2 performance. Even when b is strictly positive, second-period compensation continues to be based on the deviation of actual performance from its expected level, but this term is now augmented by the term $bE(a_i | x_{1i}, x_{1j})$, which ensures that A_i's period 2 certainty equivalent varies with first-period outcomes at the appropriate rate.

To derive the optimal contract in period 1, define $\tilde{\epsilon}_1$, by analogy with $\tilde{\beta}_1$, as the coefficient on x_{1j} in i's overall wage $w_{1i} + w_{2i}$. Then given the preceding analysis, we can express P's problem as choosing the effective coefficients $\tilde{\beta}_1$ and $\tilde{\epsilon}_1$ to minimize L, as defined in (10), subject to (15) and (18). Now

$$\mathrm{var}(w_{1i} + w_{2i}) = \mathrm{var}(\tilde{\beta}_1 x_{1i} + \tilde{\epsilon}_1 x_{1j} + \beta_2 x_{2i} + \epsilon_2 x_{2j}),$$

and with ϵ_2 set optimally at $-\beta_2\delta_2$, we have

$$\begin{aligned}\mathrm{var}(w_{1i} + w_{2i}) &= \mathrm{var}[(\tilde{\beta}_1 + \beta_2\gamma)x_{1i} + (\tilde{\epsilon}_1 + \beta_2\delta_1)x_{1j} \\ &\quad + \beta_2(x_{2i} - \gamma x_{1i} - \delta_1 x_{1j} - \delta_2 x_{2j})] \\ &= \mathrm{var}[(\tilde{\beta}_1 + \beta_2\gamma)x_{1i} + (\tilde{\epsilon}_1 + \beta_2\delta_1)x_{1j}] \\ &\quad + \mathrm{var}\{\beta_2[x_{2i} - E(x_{2i}|x_{1i}, x_{1j}, x_{2j})]\}.\end{aligned}$$

For any given $\tilde{\beta}_1$, the optimal $\tilde{\epsilon}_1$ minimizes the first term above, so

$$\tilde{\epsilon}_1 + \beta_2\delta_1 = -(\tilde{\beta}_1 + \beta_2\gamma)\frac{\mathrm{cov}(x_{1i}, x_{1j})}{\mathrm{var}(x_{1j})}.$$

Thus $\mathrm{var}(w_{1i} + w_{2i})$ reduces to

$$\begin{aligned}&(\tilde{\beta}_1 + \beta_2\gamma)^2\,\mathrm{var}(x_{1i}|x_{1j}) + (\beta_2)^2\,\mathrm{var}(x_{2i}|x_{1i}, x_{1j}, x_{2j}) \\ &\quad = (\tilde{\beta}_1 + \beta_2\gamma)^2 v_1\sigma^2 + (\beta_2)^2 v_2\sigma^2.\end{aligned}$$

The principal's problem in period 1 therefore reduces to choosing the effective incentive coefficient $\tilde{\beta}_1$ to minimize

$$L = \tfrac{1}{2}\{(1 - \tilde{\beta}_1)^2 + (1 - \beta_2)^2 + r\sigma^2[(\tilde{\beta}_1 + \beta_2\gamma)^2 v_1 + (\beta_2)^2 v_2]\} \tag{19}$$

subject to (15). Since $e_{1i} = \tilde{\beta}_1$ and $e_{2i} = \beta_2$, the following result is immediately apparent.

PROPOSITION 3. With sequentially optimal contracts, the effort levels e_{1i} and e_{2i} and the welfare loss L are independent of A_i's period 2 bargaining power b.

For given first-period explicit coefficients β_1 and ϵ_1, a change in b, which alters the strength of the reputation effect $b\psi$, alters both A_i's first-period effort e_{1i} and the risk premium $\tfrac{1}{2}r\,\mathrm{var}(w_{1i} + w_{2i})$. (There is no effect on second-period incentives or risk.) Nevertheless, both e_{1i} and $\tfrac{1}{2}r\,\mathrm{var}(w_{1i} + w_{2i})$ depend only on the *overall* wage $w_{1i} + w_{2i}$. Therefore, the welfare impact of a change in b can be costlessly offset by appropriate changes in β_1 and ϵ_1 that leave the effective coefficients $\tilde{\beta}_1$ and $\tilde{\epsilon}_1$, and hence $w_{1i} + w_{2i}$, unchanged.

Unlike changes in the reputation effect, changes in the ratchet effect $\beta_2\gamma$ do affect welfare. When $\beta_2\gamma$ increases (other things equal), although it is feasible for P to keep $\tilde{\beta}_1 = e_{1i}$ unchanged by increasing β_1 by the same amount, such an adjustment is not costless: it raises the risk cost of providing first-period incentives, $\frac{1}{2}r\sigma^2(\tilde{\beta}_1 + \beta_2\gamma)^2 v_1$ (see [19]).[10]

Proposition 3 reveals an important difference between the model of this section and the framework of Section II*C*. When we introduce sequentially optimal explicit contracts and risk-averse agents, the reputation and ratchet effects no longer have equal significance: only the ratchet effect influences effort levels and welfare.[11] This is true even though the net implicit incentive, $b\psi - \beta_2\gamma$, and hence the optimal explicit incentive $\tilde{\beta}_1$ depend on both the reputation and ratchet effects. Since we are focusing on the consequences of CPI for welfare and effort levels, all our results will be independent of the strength of the reputation effect $b\psi$.[12]

The optimal choice of $\tilde{\beta}_1$ satisfies

$$\tilde{\beta}_1 = \frac{1 - r\sigma^2 v_1 \beta_2\gamma}{1 + r\sigma^2 v_1}, \tag{20}$$

and the minimized value of L, when the correlation parameters are η and ρ, is

$$L(\eta, \rho) = \lambda(v_1)\{1 + \beta_2\gamma\}^2 + \lambda(v_2) \tag{21a}$$

$$= \lambda(v_1)\left\{\frac{\lambda(v_2)}{\lambda[(1-\tau)(1-\rho^2)]}\right\}^2 + \lambda(v_2). \tag{21b}$$

The final term in (21a), $\lambda(v_2)$, is the loss resulting from the period 2 uncertainty in x_{2i} conditional on x_{1i}, x_{1j}, and x_{2j}. The term $\lambda(v_1)$ is the loss resulting from the period 1 uncertainty in x_{1i} conditional on x_{1j} (compare [13] for the one-period model). In (21a), $\lambda(v_1)$ is multiplied by $\{1 + \beta_2\gamma\}^2 \geq 1$. The larger the ratchet effect $\beta_2\gamma$, the more costly it is to provide first-period incentives, and the larger the first period's contribution to the overall welfare loss.

[10] Since neither β_2 nor γ is an exogenous parameter, unlike b, changes in the ratchet effect are in practice accompanied by changes in other variables, in particular the conditional variances v_1 and v_2 (as detailed below). Nevertheless, the argument above isolates the effect on welfare of changes in the ratchet effect per se.

[11] With risk-neutral agents, sequentially optimal contracts could achieve the first-best outcome, so welfare would be independent of both the ratchet and the reputation effects (see Lazear 1986).

[12] Gibbons and Murphy (1992) examine the implication of changes in implicit incentives over the course of a manager's career for optimal *explicit* incentives. They do not consider CPI, nor do they examine welfare. Assuming that $b = 1$, they show that, over time, net implicit incentives decline and, in consequence, optimal explicit incentives rise.

C. *Analysis of the Effects of Comparative Performance Information*

We now examine the impact of CPI on incentives, risk, and overall welfare in our model of sequentially optimal contracts. We consider how dynamic incentives—notably the ratchet effect—modify the conclusion from the static analysis that CPI is always beneficial. In the dynamic setting, there may be a welfare cost from the impact of CPI on the ratchet effect that outweighs the static benefit of improved insurance, so CPI may lower overall welfare. On the other hand, there are circumstances in which not only is CPI beneficial for the ratchet effect, but this dynamic benefit exceeds the static one of improved insurance. We thus show that the dynamic effect of CPI can be larger in magnitude than the static effect, whether it opposes or reinforces it, and we characterize these possibilities in terms of the parameters η, ρ, τ, and $r\sigma^2$.

The absence of CPI is equivalent to the case $\eta = \rho = 0$, when the welfare loss relative to the first-best is

$$L(0, 0) = \lambda(1)\{1 + \beta_2^0\tau\}^2 + \lambda(1 - \tau^2)$$
$$= \lambda(1)\left\{\frac{\lambda(1 - \tau^2)}{\lambda(1 - \tau)}\right\}^2 + \lambda(1 - \tau^2), \tag{22}$$

where $\beta_2^0 = 1/[1 + r\sigma^2(1 - \tau^2)]$. The net welfare gain (or loss if negative) from CPI is therefore defined as

$$G(\eta, \rho) \equiv L(0, 0) - L(\eta, \rho). \tag{23}$$

Comparative performance information affects the net welfare gain in two ways. First, there is the *insurance gain*, which we define as

$$G_I(\eta, \rho) \equiv L(0, 0) - \lambda(v_1)\{1 + \beta_2^0\tau\}^2 - \lambda(v_2). \tag{24}$$

This is the net welfare change, with the term in braces in (22) held constant, from the reduction in the conditional variances of A_i's outputs x_{1i} and x_{2i} when A_j's outputs become observable. That is, G_I measures the gain from improved estimation of A_i's efforts e_{1i} and e_{2i}, abstracting from any change in the size of the ratchet effect, which is $\beta_2^0\tau$ in the absence of CPI. It is always the case that $G_I \geq 0$, because the conditional variances of both x_{1i} and x_{2i} fall as a result of CPI: the former from σ^2 to $v_1\sigma^2$ just as in the one-period model of subsection *A* and the latter from $(1 - \tau^2)\sigma^2$ to $v_2\sigma^2$.

Second, there is the *ratchet gain* (or *loss*), which we define as

$$G_R(\eta, \rho) \equiv \lambda(v_1)\{1 + \beta_2^0\tau\}^2 + \lambda(v_2) - L(\eta, \rho). \tag{25}$$

This measures the net gain or loss resulting from the influence of CPI on the ratchet effect, after allowing for the insurance gain: CPI changes the size of the ratchet effect from $\beta_2^0\tau$ to $\beta_2\gamma$.

Since $G = G_I + G_R$ and $G_I \geq 0$, it is clear that $G_R \geq 0$ is a sufficient condition for $G \geq 0$. The ratchet gain G_R is nonnegative if

$$\beta_2\gamma \leq \beta_2^0\tau. \tag{26}$$

However, $\beta_2 > \beta_2^0$, since period 2 incentives are inversely related to the conditional variance of x_{2i} (see [15]), which CPI reduces. As to the relative sizes of γ and τ, it is evident from (4) that

$$\gamma - \tau \overset{\text{sgn}}{=} (\rho - \eta)(\rho + \eta\tau). \tag{27}$$

It follows that if $\rho + \eta\tau \geq 0$, then the ratchet gain $G_R \leq 0$ whenever $\rho \geq \eta$. Under these conditions, therefore, the adverse consequences of CPI from the strengthening of the ratchet effect conflict with the beneficial effect of improved insurance, and the net welfare impact of CPI is ambiguous.

Two special cases in which the ratchet gain is zero, so that CPI is certain to increase welfare, are the cases $\tau = 0$ and $\tau = 1$. The parameter τ reflects the informativeness of period 1 outputs about period 2 performance and hence the strength of the intertemporal linkage. If $\tau = 0$, then $\gamma = 0$; so (26) is an equality, and $G_R = 0$. In this case, $G = 2[\lambda(1) - \lambda(1 - \rho^2)] \geq 0$, with strict inequality if $\rho \neq 0$. With $\tau = 0$, period 1 outputs provide no information about period 2 performance, so each period of the two-period model is equivalent to the static model (in which, with $\tau = 0$, $\kappa = \rho$).

If $\tau = 1$, then $\beta_2^0 = 1$ and also $\gamma = 1$ and $\beta_2 = 1$, for all (η, ρ). Therefore, (26) is an equality, and $G_R = 0$. In this case, $G = 4[\lambda(1) - \lambda(1 - \eta^2)] \geq 0$, with strict inequality if $\eta \neq 0$. With $\tau = 1$, the informativeness of period 1 outputs about period 2 performance is at its greatest ($\text{var}[x_{2i}|x_{1i}] = 0$), so the ratchet effect is at its strongest ($\beta_2\gamma = 1$), even in the absence of CPI. Therefore, CPI makes the ratchet effect no worse, and so operates only via the insurance gain (in period 1). Hence $G \geq 0$.

Another case in which the ratchet gain $G_R = 0$, so that the overall gain $G > 0$, occurs when $\kappa^2 = 1$ (which arises when $\eta = \rho = 1$ or $\eta = \rho = -1$). In this case, CPI reduces the conditional variances of both x_{1i} and x_{2i} to zero and hence allows the first-best outcome to be achieved.

We can summarize the results above in the following proposition.

PROPOSITION 4. Each of the following conditions is sufficient for the ratchet gain G_R to be zero and for G, the net welfare impact of

CPI, to be strictly positive: (i) $\tau = 0$ and $\rho \neq 0$, (ii) $\tau = 1$ and $\eta \neq 0$, or (iii) $\kappa^2 = 1$.

Of course, G can be positive even if $G_R < 0$. For example, as long as $1 - \kappa^2 \leq 1/7$ and $r\sigma^2 \leq 1$, we can be certain that $G > 0$. These conditions ensure that $\lambda(1)/\lambda(v_1) \geq 4$ and hence that the insurance benefit in the first period outweighs even the largest possible strengthening of the ratchet effect (from $\beta_2^0\tau = 0$ to $\beta_2\gamma = 1$).

On the other hand, it is entirely possible that CPI reduces welfare (which cannot happen in the static model). For example, if $\eta = 0$, $\rho = 1$, and $\tau \in [1/2, 1)$, then $G < 0$ (see proposition 7 below). To characterize the range of parameter values for which CPI is good/bad, we shall focus on η and ρ in the "positive quadrant," that is, $(\eta, \rho) \in [0, 1]^2$, with τ and $r\sigma^2$ held fixed for the time being. Of particular interest is the locus of points in (η, ρ) space such that $G(\eta, \rho) = 0$. The example just given in which $G < 0$ illustrates that such a locus will pass through the interior of the positive quadrant for at least some values of τ and $r\sigma^2$. The following statement summarizes some comparative static results on η and ρ.

PROPOSITION 5. Given $\eta \in [0, 1]$ and $\rho \in [0, 1]$, (i) $\text{var}(x_{1i}|x_{1j})$ and $\text{var}(x_{2i}|x_{1i}, x_{1j}, x_{2j})$ are both decreasing in η and in ρ; (ii) γ is decreasing in η; (iii) γ is increasing in ρ if $\rho \geq \eta$; (iv) $L(\eta, \rho)$ is decreasing in η; and (v) $L(\eta, \rho)$ is decreasing in ρ if $\eta = 1$, and at $\rho = 0$ for all $\eta > 0$.

Results iv and v imply that in (η, ρ) space, iso-welfare contours cut the $\eta = 1$ boundary and the $\rho = 0$ axis with a negative slope.

To develop additional sufficient conditions for CPI to increase welfare, an informative case to examine is $\rho = \eta$. If either (i) risk aversion is not too important ($r\sigma^2 \leq 1$, which also guarantees that $e_{1i} = \tilde{\beta}_1 \geq 0$) or (ii) the intertemporal linkage is not too strong ($\tau \leq \sqrt{3} - 1$), then we can show that welfare increases as the common value of ρ and η, which is simply κ, increases. Although increasing κ strengthens the ratchet effect (because β_2 increases while γ remains unchanged at τ), the bound on either $r\sigma^2$ or τ ensures that this change is outweighed by the insurance gain. Combining this result with result iv in proposition 5 gives the following proposition.

PROPOSITION 6. The net welfare impact of CPI, G, is strictly positive if $\eta \geq \rho \geq 0$ (with at least one strict inequality) and either $r\sigma^2 \leq 1$ or $\tau \leq \sqrt{3} - 1$.

When CPI weakens the ratchet effect, this dynamic benefit can sometimes even exceed the static insurance gain: G_R can actually exceed G_I. For example, if $\rho = 0$, then as $\tau \to 0$, $G_R/G_I \to 4/3$ for all $r\sigma^2$ and for all $\eta \neq 0$.

On the other hand, CPI reduces welfare if G_R is larger in magnitude than G_I but negative. Since welfare is increasing in η, $G_R/G_I < -1$ is most likely when $\eta = 0$. In fact, if $\eta = 0$, then as $\tau \to 1$, $G_R/G_I \to -2(1 + 2r)/(1 + r) < -2$ for all $r\sigma^2$ and for all $\rho \neq 0$.

We can learn more about when CPI is likely to reduce welfare. Numerical analysis suggests that, when $\eta = 0$, the net welfare impact of CPI is more likely to be negative the larger ρ is. (As ρ approaches one, the ratchet effect $\beta_2\gamma$ approaches its maximum size of one.) For ($\eta = 0$, $\rho = 1$), the following proposition shows that CPI is more likely to lower welfare, the stronger the intertemporal linkage is (the larger τ is) and the more important risk aversion is (the larger $r\sigma^2$ is).

PROPOSITION 7. Assume that $\eta = 0$ and $\rho = 1$. Then for all $r\sigma^2 \in (0, \infty)$, there exists $\bar{\tau}(r\sigma^2) \in (0, 1/2)$ such that G, the net welfare impact of CPI, is (i) strictly positive for $\tau < \bar{\tau}(r\sigma^2)$, (ii) strictly negative for $\tau \in (\bar{\tau}(r\sigma^2), 1)$, and (iii) zero for $\tau = \bar{\tau}(r\sigma^2)$ or $\tau = 1$. Furthermore, $\bar{\tau}(r\sigma^2)$ is a strictly decreasing function of $r\sigma^2$, $\lim_{r\sigma^2 \to 0} \bar{\tau}(r\sigma^2) = 1/2$, and $\lim_{r\sigma^2 \to \infty} \bar{\tau}(r\sigma^2) = 0$.

Proposition 7 highlights how the welfare impact of CPI can differ according to whether or not implicit incentives are present. In the one-period version of the model, in which all incentives are explicit, the welfare loss is $\lambda(0)$ in the absence of CPI and $\lambda[1 - (1 - \tau)^2]$ given CPI with ($\eta = 0$, $\rho = 1$). Thus for all $\tau < 1$ and for all $r\sigma^2$, CPI improves welfare, and the improvement is greater the smaller τ is. (Smaller τ means that relatively more of the uncertainty in outputs stems from the transient shocks, which are perfectly correlated when $\rho = 1$.) In the dynamic model, with implicit incentives affecting welfare through the ratchet effect, it remains true that for τ small, CPI with ($\eta = 0$, $\rho = 1$) improves welfare. However, for any $r\sigma^2$, there is a critical value of τ above which CPI with ($\eta = 0$, $\rho = 1$) reduces welfare, and this critical value is decreasing in $r\sigma^2$.

Figure 1 illustrates the welfare effects of the parameters of the model. Each box in the figure shows iso-welfare contours in (η, ρ) space for $(\eta, \rho) \in [0, 1]^2$. Welfare is increasing in η (by part iv of proposition 5), and the first-best is attained when $\eta = \rho = 1$. Risk aversion is more important in the right-hand column ($r\sigma^2 = 0.5$) than in the left ($r\sigma^2 = 0.1$). The parameter τ decreases in steps from one to zero as one moves down the rows. In the top (bottom) row, welfare is independent of ρ (η), and so the contours are vertical (horizontal). In both cases, $G_R = 0$. The contours flatten as τ decreases, reflecting the growing significance of ρ relative to η. (Recall that $\text{var}[u_{tk}]/\text{var}[a_k] = [1 - \tau]/\tau$.) In the shaded regions, which exist if $\tau \in [1/2, 1)$ (by proposition 7), CPI reduces welfare ($G < 0$). The contours have a negative slope if $\rho = 0$ or $\eta = 1$ (and $\tau \in (0,$

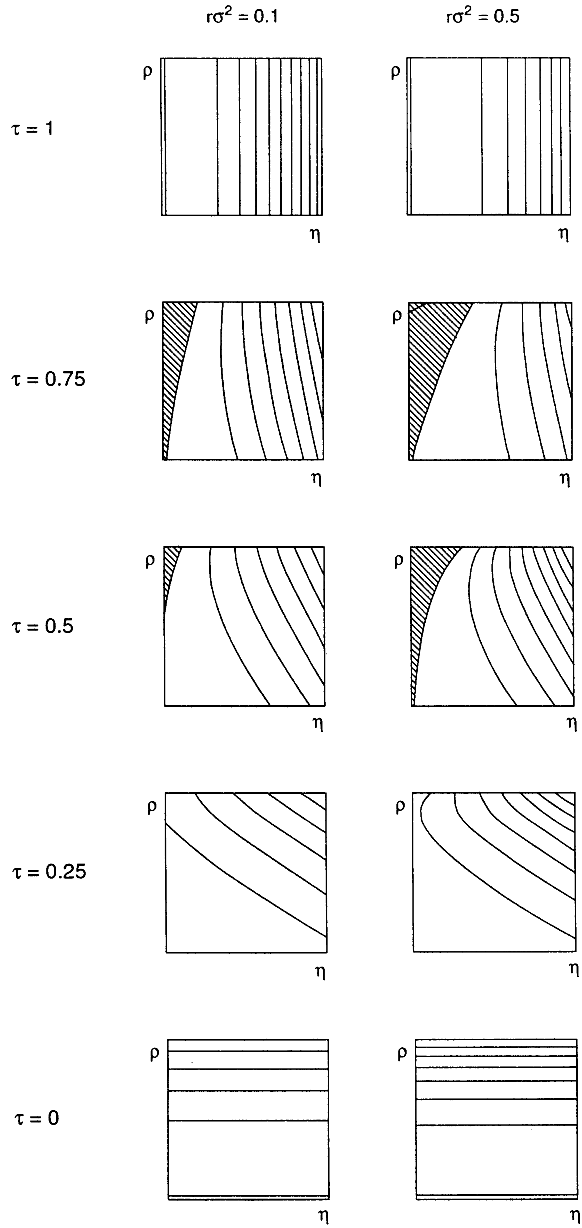

FIG. 1.—Iso-welfare contours in (η, ρ) space; in shaded regions, CPI reduces welfare.

1))—see part v of proposition 5—but can have either slope if $\rho = 1$ or $\eta = 0$. Thus it is ambiguous whether welfare is increasing in ρ in the latter cases.

We now summarize some of our main findings. Comparative performance information is always beneficial if $\tau = 0$ (no intertemporal linkage) or $\tau = 1$ (no transient shocks) because then it causes no change in the ratchet effect. Higher correlation in intrinsic productivities, η, increases the insurance gain and also eases the ratchet problem. Higher correlation in transient shocks, ρ, increases the insurance gain but usually worsens the ratchet problem. The ratchet gain/loss can be more important than the insurance gain in the sense that $G_R > G_I$ and $-G_R > G_I$ are both possible. Thus the dynamic effect of CPI can have significance comparable to that of the static insurance effect.

IV. Applications and Extensions

A. *An Application to Job Swapping*

The analysis so far has assumed that, whether or not CPI is available, agents continue to perform the *same* job for two periods. However, if A_i and A_j both work for P, then P has available an alternative means of generating CPI: she can have the agents swap jobs between periods 1 and 2. We now examine the consequences, for implicit incentives and welfare, of CPI when agents swap jobs.

If each agent performs the same job for two periods, then it is irrelevant whether the random variable a represents a characteristic of the agent or of his job. By contrast, if agents swap jobs between periods, then incentives and risk are affected by the source of the variation in a. For simplicity, we assume here that a is a purely job-specific productivity variable. That is, if agents swap jobs between periods and if we let a_k denote the uncertain productivity of the job performed by agent k in period 1, then period 2 outputs are given by

$$\begin{aligned} \tilde{x}_{2i} &= e_{2i} + a_j + u_{2i}, \\ \tilde{x}_{2j} &= e_{2j} + a_i + u_{2j}. \end{aligned} \tag{28}$$

Period 1 outputs are still given by (1), and the parameters η, ρ, and τ are still defined as in Section III.

With job swapping, A_i's period 2 contract depends on P's beliefs about a_j, and A_i's implicit incentives in period 1 arise from the sensitivity of these beliefs to i's first-period output. The ratchet effect is represented by $\beta_2(\partial/\partial x_{1i})E(\tilde{x}_{2i}|x_{1i}, x_{1j}, \tilde{x}_{2j})$, which, in terms of the outputs in the original model, is $\beta_2(\partial/\partial x_{1j})E(x_{2i}|x_{1i}, x_{1j}, x_{2j}) \equiv \beta_2\delta_1$.

The optimal value of β_2 is still given by (15). The period 1 and period 2 conditional variances of A_i's output are each the same with job swapping as without and equal $v_1\sigma^2$ and $v_2\sigma^2$, respectively. It follows that with job swapping, the welfare loss in P's relationship with A_i, $L^{js}(\eta, \rho)$, is given by

$$L^{js}(\eta, \rho) = \lambda(v_1)\{1 + \beta_2\delta_1\}^2 + \lambda(v_2). \tag{29}$$

Comparison of (29) with (21a) shows that the insurance gain from introducing CPI is the same whether or not P employs job swapping. On the other hand, the ratchet effect is always weaker with job swapping than without. This is so since, by (15), β_2 is the same in the two cases, and $\delta_1 \leq \gamma$: when agents swap jobs, each one's expected period 2 output is less sensitive to his period 1 performance than when they do not. It thus follows that, when CPI is available, welfare is higher with job swapping than without.

We can also show that the introduction of CPI with job swapping is certain to raise welfare relative to the case in which the second agent is absent, as long as risk aversion is not too important. This follows because either δ_1 is negative, so that CPI with job swapping actually generates a ratchet effect that enhances incentives, or CPI raises welfare even without job swapping.

These observations are summarized in the following proposition.

PROPOSITION 8. Let the random variable a represent a purely job-specific attribute. (i) When CPI is available, welfare when agents swap jobs between periods 1 and 2 is at least as high as when they do not, for all $(\eta, \rho) \in [-1, 1]^2$, for all $\tau \in [0, 1]$, and for all $r\sigma^2 > 0$. (ii) If $r\sigma^2 \leq 1$, then welfare from P's relationship with A_i is at least as high when CPI is available and agents swap jobs as in the absence of A_j, for all $(\eta, \rho, \tau) \in [0, 1]^3$.

Other authors (e.g., Ickes and Samuelson 1987) have noted that job rotation can be useful in organizations as a means of mitigating the ratchet effect arising when employees possess job-specific private information. Our analysis extends this observation by showing that, when CPI is available, a special type of job rotation, namely job swapping, is attractive: job-swapping leaves the risk-reduction benefits of CPI unchanged, while reducing the severity of the ratchet effect. Furthermore, when risk aversion is not too important and job swapping is feasible, then the introduction of CPI always raises welfare.

B. *Extensions*

In the regulation model of Section II*B*, we assumed that the two firms were separately owned and behaved noncooperatively. A natural question to ask in that setting is how joint ownership of the firms would affect incentives and efficiency if under joint ownership externalities between firms were internalized. It is straightforward to show that the ratchet effect is weaker, and hence efficiency is greater, under joint ownership than under separate ownership if and only if the correlation in transient shocks, ρ, exceeds the correlation in intrinsic efficiencies, η.

A very similar issue arises in the employment model of Section III if both agents work for the same principal. Instead of paying the agents on the basis of relative performance and encouraging them to compete (i.e., behave noncooperatively), the principal could base their wages on their joint output and enable them to monitor one another, thus encouraging them to behave cooperatively. In static analyses comparing "competition" and "cooperation," the preferred mode of job design is the one under which the insurance effect is stronger (see Holmstrom and Milgrom 1990; Itoh 1992). Meyer (1995) extends the model of Section III to compare competition and cooperation in a dynamic setting with limited precommitment. In such a setting, the relative efficiency of these different approaches to designing jobs depends not only on their contributions to the static insurance effect but also on their dynamic impacts on the ratchet effect. Just as we found in Section III, the dynamic effects can be the more important ones.

Dynamic incentive effects are also important for other issues in job design, for example, the assignment of tasks to agents. When should two agents be made jointly responsible for the performance of a single task? When tasks differ, should those assigned to each agent be homogeneous or heterogeneous? Static analyses of these questions have hinged on the effect of task assignments on the insurance-incentive trade-off (see Holmstrom and Milgrom 1991; Itoh 1992). Meyer, Olsen, and Torsvik's (1996) extension of our analysis shows that, in a dynamic setting with limited precommitment, the prescriptions from static analyses can be reversed. Decisions about task assignments have important dynamic impacts on the ratchet effect, and in environments in which the ratchet problem is severe, these dynamic effects can outweigh the static insurance considerations.[13]

A final extension is to examine whether the results of our two-agent analysis of CPI are robust to the introduction of additional agents. In our discussion paper (Meyer and Vickers 1994), we sup-

[13] In a related vein, Olsen and Torsvik (1995) and Olsen (1996) emphasize the importance of the ratchet effect in evaluating organizational design decisions such as decentralization and vertical integration.

pose that P can observe the performance of N symmetrically placed agents, each producing output according to (1), and examine the limiting case as $N \rightarrow \infty$. The main lesson we draw is that our earlier conclusions are indeed robust. With $N \rightarrow \infty$, the effects of CPI on implicit incentives can still have the same order of significance as its risk-reduction benefits. Just as with $N = 2$, CPI with $N \rightarrow \infty$ can exacerbate the ratchet effect by so much as to outweigh the insurance gain. Nevertheless, we do find one contrast: whereas CPI with $N = 2$ can reduce welfare even for arbitrarily small values of risk aversion, CPI with $N \rightarrow \infty$ can never be harmful if risk aversion is small.

V. Summary and Conclusion

The analysis in this paper illustrates a general point about the value of comparative performance information in agency relationships. If, as in the standard one-period principal-agent problem, the principal faces information constraints only, then CPI always has positive value because it effectively eases the information constraints. In particular, CPI allows explicit incentives to be provided at lower cost in terms of risk; we have termed this the insurance effect. If, however, the principal in addition faces constraints on precommitment possibilities, then the effect of CPI on these constraints can also have considerable significance.

In the simple models in Section II, no explicit incentives can be designed. Comparative performance information influences incentives and efficiency only via implicit incentives, and specifically via the reputation effect and the ratchet effect. (In the model of managerial career concerns, only the reputation effect, which enhances incentives, is present; in the yardstick regulation model, the only source of implicit incentives is the ratchet effect, which dampens incentives.) Whether CPI raises or lowers efficiency is shown to depend on the correlation between agents' intrinsic characteristics relative to that between the transitory shocks affecting their performances. Also important is the relative strength of the reputation and ratchet effects when they are simultaneously present.

Section III studied a two-period model in which explicit incentives are choice variables but contracts are constrained to be sequentially optimal. That is, the principal cannot commit ex ante not to set period 2 contract terms other than optimally in the light of information gained in period 1. It is shown that, with this dynamic consistency constraint, CPI can significantly influence the ratchet effect, which is a key determinant of overall incentives and of welfare.

(While CPI also influences the reputation effect on incentives, this effect is now irrelevant for welfare as long as explicit contracts can be designed.) Comparative performance information reduces the ratchet effect, and through this route improves welfare, if the correlation between the transitory shocks to agents' performances is not too great relative to that between their intrinsic characteristics. However, the influence of CPI on the ratchet effect can be detrimental to welfare in other circumstances, and it is quite possible for this cost to outweigh the familiar benefit of CPI in terms of risk reduction. Thus, when there are constraints on the principal's powers of precommitment, it is essential in evaluating the impact of CPI to take account of its influence on the ratchet effect.

Appendix

A. Derivation of Equations (3d) and (4)

The conditional expectation $E(x_{2i}|x_{1i}, x_{1j}, x_{2j})$ has the linear form

$$E(x_{2i}|x_{1i}, x_{1j}, x_{2j}) = \hat{e}_{2i} + \gamma(x_{1i} - \hat{e}_{1i}) + \delta_1(x_{1j} - \hat{e}_{1j}) + \delta_2(x_{2j} - \hat{e}_{2j}). \tag{A1}$$

Then $\text{var}(x_{2i}|x_{1i}, x_{1j}, x_{2j}) \equiv v_2\sigma^2$ and $(\gamma, \delta_1, \delta_2)$ are derived by solving

$$\begin{aligned}\min_{\gamma,\delta_1,\delta_2} E\{[x_{2i} - \hat{e}_{2i} - \gamma(x_{1i} - \hat{e}_{1i}) - \delta_1(x_{1j} - \hat{e}_{1j}) - \delta_2(x_{2j} - \hat{e}_{2j})]^2\} \\ = \min_{\gamma,\delta_1,\delta_2} [1 + \gamma^2 + (\delta_1)^2 + (\delta_2)^2 - 2\kappa(\delta_2 - \gamma\delta_1) \\ - 2\tau(\gamma - \delta_1\delta_2) - 2\eta\tau(\delta_1 - \gamma\delta_2)]\sigma^2.\end{aligned} \tag{A2}$$

The first-order conditions associated with γ, δ_1, and δ_2, respectively, are

$$\begin{aligned} 0 &= \gamma + \kappa\delta_1 - \tau + \eta\tau\delta_2, \\ 0 &= \delta_1 + \kappa\gamma + \tau\delta_2 - \eta\tau, \\ 0 &= \delta_2 - \kappa + \tau\delta_1 + \eta\tau\gamma. \end{aligned} \tag{A3}$$

Equation (4) solves these simultaneous equations for γ. With (A2) they combine to give the expression for $v_2\sigma^2$ in (3d).

B. Derivation of Equations (20) and (21)

Differentiate the welfare loss L in (19) with respect to $\tilde{\beta}_1$ to get the first-order condition

$$0 = -(1 - \tilde{\beta}_1) + r\sigma^2(\tilde{\beta}_1 + \beta_2\gamma)v_1, \tag{A4}$$

which implies (20). Multiply the right-hand side of (A4) by $\frac{1}{2}(1 - \tilde{\beta}_1)$, and add the result, which equals zero, to (19) to obtain

$$\begin{aligned}L &= \tfrac{1}{2}\{(1-\beta_2)^2 + r\sigma^2[(1+\beta_2\gamma)(\tilde{\beta}_1+\beta_2\gamma)v_1 + (\beta_2)^2 v_2]\}\\ &= \tfrac{1}{2}[(1-\beta_2)^2 + r\sigma^2(\beta_2)^2 v_2] + \tfrac{1}{2}r\sigma^2 v_1(1+\beta_2\gamma)(\tilde{\beta}_1 - 1 + 1 + \beta_2\gamma)\\ &= \lambda(v_2) + \tfrac{1}{2}r\sigma^2 v_1(1+\beta_2\gamma)^2\left(\frac{-r\sigma^2 v_1}{1+r\sigma^2 v_1}+1\right) \qquad \text{(A5)}\\ &= \lambda(v_2) + (1+\beta_2\gamma)^2\lambda(v_1),\end{aligned}$$

which is (21a). (The derivation of the penultimate line above used [15] and [20].) Furthermore,

$$\begin{aligned}1+\beta_2\gamma &= 1-\beta_2+\beta_2(1+\gamma)\\ &= 1-\beta_2+\frac{\beta_2 v_2}{(1-\tau)(1-\rho^2)} \quad \text{from (3d)}\\ &= \frac{r\sigma^2 v_2}{1+r\sigma^2 v_2}\left[1+\frac{1}{r\sigma^2(1-\tau)(1-\rho^2)}\right] \quad \text{from (15)} \qquad \text{(A6)}\\ &= \frac{\lambda(v_2)}{\lambda[(1-\tau)(1-\rho^2)]}.\end{aligned}$$

Equations (A5) and (A6) imply (21b).

C. *Proof of Proposition 5*

Part i.—For $\eta \geq 0$ and $\rho \geq 0$, $\kappa^2 = [\eta\tau + \rho(1-\tau)]^2$ is increasing, and therefore $v_1 = (1-\kappa^2)$ is decreasing, in both η and ρ. Applying the envelope theorem to (A2) yields

$$\frac{\partial v_2}{\partial \eta} = -2\tau(\delta_2 - \gamma\delta_1 + \delta_1 - \gamma\delta_2) = -2\tau(1-\gamma)(\delta_2+\delta_1).$$

Equations (A3) imply that $\delta_2 + \delta_1 \geq 0$ and $\gamma \leq 1$. Therefore, v_2 is decreasing in η. The same method implies that

$$\frac{\partial v_2}{\partial \rho} = -2(1-\tau)[(1-\gamma)\delta_1 + (1+\gamma)\rho] \leq 0$$

and hence that v_2 is decreasing in ρ also.

Parts ii and iii.—Differentiate γ as given by (4).

Part iv.—Using part i of the proposition and the definition of $\lambda(\cdot)$, we know that $\lambda(v_1)$ and $\lambda(v_2)$ are decreasing in η. Since $(1-\tau)(1-\rho^2)$ is independent of η, it follows from (21b) that L is decreasing in η.

Part v.—We set $\eta = 1$ and use (21a) to evaluate

$$\frac{\partial L}{\partial \rho} = \frac{\partial L}{\partial v_1}\frac{\partial v_1}{\partial \rho} + \frac{\partial L}{\partial v_2}\frac{\partial v_2}{\partial \rho} + \frac{\partial L}{\partial \gamma}\frac{\partial \gamma}{\partial \rho}. \qquad \text{(A7)}$$

From part i, $\partial v_1/\partial\rho \leq 0$ and $\partial v_2/\partial\rho \leq 0$. Differentiation of (4) when $\eta = 1$ shows that $\partial\gamma/\partial\rho \leq 0$. Clearly, $\partial L/\partial v_1 \geq 0$ and $\partial L/\partial\gamma \geq 0$. We show below that when $\eta = 1$, $\partial L/\partial v_2 \geq 0$. It then follows from (A7) that $\partial L/\partial\rho \leq 0$.

From (21a),

$$\frac{\partial L}{\partial v_2} = 2\gamma\lambda(v_1)(1 + \beta_2\gamma)\frac{\partial \beta_2}{\partial v_2} + \lambda'(v_2)$$

$$= \frac{\partial \beta_2}{\partial v_2}[2\gamma\lambda(v_1)(1 + \beta_2\gamma) - 1/2] \quad \text{from (15)}$$

$$\stackrel{\text{sgn}}{=} 1/2 - 2\gamma\lambda(v_1)(1 + \beta_2\gamma).$$

To show that $\partial L/\partial v_2 \geq 0$, it is sufficient to prove that $\lambda(v_1)(1 + \beta_2\gamma) \leq 1/2$ since, for $\eta = 1$, (4) implies that $\gamma \leq 1/2$:

$$1/2 - \lambda(v_1)(1 + \beta_2\gamma) \stackrel{\text{sgn}}{=} 1 - \frac{r\sigma^2 v_1}{1 + r\sigma^2 v_1}(1 + \beta_2\gamma) \tag{A8}$$

$$\stackrel{\text{sgn}}{=} 1 + r\sigma^2 v_2 - r\sigma^2 v_1\gamma.$$

We now show that $v_2 \geq v_1\gamma$, which implies that (A8) ≥ 0 and hence that $\partial L/\partial v_2 \geq 0$:

$$v_2 \geq v_1\gamma \Leftrightarrow (1 + \gamma)(1 - \tau)(1 - \rho^2)$$

$$\geq \gamma(1 + \kappa)(1 - \tau)(1 - \rho) \quad \text{when } \eta = 1$$

$$\Leftrightarrow (1 + \gamma)(1 + \rho) \geq \gamma(1 + \kappa)$$

$$\Leftrightarrow 1 + \rho \geq \gamma\tau(1 - \rho) \quad \text{when } \eta = 1.$$

This last inequality is always true.

To prove the second part of part v, fix $\eta > 0$ and consider $\partial L/\partial \rho$ at $\rho = 0$. From part i, we know that $\lambda(v_1)$ and $\lambda(v_2)$ are decreasing in ρ. The derivative with respect to ρ of the term in braces in (21b) is

$$\frac{\lambda'(v_2)(\partial v_2/\partial \rho)}{\lambda[(1 - \tau)(1 - \rho^2)]} + \frac{2\rho(1 - \tau)\lambda'[(1 - \tau)(1 - \rho^2)]\lambda(v_2)}{\{\lambda[(1 - \tau)(1 - \rho^2)]\}^2}.$$

With $\rho = 0$ and $\eta > 0$, the first of these terms is negative and the second is zero. Thus the term in braces is decreasing in ρ, and therefore the same is true of L.

D. *Proof of Proposition 6*

Given part iv of proposition 5, it suffices to establish that if $\rho = \eta = \kappa$ and either $r\sigma^2 \leq 1$ or $\tau \leq \sqrt{3} - 1$, then L is decreasing in κ. Since, with $\rho = \eta$, $v_1 = 1 - \kappa^2$ and $v_2 = (1 - \tau^2)v_1$, this is equivalent to establishing that

$$\tilde{L}(v_1) \equiv \lambda(v_1)\left\{\frac{\lambda[(1 - \tau^2)v_1]}{\lambda[(1 - \tau)v_1]}\right\}^2 + \lambda[(1 - \tau^2)v_1]$$

is increasing in v_1. From the fact that $z\lambda'(z) = \lambda(z)[1 - 2\lambda(z)]$, it follows that

$$\frac{v_1}{\lambda(v_2)}\frac{d\tilde{L}}{dv_1} = \frac{\lambda(v_1)\lambda(v_2)}{\{\lambda[(1-\tau)v_1]\}^2}\{1 - 2\lambda(v_1) - 4\lambda(v_2) + 4\lambda[(1-\tau)v_1]\} + [1 - 2\lambda(v_2)]. \tag{A9}$$

Suppose that $r\sigma^2 \leq 1$. Then it can be shown that $\lambda[(1-\tau)v_1] > 2\lambda(v_1)\lambda(v_2)$ and $\lambda(v_2) < 1/4$. Therefore,

$$1 - 2\lambda(v_1) - 4\lambda(v_2) + 4\lambda[(1-\tau)v_1]$$
$$> 1 - 2\lambda(v_1) - 4\lambda(v_2) + 8\lambda(v_1)\lambda(v_2) = [1 - 2\lambda(v_1)][1 - 4\lambda(v_2)] > 0.$$

Hence the right-hand side of (A9) is positive if $r\sigma^2 \leq 1$, so L is decreasing in κ, as required.

Finally, define $Q \equiv r\sigma^2(1-\kappa^2)$. Straightforward algebraic manipulation shows that

$$\frac{d\tilde{L}}{dv_1} \overset{\text{sgn}}{=} -h(Q, \tau),$$

where $h(Q, \tau) = a_0(\tau) + a_1(\tau)Q + a_2(\tau)Q^2 + a_3(\tau)Q^3$; $a_0(\tau) = -2$; $a_1(\tau) = 2(1-\tau)(\tau^2 - \tau - 3) \leq 0$ if $\tau \in [0, 1]$; $a_2(\tau) = (1-\tau)(7\tau^2 + 2\tau - 6) \leq 0$ if $\tau \in [0, (\sqrt{43} - 1)/7]$; and $a_3(\tau) = (1-\tau)(1-\tau^2)(\tau^2 + 2\tau - 2) \leq 0$ if $\tau \in [0, \sqrt{3} - 1]$. Note that $\sqrt{3} - 1 < (\sqrt{43} - 1)/7 < 1$. If $\tau \leq \sqrt{3} - 1$, then $a_i(\tau) \leq 0$ for $i = 1, 2, 3$; so for all $Q \geq 0$, $h_Q(Q, \tau) \leq 0$ and therefore $h(Q, \tau) < 0$. Therefore, if $\tau \leq \sqrt{3} - 1$, L is decreasing in κ, as required.

E. *Proof of Proposition 7*

Define $R \equiv r\sigma^2$. The expression for $L(0, 0)$ is given by (22). With ($\eta = 0$, $\rho = 1$), $\gamma = 1$, $v_1 = \tau(2-\tau)$, $v_2 = 0$, and $\beta_2 = 1$; so, from (21a), $L(0, 1) = 2R\tau(2-\tau)/[1 + R\tau(2-\tau)]$. Lengthy but straightforward algebraic manipulation shows that

$$L(0, 0) - L(0, 1) \overset{\text{sgn}}{=} (1-\tau)g(\tau, R), \tag{A10}$$

where $g(\tau, R) = \sum_{k=0}^{5} b_k(R)\tau^k$; $b_0(R) = 2 + 4R + 2R^2$; $b_1(R) = -(4 + 14R + 14R^2 + 4R^3)$; $b_2(R) = -(5R + 6R^2 + 2R^3)$; $b_3(R) = 7R + 14R^2 + 6R^3$; $b_4(R) = R^2 + 2R^3$; and $b_5(R) = -(3R^2 + 2R^3)$. Clearly, if $\tau = 1$, $L(0, 0) = L(0, 1)$.

The argument below, which is valid for all $R \geq 0$, establishes that $g(\tau, R)$ has a unique τ-root in (0, 1) and crosses the τ-axis from above.

Step i.—$g_{\tau\tau\tau}(\tau, R)$ is quadratic in τ with a negative coefficient on τ^2, and $g_{\tau\tau\tau}(0, R) > 0$. Hence $g_{\tau\tau\tau}(\tau, R)$ has at most one τ-root in (0, 1] and at such a root crosses the τ-axis from above.

Step ii.—Step i, coupled with $g_{\tau\tau}(0, R) < 0$ and $g_{\tau\tau}(1, R) > 0$, implies that $g_{\tau\tau}(\tau, R)$ has a unique τ-root in (0, 1) and crosses the τ-axis from below.

Step iii.—Step ii, coupled with $g_{\tau}(0, R) < 0$, implies that $g_{\tau}(\tau, R)$ has at most one τ-root in (0, 1] and at such a root crosses the τ-axis from below.

Step iv.—Step iii, coupled with $g(0, R) > 0$ and $g(1, R) < 0$, implies that $g(\tau, R)$ has a unique τ-root in (0, 1). This root, $\bar{\tau}(R)$, is such that if $\tau \lesseqgtr \bar{\tau}(R)$, $g(\tau, R) \gtreqless 0$. This proves parts i–iii of the proposition.

Now express $g(\tau, R)$ in the form

$$g(\tau, R) = c_0(\tau) + c_1(\tau)R + c_2(\tau)R^2 + c_3(\tau)R^3, \quad \text{(A11)}$$

where $c_0(\tau) = 2 - 4\tau$, $c_1(\tau) = 4 - 14\tau - 5\tau^2 + 7\tau^3$, $c_2(\tau) = 2 - 14\tau - 6\tau^2 + 14\tau^3 + \tau^4 - 3\tau^5$, and $c_3(\tau) = -4\tau - 2\tau^2 + 6\tau^3 + 2\tau^4 - 2\tau^5$. Since $\lim_{R\to 0} g(\tau, R) = 2 - 4\tau$, $\lim_{R\to 0} \bar{\tau}(R) = 1/2$. Also, since $c_i(1/2) < 0$ for $i = 1, 2, 3$, it follows, from step iv above, that $\bar{\tau}(R) \in (0, 1/2)$ for all $R > 0$.

To complete the proof, we use the following facts, which are easily verified:

$$c_0(\tau) > 0 \quad \forall\, \tau \in (0, 1/2), \quad \text{(A12a)}$$

$$c_3(\tau) < 0 \quad \forall\, \tau \in (0, 1/2], \quad \text{(A12b)}$$

$$\exists\, \hat{\tau} \in (0, 1/2) \text{ such that if } \tau \in (0, \hat{\tau}],\ c_1(\tau) \geq 0 \text{ and if } \tau \in (\hat{\tau}, 1/2],\ c_1(\tau) < 0, \quad \text{(A12c)}$$

$$c_2(\tau) < c_1(\tau) \quad \forall\, \tau \in (0, 1/2]. \quad \text{(A12d)}$$

To prove that $\bar{\tau}(R)$ is strictly decreasing in R, it is sufficient to show that $g_\tau(\tau, R) < 0$ for all $\tau \in [0, 1/2]$, for all R, and $g_R(\bar{\tau}(R), R) < 0$ for all R.

From (A10), it is easily checked that $g_\tau(1/2, R) < 0$ for all R and, from step iii above, it follows that $g_\tau(\tau, R) < 0$ for all $\tau \in [0, 1/2]$, for all R.

To show that $g_R(\bar{\tau}(R), R) < 0$ for all R, use (A11) to write

$$g_R(\bar{\tau}(R), R) = c_1(\bar{\tau}(R)) + 2Rc_2(\bar{\tau}(R)) + 3R^2 c_3(\bar{\tau}(R)) \quad \text{(A13a)}$$

$$= \frac{1}{R}[g(\bar{\tau}(R), R) - c_0(\bar{\tau}(R))] + \frac{1}{R}[g(\bar{\tau}(R), R) - c_0(\bar{\tau}(R)) - Rc_1(\bar{\tau}(R))] + R^2 c_3(\bar{\tau}(R)). \quad \text{(A13b)}$$

Suppose that R is such that $\bar{\tau}(R) \in (\hat{\tau}, 1/2]$. Then from (A12), $c_i(\bar{\tau}(R)) < 0$ for $i = 1, 2, 3$. Hence, from (A13a), $g_R(\bar{\tau}(R), R) < 0$. The only other possibility is that R is such that $\bar{\tau}(R) \in (0, \hat{\tau}]$. Then from (A12), $c_0(\bar{\tau}(R)) > 0$, $c_1(\bar{\tau}(R)) \geq 0$, and $c_3(\bar{\tau}(R)) < 0$. Since $g(\bar{\tau}(R), R) = 0$ by definition of $\bar{\tau}(R)$, it follows from (A13b) that $g_R(\bar{\tau}(R), R) < 0$. This completes the proof that $\bar{\tau}(R)$ is strictly decreasing in R.

To show that $\lim_{R\to\infty}\bar{\tau}(R) = 0$, observe, using (A11) and (A12b), that for any $\tau \in (0, 1/2)$, by choosing R large enough, we can ensure that $g(\tau, R) < 0$. Therefore, no strictly positive value of τ can equal $\lim_{R\to\infty}\bar{\tau}(R)$. Since $\bar{\tau}(R) \in (0, 1/2)$ for all $R > 0$, it follows that $\lim_{R\to\infty}\bar{\tau}(R) = 0$.

F. Proof of Proposition 8

Part i.—The coefficients γ, δ_1, and δ_2 are defined by (A1). From the formulas for ordinary least squares regression coefficients, it is easily checked that $(\gamma, \delta_1, \delta_2) \in [-1, 1]^3$ for all $\tau \in [0, 1]$ and for all $(\eta, \rho) \in [-1, 1]^2$. Manipulation of (A3) yields

$$(\gamma - \delta_1)(1 - \kappa) = \tau(1 - \eta)(1 + \delta_2).$$

Thus, for all $\tau \in [0, 1]$ and for all $(\eta, \rho) \in [-1, 1]^2$, $\gamma \geq \delta_1$, and hence $0 \leq 1 + \beta_2\delta_1 \leq 1 + \beta_2\gamma$. It then follows from comparison of (21a) and (29) that welfare with job swapping is at least as high as without job swapping, for all $r\sigma^2 > 0$.

Part ii.—Welfare in the absence of A_j is $L(0, 0)$, as given by (22). Suppose first that $\eta > \rho$ and $(\eta, \rho, \tau) \in [0, 1]^3$. Then

$$L(0, 0) > L(\eta, \rho) \quad \text{if } r\sigma^2 \leq 1, \text{ by proposition 6}$$
$$\geq L^{js}(\eta, \rho) \quad \text{by part i of proposition 8.}$$

Suppose instead that $\eta \leq \rho$ and $(\eta, \rho, \tau) \in [0, 1]^3$. Then it follows from (A3) that $\delta_1 \leq -\eta\tau \leq 0$, and hence $0 \leq 1 + \beta_2\delta_1 \leq 1 + \beta_2^0\tau$. This inequality, coupled with part i of proposition 5, implies that $L(0, 0) \geq L^{js}(\eta, \rho)$.

References

Antle, Rick, and Smith, Abbie. "An Empirical Investigation of the Relative Performance Evaluation of Corporate Executives." *J. Accounting Res.* 24 (Spring 1986): 1–39.

Armstrong, C. Mark; Cowan, Simon G. B.; and Vickers, John S. *Regulatory Reform: Economic Analysis and British Experience.* Cambridge, Mass.: MIT Press, 1994.

Baker, George; Gibbons, Robert; and Murphy, Kevin J. "Subjective Performance Measures in Optimal Incentive Contracts." *Q.J.E.* 109 (November 1994): 1125–56.

Besley, Timothy J., and Case, Anne. "Incumbent Behavior: Vote-Seeking, Tax-Setting, and Yardstick Competition." *A.E.R.* 85 (March 1995): 25–45.

Cremer, Jacques. "Arm's Length Relationships." *Q.J.E.* 110 (May 1995): 275–95.

Fama, Eugene F. "Agency Problems and the Theory of the Firm." *J.P.E.* 88 (April 1980): 288–307.

Freixas, Xavier; Guesnerie, Roger; and Tirole, Jean. "Planning under Incomplete Information and the Ratchet Effect." *Rev. Econ. Studies* 52 (April 1985): 173–91.

Gibbons, Robert, and Murphy, Kevin J. "Relative Performance Evaluation for Chief Executive Officers." *Indus. and Labor Relations Rev.* 43 (suppl.; 1990): S30–S51.

———. "Optimal Incentive Contracts in the Presence of Career Concerns: Theory and Evidence." *J.P.E.* 100 (June 1992): 468–505.

Green, Jerry R., and Stokey, Nancy L. "A Comparison of Tournaments and Contracts." *J.P.E.* 91 (June 1983): 349–64.

Holmstrom, Bengt. "Managerial Incentive Schemes—A Dynamic Perspective." In *Essays in Economics and Management in Honour of Lars Wahlbeck.* Helsinki: Swenska Handelshogkolan, 1982. (*a*)

———. "Moral Hazard in Teams." *Bell J. Econ.* 13 (Autumn 1982): 324–40. (*b*)

Holmstrom, Bengt, and Milgrom, Paul R. "Aggregation and Linearity in the Provision of Intertemporal Incentives." *Econometrica* 55 (March 1987): 303–28.

———. "Regulating Trade among Agents." *J. Institutional and Theoretical Econ.* 146 (March 1990): 85–105.
———. "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design." *J. Law, Econ. and Organization* 7 (suppl.; 1991): 24–52.
Holmstrom, Bengt, and Tirole, Jean. "The Theory of the Firm." In *Handbook of Industrial Organization,* vol. 1, edited by Richard Schmalensee and Robert D. Willig. Amsterdam: North-Holland, 1989.
Ickes, Barry W., and Samuelson, Larry. "Job Transfers and Incentives in Complex Organizations: Thwarting the Ratchet Effect." *Rand J. Econ.* 18 (Summer 1987): 275–86.
Itoh, Hideshi. "Cooperation in Hierarchical Organizations: An Incentive Perspective." *J. Law, Econ. and Organization* 8 (April 1992): 321–45.
Laffont, Jean-Jacques, and Tirole, Jean. "Adverse Selection and Renegotiation in Procurement." *Rev. Econ. Studies* 57 (October 1990): 597–625.
———. *A Theory of Incentives in Procurement and Regulation.* Cambridge, Mass.: MIT Press, 1993.
Lazear, Edward P. "Salaries and Piece Rates." *J. Bus.* 59 (July 1986): 405–31.
Lazear, Edward P., and Rosen, Sherwin. "Rank-Order Tournaments as Optimum Labor Contracts." *J.P.E.* 89 (October 1981): 841–64.
Meyer, Margaret A. "Cooperation and Competition in Organizations: A Dynamic Perspective." *European Econ. Rev.* 39 (April 1995): 709–22.
Meyer, Margaret A., and Mookherjee, Dilip. "Incentives, Compensation, and Social Welfare." *Rev. Econ. Studies* 54 (April 1987): 209–26.
Meyer, Margaret A.; Olsen, Trond E.; and Torsvik, Gaute. "Limited Intertemporal Commitment and Job Design." *J. Econ. Behavior and Organization* (December 1996).
Meyer, Margaret A., and Vickers, John. "Performance Comparisons and Dynamic Incentives." Discussion Paper no. 96. Oxford: Nuffield Coll., 1994.
Milgrom, Paul R., and Roberts, D. John. *Economics, Organization, and Management.* Englewood Cliffs, N.J.: Prentice-Hall, 1992.
Mookherjee, Dilip. "Optimal Incentive Schemes with Many Agents." *Rev. Econ. Studies* 51 (July 1984): 433–46.
Nalebuff, Barry J., and Stiglitz, Joseph E. "Prizes and Incentives: Towards a General Theory of Compensation and Competition." *Bell J. Econ.* 14 (Spring 1983): 21–43.
Olsen, Trond E. "Agency Costs and the Limits of Integration." *Rand J. Econ.* 27 (Autumn 1996): 479–501.
Olsen, Trond E., and Torsvik, Gaute. "Intertemporal Common Agency and Organizational Design: How Much Decentralization?" *European Econ. Rev.* 39 (August 1995): 1405–28.
Rosen, Sherwin. "Prizes and Incentives in Elimination Tournaments." *A.E.R.* 76 (September 1986): 701–15.
Shleifer, Andrei. "A Theory of Yardstick Competition." *Rand J. Econ.* 16 (Autumn 1985): 319–27.
Theil, Henri. *Principles of Econometrics.* Amsterdam: North-Holland, 1971.
Zwiebel, Jeffrey. "Corporate Conservatism and Relative Compensation." *J.P.E.* 103 (February 1995): 1–25.