

# Incentives Between Firms (and Within)

Robert Gibbons\*

February 6, 2001

This paper reviews the significant progress in “agency theory” (*i.e.*, the economic theory of incentives) during the 1990s, with a special focus on applications to supply transactions. I begin with a brief summary of the classic agency model, which explores the tradeoff between incentives and insurance, but then emphasize six recent models that all abstract from the issue of risk-aversion. These six models come in three pairs: (1) new foundations for the theory of incentive contracts, including “get what you pay for” problems and subjective performance assessments; (2) new directions in incentive theory (*i.e.*, beyond incentive contracts), including investing in capabilities and career concerns; and (3) new applications to supply transactions, including hold-up and relational contracts within versus between firms.

Let me emphasize from the beginning that my knowledge of supply transactions is rudimentary and my mathematics worse. What I therefore hope to bring to this discussion are (a) reasonable familiarity with the recent economics literature on incentives and (b) some ability to organize and translate these models into a form that can travel. For example, in the discussion of investing in capabilities I borrow a model from labor economics (of promotion decisions for workers in firms) and translate it into a model of supplier management.

## I. The Classic Agency Model

To set the stage, I begin with a brief description of the classic model in agency theory, which explores the tradeoff between incentives and insurance. By treating this model briefly, I do not mean to imply that this tradeoff is irrelevant, but rather that other issues are now deemed at least as important, especially for incentive problems between firms (where risk-aversion might be less important than for individuals).<sup>1</sup> Because this classic model is merely a point of departure for this survey, I also do not explore any applications of this model to supply relationships.

---

\* Sloan Distinguished Professor of Management, MIT’s Sloan School (rgibbons@mit.edu) and Research Associate, National Bureau of Economic Research. This paper draws some of its non-mathematical exposition from Gibbons (1997) and Gibbons (1998), but those surveys omitted the mathematical analyses and the focus on supply relationships presented here. I thank Bengt Holmstrom, Ed Lazear, Canice Prendergast, and Mike Waldman for years of discussions on these issues, and especially George Baker and Kevin J. Murphy for not only discussions but also co-authorship of many of the ideas reported here.

<sup>1</sup> See Prendergast (1998) for a full-fledged discussion of theory and evidence related to the classic agency model.

The classic model in agency theory involves an Agent who takes an action to produce output. The Principal owns the output but contracts to share it with the Agent by paying a wage contingent on output. There is noise in the production function, so the Agent's output is uncertain. The timing of events is as follows:

1. The Principal and the Agent sign a compensation contract  $w(y)$ .
2. The Agent chooses an action ( $a$ ) but the Principal cannot observe this choice.
3. Events beyond the Agent's control ( $\varepsilon$ ) occur.
4. The action and the noise term determine the Agent's output ( $y$ ).
5. The Agent receives the compensation specified by the contract.

The Agent is risk-averse so there is a tradeoff between incentives and insurance. Paying a constant wage,  $w(y) = w_0$ , would provide the Agent with full insurance but no incentive; selling the Agent the firm for a fee of  $F$  (or, equivalently, paying the Agent  $w(y) = y - F$ ) would provide the Agent with full incentives but no insurance. The efficient contract lies between these extremes.

An intuitive closed-form solution can be derived in the linear-normal-exponential case. The production function is linear,  $y = a + \varepsilon$ , where  $\varepsilon$  is a normally distributed noise term with zero mean and variance  $\sigma^2$ . The incentive contract is linear,  $w(y) = s + by$ , where the intercept  $s$  is the salary and the slope  $b$  is the bonus rate. The Agent's utility function is  $U(x) = -e^{-rx}$ , where  $r > 0$  is the Agent's coefficient of absolute risk aversion and  $x = w - c(a)$  is the Agent's net payoff—the realized wage minus the convex disutility of action  $c(a)$ . The Principal is risk-neutral and so seeks to maximize the expected value of profit,  $y - w$ .

Given the contract  $w(y) = s + by$ , the Agent's problem is

$$\max_a \int_{-\infty}^{\infty} -e^{-r[s+ba-c(a)+\varepsilon]} e^{-rb\varepsilon} \phi(\varepsilon) d\varepsilon,$$

where  $\phi(\varepsilon)$  denotes the normal density function. The Agent's optimal action, denoted  $a^*(b)$ , therefore solves  $c'(a) = b$  and the Agent's maximized certainty equivalent is

$$CE(s, b) = s + ba^*(b) - c[a^*(b)] - \frac{1}{2} rb^2 \sigma^2.$$

The Principal's expected profit is

$$E \pi(s, b) = (1 - b)a^*(b) - s,$$

so total surplus is

$$CE(s, b) + E \pi(s, b) = a^*(b) - c[a^*(b)] - \frac{1}{2} rb^2 \sigma^2.$$

We can now determine what contract slope  $b$  maximizes the total surplus. This slope, denoted  $b^*$ , is efficient: if the parties agreed to a contract with some other slope then both parties

could be made better off by switching to a contract with slope  $b^*$  and choosing an appropriate value of  $s$  to distribute the increased total surplus. The first-order condition for  $b^*$  is  $a^{*'} - c'a^{*'} - rb^{*2} = 0$ . Because  $c'[a^*(b)] = b$ , we have  $a^{*'} = 1/c''$  and hence

$$b^* = \frac{1}{1 + r\sigma^2 c''}.$$

This result makes sense. Since  $r$ ,  $\sigma^2$ , and  $c''$  are positive,  $b^*$  is between zero (full insurance) and one (full incentives). Furthermore,  $b^*$  is smaller if the Agent is more risk-averse ( $r$  is higher) or there is more uncertainty in production ( $\sigma^2$  is higher) or marginal disutility increases more quickly ( $c''$  is higher).

This solution to the classic model is tidy but flawed: Mirrlees (1974) showed that the best linear contract,  $w = s + b^*y$ , is inferior to various non-linear contracts. In particular, a step-function contract (where the Agent earns  $w_H$  if  $y \geq y_0$  but  $w_L < w_H$  if  $y < y_0$ ) can perform very well, approaching the twin goals of full incentives and full insurance in the limit (as  $y_0$  and  $w_L$  decrease in appropriate fashion, so that the Agent almost surely receives  $w_H$  and yet has incentives from fear of  $w_L$ ). Mirrlees's result prompted much research on how the optimal contract depends on the details of the utility function and the conditional distribution of output given the Agent's action. In brief, this work showed that the optimal contract in the classic agency model is extremely sensitive to these details. In particular, the optimal contract is linear only under very special assumptions about the utility function and the conditional distribution of output; see Holmstrom (1979) and Banker and Datar (1989).

Holmstrom and Milgrom (1987) reinterpreted the classic agency model so as to rescue linear contracts. Rather than a single action ( $a$ ) that influences a single outcome ( $y$ ), Holmstrom and Milgrom envision a sequence of actions influencing a corresponding sequence of outcomes (say, one action and one outcome per day, over the course of a year). There are no connections across days (*i.e.*, the action  $a_t$  on day  $t$  affects that day's outcome,  $y_t$ , but has no influence on any other day's outcome) and all past outcomes are observed before the next day's action is chosen. Under certain assumptions, it is optimal to repeat the same one-day contract every day, regardless of history. If the one-day output  $y_t$  is binary then the year-end aggregate pay ( $w = \sum y_t$ ) is a linear function of the year-end aggregate output ( $y = \sum y_t$ ).

In my view, the main contribution of this Holmstrom-Milgrom model is not that it justifies linear contracts (by imposing quite strong assumptions, most of which were not described above) but rather that its sequential-action model implicitly alerts us to gaming as a natural consequence of non-linear contracts. For example, a step-function contract of the kind studied by Mirrlees (in the classic one-action model) would induce no effort once the Agent's aggregate output to date passes the hurdle  $y_0$  (in the daily-action model). More generally, if the incentive contract for the year is a non-linear function of year-end aggregate output then the worker's incentives change from day to day, depending on the aggregate output to date.

A growing body of evidence is consistent with the prediction that non-linear contracts create history-dependent incentives: see Healy (1985) on bonus plans with ceilings and floors, Asch (1990) and Oyer (1998) on bonuses tied to quotas, and Brown, Harlow, and Starks (1996) and Chevalier and Ellison (1997) on how the convex relationship between mutual fund performance and assets under management caused risk-taking portfolio choices by ostensibly conservative funds. This issue of gaming caused by the shape of the pay plan leads naturally to our next topic: gaming caused by the performance measure itself, regardless of the contract's shape.

## II. New Foundations for the Theory of Incentive Contracts

In 1975, Steven Kerr published "On the Folly of Rewarding A, While Hoping for B." The argument was simple: you get what you pay for. Kerr distilled this unifying theme from a disparate set of examples involving politicians, soldiers, doctors, orphanage directors, professors, and students, as well as manufacturing and clerical employees and even human-resource managers. From these examples, Kerr (pp. 779-80) concluded that two main causes of distorted incentives are "fascination with an 'objective' criterion, [where] individuals seek to establish simple, quantifiable standards against which to measure and reward performance" and "overemphasis on highly visible behaviors, [when] some parts of the task are highly visible while others are not."

The following examples, drawn from a seemingly infinite supply, show that Kerr's observations on the folly of rewarding A while hoping for B remain true today.<sup>2</sup>

At the H.J. Heinz Company, division managers received bonuses only if earnings increased from the prior year. The managers delivered consistent earnings growth by manipulating the timing of shipments to customers and by prepaying for services not yet received, both at some cost to the firm. At Dun & Bradstreet, salespeople earned no commission unless the customer bought a larger subscription to the firm's credit-report services than in the previous year. In 1989, the company faced millions of dollars in lawsuits following charges that its salespeople deceived customers into buying larger subscriptions by fraudulently overstating their historical usage. In 1992, Sears abolished the commission plan in its auto-repair shops, which paid mechanics based on the profits from repairs authorized by customers. Mechanics misled customers into authorizing unnecessary repairs, leading California officials to prepare to close Sears' auto-repair business statewide. (Baker, Gibbons, and Murphy, 1994: 1125)

---

<sup>2</sup> There are of course exceptions, which I interpret as proving Kerr's rule. For instance, Lazear (2000) describes the transition from salaries to piece rates at a firm that installs auto windshields. The data show convincingly that output increased, due to two predicted effects: piece rates provided stronger incentives for hard work and also induced the self-selection towards a workforce that benefited from those incentives. Unfortunately, not all jobs are as narrowly defined and easily monitored as windshield installation. I conjecture that the same is true for supply transactions.

To my continuing embarrassment, the classic agency model cannot even express Kerr's title, not to mention to evaluate or extend his conclusions. Fortunately, recent work has brought agency theory not only to Kerr's position but beyond.

In this section I describe two recent models that reject a strong but unremarked assumption in the classic agency model: that  $y$  can be called "output," as though it could easily be measured. This usage hides the fact that in the classic model  $y$  reflects everything the Principal cares about, except for wages (*i.e.*, the Principal's payoff is  $y - w$ ). Therefore, I henceforth call  $y$  the Agent's "total contribution" and I henceforth assume that *no* contract based on  $y$  can be enforced in court, including but not limited to the linear contract  $w = s + by$ .

If the Agent's total contribution is not contractible then there are two alternatives: court-enforceable contracts based on alternative performance measures, as discussed in Section IIA, and "relational contracts" (*i.e.*, self-enforcing agreements perhaps based on total contribution), as discussed in Section IIB. I believe that these two models are the new foundations for agency theory. Since these models are foundational, I continue with the abstract terminology of "Principal" and "Agent" in this section, but switch to concrete terminology from supply chains in later sections.

#### A. *Objective Performance Measurement: Getting What You Pay For*

Even when contracts based on  $y$  are not available, other contracts can be enforced in court. Such contracts are based on alternative performance measures—such as the number of units produced, with limited adjustment made for quality, timely delivery, and so on. Let  $p$  denote such a performance measure and consider a linear incentive contract  $w = s + bp$ . As in the classic agency model, a large value of  $b$  will create strong incentives, but now the Agent's incentives are to produce a high value of  $p$ , not of  $y$ . But the Principal does not directly benefit from increased realizations of measured performance,  $p$ ; rather, the Principal benefits from increased realizations of the Agent's total contribution,  $y$ . In this sub-section, *the essence of the incentive problem is the divergence between the Agent's incentive to increase  $p$  and the Principal's desire to increase  $y$ .*

To illustrate this idea, consider first a simple extension of the classic agency model in which  $y = a_1 + a_2$  and  $p = a_1$ . In this case the contract  $w = s + bp$  creates incentives to increase  $p$  and the induced action also increases  $y$ . But now suppose that there are two kinds of actions (or "tasks") that the Agent can take,  $a_1$  and  $a_2$ . In this case the contract  $w = s + bp$  creates incentives that depend on the bonus rate  $b$  and on the way the actions  $a_1$  and  $a_2$  affect the performance measure  $p$ . For example, if  $y = a_1 + a_2$  and  $p = a_1$  then a contract based on  $p$  cannot create incentives for  $a_2$  and so misses this potential contribution to  $y$ . Alternatively, if  $y = a_1$  and  $p = a_1 + a_2$  then a contract based on  $p$  creates an incentive for the Agent to take action  $a_2$ , even though  $a_2$  is irrelevant to the Agent's total contribution. Finally, in an extreme case such as  $y = a_1$  and  $p = a_2$ , the contract  $w = s + bp$  creates no value at all.

All of the examples above are special cases of the "multi-task" model originated by Holmstrom and Milgrom (1991). As a tractable example, suppose that the technology of production

is  $y = f_1 a_1 + f_2 a_2 + \dots$ , the technology of performance measurement is  $p = g_1 a_1 + g_2 a_2 + \dots$ , the contract is  $w = s + bp$ , and the payoffs are  $y - w$  to the Principal and  $w - c(a_1, a_2)$  to the Agent.<sup>3</sup> To keep things simple, assume that

$$c(a_1, a_2) = \frac{1}{2} a_1^2 + \frac{1}{2} a_2^2,$$

but notice that this assumption rules out the potentially important case where the actions compete for the Agent's attention (*i.e.*, increasing the level of one action increases the marginal cost of the other).

The timing of events in this model mimics the classic agency model, except that it is modified to incorporate the new distinction here between  $y$  and  $p$ . First, the Principal and the Agent sign a compensation contract,  $w = s + bp$ . Second, the Agent chooses actions ( $a_1$  and  $a_2$ ) but the Principal cannot observe these choices. Third, events beyond the Agent's control (  $\epsilon$  and  $\eta$  ) occur. Fourth, the actions and the noise terms determine the Agent's total contribution ( $y$ ) and measured performance ( $p$ ). Fifth, measured performance is observed by the Principal and the Agent (and by a Court, if necessary).<sup>4</sup> Finally, the Agent receives the compensation specified by the contract.

For the remainder of this paper I depart from the classic model's central assumption that the Agent is risk-averse. In the multi-task setting described above, a risk-neutral Agent chooses the actions  $a_1$  and  $a_2$  to maximize the expected payoff  $E(w) - c(a_1, a_2)$ . The Agent's optimal actions are therefore  $a_1^*(b) = g_1 b$  and  $a_2^*(b) = g_2 b$ , so the Principal's expected payoff is

$$E(y - w) = f_1 a_1^*(b) + f_2 a_2^*(b) - s - g_1 a_1^*(b) - g_2 a_2^*(b),$$

and the Agent's expected payoff is

$$E(w) - c(a_1, a_2) = s + b[g_1 a_1^*(b) + g_2 a_2^*(b)] - \frac{1}{2} a_1^*(b)^2 - \frac{1}{2} a_2^*(b)^2.$$

The sum of these expected payoffs is the total surplus

$$E(y) - c(a_1, a_2) = f_1 a_1^*(b) + f_2 a_2^*(b) - \frac{1}{2} a_1^*(b)^2 - \frac{1}{2} a_2^*(b)^2,$$

which is maximized at the efficient value of  $b$ :

$$b^* = \frac{f_1 g_1 + f_2 g_2}{g_1^2 + g_2^2} = \frac{\|f\|}{\|g\|} \cos(\theta),$$

where  $\theta$  is the angle between the  $f = (f_1, f_2)$  and  $g = (g_1, g_2)$  vectors.

<sup>3</sup> See Feltham and Xie (1994), Kulp, Datar, and Lambert (1999), and Baker (2000) for analyses of richer versions of this example.

<sup>4</sup> For simplicity, in this model no one ever observes the Agent's total contribution, even though the Principal eventually receives the payoff  $y - w$ ! See Section II.B on relational contracts and subjective performance assessment for a more realistic approach to this issue.

There are two important features in  $b^*$ : scaling and alignment, reflected by  $\|f\|/\|g\|$  and  $\cos(\theta)$ , respectively. To understand scaling, imagine that  $g_1$  and  $g_2$  are both much larger than  $f_1$  and  $f_2$ . Then the Agent can greatly increase  $p$  by choosing high values of  $a_1$  and  $a_2$  but these actions do not greatly increase  $y$ . As a result, the efficient contract puts a small bonus rate on  $p$ , as shown by  $\|f\|/\|g\|$ . To understand alignment, imagine first that the  $f$  and  $g$  vectors lie almost on top of one another (regardless of their lengths). In this case the incentives created by paying on  $p$  are valuable for increasing  $y$ . Alternatively, imagine that the  $f$  and  $g$  vectors are almost orthogonal to each other. In this second case the incentives created by paying on  $p$  are almost useless for increasing  $y$ . As a result, the efficient contract puts a larger bonus rate on  $p$  when  $f$  and  $g$  are more closely aligned, as shown by  $\cos(\theta)$ .

Baker (2000) uses this model to dispel a persistent confusion about what makes a good performance measure. One might be tempted to guess that  $p$  is a good performance measure if it is highly correlated with  $y$ . But what determines the correlation between  $p$  and  $y$ ? Given technologies such as  $y = f_1 a_1 + f_2 a_2 + \epsilon$  and  $p = g_1 a_1 + g_2 a_2 + \eta$ , one important part of the answer involves the two variables in this model that we have not discussed thus far—the noise terms  $\epsilon$  and  $\eta$ . For example, suppose that  $p$  is a division's accounting earnings and  $y$  is the firm's stock price: both are likely to be hit by business-cycle variations (noise terms), but earnings reflect only short-run actions while the stock price also incorporates long-term actions. Thus, earnings and the stock price might be highly correlated because of their noise terms, even though paying on one creates distorted incentives for the other (namely, ignoring long-run actions). More generally,  *$p$  is a valuable performance measure if it induces valuable actions, not if it is highly correlated with  $y$* . In short, alignment is more important than noise.

The general theme of this multi-task theory of objective performance measurement is that it is no use creating strong incentives for the wrong actions. That is, if attaching a large bonus rate  $b$  to the performance measure  $p$  would create strong but distorted incentives then the optimal bonus rate may be quite small. The efficient bonus rate depends on scale and alignment. These lessons seem as applicable to contracts between firms as to those between individuals. Furthermore, the idea underlying this sub-section—that measured performance differs from total contribution—also motivates another approach to providing incentives: paying Agents for their total contributions by using subjective performance assessments rather than (only) objective performance measures. We consider this approach next.

### *B. Subjective Performance Assessments: Relational Contracts in Repeated Games*

Subjective assessments play important roles in many incentive contracts. Even foreign-exchange traders, whose books are marked to market at the end of each trading day, can have their incentive compensation tied to subjective judgments by their managers and co-workers. At Citicorp, for example, part of a trader's job is to execute orders that the bank's salespeople receive from clients, and part of the trader's bonus is based on subjective assessments by the salespeople

concerning the timing and terms on which the trader executed orders.<sup>5</sup> Similarly, Lincoln Electric is well-known for its use of piece-rate formulas that tie a worker's pay to that worker's output, but about half of a worker's compensation is a bonus based on the supervisor's subjective assessment of the worker's cooperation, innovation, dependability, and so on (Fast and Berg, 1975). Moving beyond case studies, Hayes and Schaefer (2000) report evidence consistent with the use of subjective assessments when boards of directors decide the salary and bonus of chief executives: variation in the executive's current cash compensation that is not explained by current performance measures (such as stock return, sales, and earnings) predicts future variation in these performance measures. Finally, moving from microdata to everyday experience, subjective assessments of current performance may play a crucial role in determining future compensation, promotions, and continued employment, even where current compensation does not involve incentive pay of any kind. That is, these future compensation, promotion, and employment decisions are not specified in advance as a function of objective performance measures.

But all of these examples are within firms. Are subjective assessments also important between firms? Simply put, yes. Supply chains often involve long-run, hand-in-glove supplier relationships through which the parties reach accommodations when unforeseen or uncontracted-for events occur. Similar relationships also exist in alliances, joint ventures, and business groups. In all these settings, "relational contracts" help firms circumvent difficulties in formal contracting (*i.e.*, contracting enforced by a third party, such as a court). For example, a formal contract must be specified *ex ante* in terms that can be verified *ex post* by the third party, whereas a relational contract can be based on outcomes that are observed by only the contracting parties *ex post*, and also on outcomes that are prohibitively costly to specify *ex ante*. A relational contract thus allows the parties to utilize their detailed knowledge of their specific situation and to adapt to new information as it becomes available. For the same reasons, however, relational contracts cannot be enforced by a third party and so must be self-enforcing: each party's reputation must be sufficiently valuable that neither party wishes to renege.

In this sub-section I present Bull's (1987) repeated-game model of a relational incentive contract, as distilled by Baker, Gibbons, and Murphy (1994). Following most of the literature, I will make the strong assumption that both the Principal and the Agent observe the Agent's contribution perfectly. The analysis would be essentially unchanged if the Principal and the Agent both observed an imperfect proxy for the Agent's total contribution and the Principal promised to pay a bonus based on this proxy (Baker, Gibbons and Murphy, 1994). But the analysis would be more difficult (and the predicted outcome less efficient) if the parties attempted to implement a relational contract based on a subjective assessment made by one party but observed only imperfectly by the other (Compte, 1998; Kandori and Matsushima, 1998).

---

<sup>5</sup> Personal communication in December 1997 from Julian Simmonds, Head of Foreign Exchange, Citicorp. See also Eccles and Crane (1988) for a similar description of the compensation of investment bankers.



Consider a repeated game between a single Principal and a single Agent.<sup>6</sup> In each period, the Agent chooses an unobservable action,  $a$ , that stochastically determines the Agent's total contribution,  $y$ . To keep things simple, assume that  $y$  is either High ( $y = H$ ) or Low ( $y = L < H$ ). It is then natural to define the Agent's action to be the probability that  $y$  is High:  $\text{Prob}\{y = H \mid a\} = a$ , where  $a \in [0, 1]$ . As motivated above, assume that the Agent's total contribution is too complex and subtle to be verified by a third party and so cannot be the basis of an enforceable contract. That is,  $y$  cannot be objectively measured. On the other hand, we assume that  $y$  can be subjectively assessed and used in a relational contract, as follows.

Imagine that compensation contracts consist of a base salary,  $s$ , and a relational-contract bonus  $b$  that the Principal promises to pay if the subjective assessment is  $y=H$ . The timing of events within each period is as follows. First, the Principal offers the Agent a compensation package  $(s, b)$ . Second, the Agent either accepts the compensation package or rejects it in favor of an alternative employment opportunity with payoff  $w_a$ . Third, if the Agent accepts then the Agent chooses an action  $a \in [0, 1]$  at cost  $c(a)$ . The Principal does not observe the Agent's action. Fourth, the Principal and the Agent observe the realization of the Agent's total contribution,  $y$ . Finally, if  $y = H$  then the Principal chooses whether to pay the Agent the bonus  $b$  specified in the relational contract.

The Principal's payoff when the Agent's total contribution is  $y$  and total compensation is  $w$  is  $y - w$ . The Principal's discount rate is  $r$ ; the Agent's discount rate is immaterial, because it is the Principal's reputation that is at stake. The Agent's payoff from choosing an action with convex cost  $c(a)$  and receiving total compensation  $w$  is  $w - c(a)$ , where  $c'(a) \rightarrow \infty$  as  $a \rightarrow 1$ . The first-best action,  $a^*$ , solves

$$\max_a L + a \cdot (H - L) - c(a),$$

so  $c'(a^*) = H - L$ .

In a single-period employment relationship the Principal will choose not to pay a bonus, so the Agent (anticipating the Principal's decision) will choose not to supply effort, so the Principal (anticipating the Agent's choice) will not pay a salary, so the Agent will choose not to work for the Principal. To formalize self-enforcing relational contracts, we consider an infinitely repeated relationship.<sup>7</sup> We assume that the Principal and the Agent play trigger strategies (roughly speaking, the parties begin by cooperating and then continue to cooperate unless one side defects, in which case they refuse to cooperate forever after) and ask whether these strategies are an equilibrium of the repeated game; see Baker, Gibbons, and Murphy (1994) for details.

---

<sup>6</sup> The single worker we consider could just as well be an infinite sequence of workers, each of whom lives for one period, provided that each period's worker learns the history of play before the period begins.

<sup>7</sup> The discount rate  $r$  can be reinterpreted so that the game is not infinitely repeated but instead concludes at an uncertain date: suppose that after each period is played a coin is flipped, and that if heads occurs then the game ends; if the probability of heads is  $q$  and the firm's actual discount rate is  $s$  then  $r = (s+q)/(1-q)$ .

Given a relational-contract bonus  $b$ , if the Agent believes the Principal will honor the relational contract then the Agent's problem is

$$\max_a s + a \cdot b - c(a),$$

so the Agent's optimal action satisfies  $c'(a) = b$ . The Agent will choose to work for the Principal if his expected payoff exceeds the alternative wage:

$$s + a^*(b) \cdot b - c[a^*(b)] \geq w_a.$$

If the Principal offers the minimum salary that the Agent will accept then the Principal's expected profit per period is

$$V(b) = a^*(b) \cdot b - c[a^*(b)] - w_a.$$

But should the Agent believe that the Principal will honor the relational contract? That is, if the Agent's contribution is  $y = H$ , will the Principal decide to pay the bonus  $b$ ? The answer depends on the Principal's discount rate,  $r$ . Given the Agent's strategy, if the Principal does not pay the bonus then its payoff is  $H - s$  this period but zero thereafter, whereas if the Principal does pay the bonus then its payoff is  $H - s - b$  this period but equal to the expected profit from the relationship thereafter. Thus, the Principal should pay the bonus if and only if the present value of the expected profit beginning next period exceeds the size of the bonus:

$$(H - s - b) + \frac{1}{r} V(b) \geq (H - s) + \frac{1}{r} \cdot 0,$$

or  $b \leq V(b)/r$ , where  $1/r$  is the present value of \$1 received next period and every period thereafter.<sup>8</sup> The optimal relational contract therefore sets  $b$  to maximize expected profit per period,  $V(b)$ , subject to the reneging constraint  $b \leq V(b)/r$ .

This model yields intuitive comparative statics. For example, for low interest rates, the Principal is patient and the prospect of future profits makes it willing to pay even a large bonus this period. Thus, for low interest rates, first-best incentives can be created through the choice of an appropriately high value of the bonus  $b$  (namely,  $b = H - L$ ). But for moderate interest rates, the present value of the expected profit from creating first-best incentives is not sufficiently high to make the Principal willing to pay the high bonus necessary to create first-best incentives. Thus, for moderate interest rates, the Principal can offer only a moderate bonus: the Agent's incentive falls, and so expected profits fall, but the present value of future profits remains large enough to induce the Principal to pay this moderate bonus when the Agent produces a High contribution. Finally, for

---

<sup>8</sup> The reneging constraint in this simple model illustrates a much more general result: in a broad class of models, a relational contract is an equilibrium of the repeated game if the maximum total reneging temptation is less than the present value of the total surplus from the contract. (Note that which player is tempted to renege and who gets how much of the total surplus do not appear in this statement.) See equations (3) and (6) in Section IV.B for richer illustrations of this result and MacLeod and Malcomson (1989, 1998) and Levin (2000) for more general statements.

high interest rates, the present value of future profits is so small that the Principal is not willing to pay even a moderate bonus. In fact, there may be no value of the bonus low enough that it is worthwhile for the Principal to pay it, because a small bonus creates only a small incentive for the Agent and hence only a small present value of future profits for the Principal. Thus, for high interest rates, the bonus must be either small or nonexistent.

Some firms use formal and relational incentive contracts in combination. For example, recall that Lincoln Electric uses both piece rates and subjective bonuses. Baker, Gibbons, and Murphy (1994) explore the simultaneous use of formal contracts based on objective performance measures ( $p$ ) and relational contracts based on subjective assessments of total contribution ( $y$ ). The relational contract can reduce the distortionary incentives that would be created by the formal contract on its own, while the formal contract can reduce the size of the relational-contract bonus that the Principal would need to offer if it used only a relational contract. Thus, the combination of formal and relational contracts can reduce distortion in the Agent's incentives and reduce the Principal's temptation to renege on a promised bonus.

Many well-known relational contracts have come under substantial stress (and sometimes failed) when the world has changed important parameters, such as the expected profit for the firm. For example, for several decades IBM made a "no layoffs" pledge to its employees. This was not a formal contract, enforceable by a court, but it was part of "the deal" at IBM: a shared understanding between the firm and its employees about how employment would proceed. As innovations in personal computers and workstations reduced the demand for mainframe computers, however, one could imagine that the value to IBM of living up to this pledge fell. Eventually, IBM abandoned the policy. A similar story can be told about bonus payments to investment bankers at First Boston in the early 1990s (Stewart, 1993). Many bankers left the firm after a second consecutive year in which bankers claimed that bonuses were unexpectedly low. The new parent company, Credit Suisse, claimed that bonuses were low because performance was low (and First Boston had indeed performed worse than its competitors), but many bankers argued that they should be paid the same as bankers at other firms. There may have been a legitimate misunderstanding between the Swiss's view (payment for results) and the bankers' (match the market). Alternatively, the Swiss may have claimed such a misunderstanding as a way to cover their decision to lower bonuses after the junk-bond market collapsed, which caused bankers specializing in mergers and acquisitions become less valuable, reducing the present value of future profits to the Swiss from their relationship with the bankers.

In this model a relational contract is a repeated-game equilibrium. Designing and managing such a contract involves determining a value of  $b$ , communicating to the Agent that  $b$  will be paid if  $y = H$ , assessing whether  $y = H$ , and deciding whether to pay  $b$  if  $y = H$ . In the model, these tasks are trivial; in reality, each requires judgment and knowledge of the specifics of complex situations. Furthermore, hard as these tasks are in practice, much evidence suggests that changing a relational contract is harder still: managers must end one relational contract but preserve enough credibility to begin another, and the new contract they seek to begin often looks suspiciously like renegeing on the old contract they seek to end! See Baker, Gibbons, and Murphy (2001a) for more on these issues.

### III. New Directions in Incentive Theory (Beyond Incentive Contracts)

I think it is important that the theory of incentive contracts has progressed beyond its classic model to the new foundations just described. I expect these new foundations to be the bedrock for the emerging literature on incentives between firms. But there have been other innovations in incentive theory that may also prove useful to those analyzing supply transactions. In this section I sketch two of these new directions in incentive theory: investing in capabilities and career concerns. Both of these models illustrate that incentive contracts are not the only source of incentives.

Since we are now moving from the foundations of agency theory towards applications to supply chains, I abandon the abstract terminology of “Agent” and “Principal” for more concrete labels such as “supplier” and “purchaser.”

#### A. *Investing in Capabilities: Paying for Future Performance*

In this sub-section I temporarily shift the focus from incentives for effort to incentives for investment. In particular, I consider a supplier’s investments in purchaser-specific capabilities. In this new setting performance evaluation is trickier than above because the purchaser must now evaluate a supplier’s potential future value-added to the purchaser’s production process, rather than the supplier’s realized contribution to date. Accordingly, all the difficulties of objective performance measurement described above continue to apply, but probably with more force. Relational incentive contracts based on subjective performance assessments may again be attractive, but the shift from realized performance to date to potential future value-added probably makes it more difficult to find a subjective assessment that both the purchaser and the supplier can observe. I therefore explore a third class of models, based on what a labor economist might call “the promise of promotion” rather than on formal or relational incentive contracts.

If a court could verify a supplier’s potential future value-added to the purchaser’s production process then it would be simple to achieve efficient capability investments. This full-contracting world is the case analyzed in Becker’s (1962) seminal work in labor economics on workers’ decisions to invest in “human capital” (*i.e.*, skill). But I have argued that it is often difficult for a court (or other third party) to determine a worker’s/supplier’s realized value-added to the employer’s/purchaser’s production process. And I now suggest that it is even more difficult for a court to determine a worker’s/supplier’s potential future value-added. Accordingly, just as the recent literature on objective performance measurement has rejected the strong but unremarked assumption in the classic agency model that the Agent’s total contribution is simply “output” that can easily be measured, the recent literature on human capital has likewise rejected the assumption that a worker’s skill can be easily measured. I therefore make the analogous assumption in this sub-section: the supplier’s capability (*i.e.*, potential future value-added to the purchaser’s production process) cannot be easily measured.

To be concrete, in this sub-section let  $v$  denote the purchaser's assessment of the supplier's potential value-added to the purchaser's production process, based on the supplier's performance in an unmodeled probationary period. If the supplier spends extra time during the probationary period learning about the purchaser's markets, competitors, technology, culture, and so on then the supplier will have a high potential value-added for the future. More specifically, suppose the supplier's potential value-added is Low ( $v = L$ ) if the supplier does not invest in such purchaser-specific capabilities but is High ( $v = H > L$ ) if the supplier does invest. To make such an investment, the supplier incurs an opportunity cost of  $c$ . Finally, suppose that the investment is efficient: the value of the investment (to the purchaser) exceeds the cost of the investment (to the supplier), or  $H - L > c$ .

If a contract based on the supplier's potential contribution could be enforced, it would be simple to achieve efficient capability investments in this setting. For example, suppose that  $w_H$  will be paid if  $v = H$  but only  $w_L$  will be paid if  $v = L$ . Then the supplier will choose to invest if  $w_H - w_L > c$  and the purchaser will want to induce investment if  $H - L > w_H - w_L$ . Because the investment is efficient ( $H - L > c$ ), it would be simple to find wages that satisfy  $H - L > w_H - w_L > c$ . As for the participation constraints, let  $r$  denote the supplier's best alternative opportunity after the probationary period. Then the supplier requires  $w_H - c > r$  and the purchaser requires  $H - w_H > 0$ . But this investment would not be efficient (this time relative to alternative opportunities) unless  $H - c > r$ , so it would again be simple to satisfy  $H > w_H > r + c$ . In short, if a contract based on the supplier's potential contribution could be enforced, efficient capability investments could be induced by contract, without recourse to devices such as the promise of promotion.

But suppose (for the remainder of this sub-section) that contracts based on the supplier's potential contribution cannot be enforced. Then we need an indirect way to induce the supplier to invest *and* to induce the purchaser to reward such an investment. That is, non-contractible purchaser-specific capabilities create a two-sided incentive problem: the supplier is concerned that the purchaser cannot be trusted to reward investment properly, and the purchaser is concerned that the supplier will not invest unless such rewards are anticipated.<sup>9</sup> Prendergast (1993) shows that, under certain circumstances, the promise of promotion can solve this two-sided incentive problem, as follows.

As before, suppose that the supplier's capabilities are Low ( $L$ ) if the supplier does not invest in purchaser-specific capabilities but High ( $H$ ) if the supplier does invest, and that the supplier bears the opportunity cost  $c$  from such an investment. Now imagine that the purchaser has two products to which the supplier could be assigned. Call these products "Easy" and "Difficult," meaning that the supplier's capabilities are more important in the production process for the Difficult ( $D$ ) product than for the Easy ( $E$ ) product. More specifically, suppose that a supplier with skill  $s$  ( $= L$  or  $H$ ) assigned to product  $p$  ( $= E$  or  $D$ ) has productivity  $v_{sp}$  satisfying: (1) an untrained supplier is more productive when assigned to the Easy product ( $v_{LE} > v_{LD}$ ); (2) a trained supplier is

---

<sup>9</sup> Note that non-contractible general-purpose capabilities do not create a two-sided incentive problem, because with general-purpose capabilities the supplier knows that competition from other purchasers will cause the supplier's investments in capabilities to be rewarded.

more productive when assigned to the Difficult product ( $v_{HD} > v_{HE}$ ); and (3) training is efficient, because the productivity difference between a trained supplier assigned to the Difficult product and an untrained supplier assigned to the Easy product exceeds the opportunity cost of training ( $v_{HD} - v_{LE} > c$ ).

Suppose that the purchaser can commit to paying a high fee ( $w_D$ ) to the supplier if the supplier is assigned to the Difficult product and a low fee ( $w_E < w_D$ ) if the supplier is assigned to the Easy product. (In a richer model this fee might be the price per unit supplied, but in this subsection we will take the quantity supplied to be fixed.) If the supplier believes that investing in skills will yield assignment to the Difficult product then the supplier will invest if the difference between the high and low fees exceeds the opportunity cost of training ( $w_D - w_E > c$ ). The purchaser, for its part, will choose to assign a trained supplier to the Difficult product if doing so is more profitable than leaving the trained supplier on the easy product ( $v_{HD} - w_D > v_{HE} - w_E$ ). Unfortunately, these two conditions may be incompatible, even if investment is efficient: the supplier and purchaser incentive conditions require  $v_{HD} - v_{HE} > w_D - w_E > c$ , but the assumption that training is efficient guarantees only that  $v_{HD} - v_{LE} > c$ , not that  $v_{HD} - v_{HE} > c$ .

In short, the promise of promotion creates a tension between needing a large enough difference in fees to induce the supplier to invest but a small enough difference in fees to induce the purchaser to assign the supplier to the Difficult product after the supplier has invested. If the values-added from assigning a trained supplier to the two products' production processes are sufficiently different ( $v_{HD} - v_{HE} > c$ ) then there exists a difference in fees that meets both these constraints. But if the two products in question share the same underlying technology of production (or even just satisfy  $v_{HD} - v_{HE} < c$ ) then the promise of promotion cannot simultaneously provide an incentive for the supplier to invest and an incentive for the purchaser to reward investment.

I turn next to up-or-out rules—contracts specifying that after some fixed probationary period the purchaser must either pay the supplier a high fee ( $w^*$ ) or fire the supplier. Up-or-out rules can induce the supplier to invest in purchaser-specific capabilities under circumstances where the promise of promotion would fail, such as where the two products in question share the same underlying technology of production. On the other hand, up-or-out rules have other drawbacks.

Kahn and Huberman (1988) develop a model in which an up-or-out rule solves the two-sided incentive problem created by non-contractible specific investments. If the supplier anticipates a promotion for investing, the supplier will invest if the high wage ( $w^*$ ) minus the opportunity cost exceeds the best alternative ( $w^* - c > r$ ). The purchaser, for its part, will promote a supplier who invested if the supplier's productivity exceeds the high wage ( $H - w^* > 0$ ). Assuming that the investment is efficient ( $H - c > r$ ), in this simple model there is always a high wage that is both high enough to induce the supplier to invest and yet low enough to induce the purchaser to promote a trained supplier ( $H > w^* > r + c$ ).

This example of an up-or-out rule works too neatly. In richer settings, an up-or-out rule may have big costs, even if it solves the two-sided incentive problem discussed here. Suppose, for example, that suppliers who make the appropriate investment could realize any one of several

different levels of capability. If some of the low realizations make the supplier worth less than the high fee attached to the promotion then these suppliers will be fired, even though they undertook the appropriate investment. Firing these suppliers wastes their capabilities. A similar problem arises if there is more than one up-or-out rung in a product ladder—suppliers who clear the first hurdle but fail the second also represent wasted capabilities.

These examples identify a tradeoff between the promise of promotion and an up-or-out rule. The promise of promotion induces the supplier to invest in purchaser-specific capabilities only if the two products are sufficiently different in how they utilize these capabilities, whereas the up-or-out rule can induce investment even if investment yields identical productivity increases for all products. On the other hand, the promise of promotion never wastes the acquired capabilities of those not promoted, whereas an up-or-out rule has this problem in many settings.

Both the promise of promotion and an up-or-out rule illustrate a new kind of contract, different from the formal and relational contracts analyzed in Sections II.A and II.B. In a promise of promotion, for example, the contract links the purchaser's product-assignment decision to the fee that the supplier receives. That is, the contract is based on the Principal's decision rather than on the Agent's outcome, and incentives for the Agent then arise from the Agent's understanding of how the Principal's decision will depend on this outcome. Formally, if the decision is  $d$  and the outcome  $v$  then the contract is  $w(d)$  and incentives arise from  $w(d^*(v))$ .

Obviously, this indirect approach to creating incentives would fail if  $d^*(v) = d_0$  (e.g., if the Principal never promoted the Agent). This is where incentives for effort differ from incentives for investment: if "effort" means that the Agent's total contribution is already realized then the Principal has no incentive to make a decision that depends on the Agent's outcome, whereas if "investment" means that (some of) the Agent's total contribution is still to come then the Principal may have an incentive to tailor the decision to the outcome. Whether the Principal has any (useful) incentive to tailor the decision to the outcome depends on how the Agent's eventual total contribution ( $y$ ) depends on the Principal's decision and the Agent's outcome:  $y(d, v)$ . A promise of promotion works the two products are sufficiently different in how they utilize the supplier's capabilities. More generally, in this indirect approach to creating incentives the Principal chooses  $d$  to maximize  $y(d, v) - w(d)$ , so the resulting  $d^*(v)$  creates useful incentives for the Agent only if  $y(d, v)$  creates appropriate incentives for the Principal.<sup>10</sup>

In this sub-section we have examined indirect incentives in a two-party setting. We turn next to indirect incentives created by competition between (at least) two Principals. Of course, such incentives arise only if both Principals are interested in the Agent's outcome. The simplest case is thus general-purpose performance, rather than the purchaser-specific investments analyzed here.

---

<sup>10</sup> The "implementation" literature in contract theory has asked whether there exists a contract  $w(d)$  that induces the Principal to implement a desired decision rule  $d^*(v)$ , in a world where  $v$  is uncertain but exogenous; see Moore and Repullo (1988), for example. The Prendergast and Kahn-Huberman models described here can be seen as endogenizing  $v$ , by allowing the Agent to choose  $v$  to maximize  $w(d) - c(v)$ . See Aghion, Dewatripont, and Rey (1994) for more abstract progress along the Prendergast-Kahn-Huberman line.

### B. *Career Concerns: Learning without Commitment*

When a worker's current performance affects the market's belief about the worker's ability, and hence the worker's future compensation, we say that the worker has current incentives from "career concerns." In Holmstrom's (1982) model, for example, a worker's total contribution depends not only on the worker's effort but also on the worker's ability. Firms are initially uncertain about the worker's ability and so use observed performance to update their beliefs. Competition among prospective employers makes the worker's future wage depend on firms' updated beliefs about the worker's ability, so the worker has an incentive to perform well to influence these beliefs.

Holmstrom shows that, in the absence of incentive contracts, career concerns produce an inefficient (but intuitive) pattern of effort over time: managers typically work too hard in early years (while the market is still assessing the manager's ability and the manager has a long time to reap the rewards of improving the market's belief) and not hard enough in later years (when the market's belief is hard to budge and the manager has little time to go). Gibbons and Murphy (1992) add incentive contracts (like those in the classic agency model) to the Holmstrom model. Because career-concern incentives decline as the manager approaches retirement, the slope of the optimal incentive contract increases, keeping total incentives (roughly) constant. Gibbons and Murphy present evidence on CEO compensation consistent with this prediction. Chevalier and Ellison (1999) present evidence that mutual fund managers take actions driven by career concerns.

I believe that some suppliers may have career concerns, so in this sub-section I sketch Holmstrom's model. The ingredients in this model are standard: the technology of production, the payoffs, the information structure, and the timing of moves within each period. Again, this model has no incentive contracts whatsoever; adding formal contracts based on an objective performance measure as in Section II.A (*e.g.*,  $w = s + bp$ ) to this basic career-concerns model would be simple and perhaps productive.

*Production:* Consider a T-period model where a supplier's total contribution in period  $t$  ( $y_t$ ) depends on "effort" ( $a_t$ ), noise ( $\epsilon_t$ ), and "ability" ( $\theta$ ). In particular, assume that

$$y_t = \theta + a_t + \epsilon_t,$$

where  $\epsilon_t$  is i.i.d.  $N(0, h)$ .

*Payoffs:* The supplier is risk-neutral but has effort cost  $c(a_t)$  with the usual properties, so the supplier's payoff is

$$\sum_{t=1}^T (w_t - c(a_t)),$$



where  $w_t$  is the “wage” paid to the supplier in period  $t$  and  $\delta$  is the supplier’s discount factor. There are at least two identical purchaser firms. These firms are risk-neutral, with payoff

$$\sum_{t=1}^T \delta^{t-1} (y_t - w_t).$$

For simplicity, the purchaser firms have the same discount factor as the supplier.

*Information:* This model has symmetric uncertainty (as opposed to asymmetric information, such as where the supplier knows its costs but the purchasers do not). The supplier and the purchasers begin with the prior belief that  $\theta$  is  $N(m_0, h_0)$  and independent of the noise terms. As in all agency models, the purchasers cannot observe the supplier’s effort choice, but at the end of each period the supplier and the purchasers all observe the supplier’s total contribution for that period,  $y_t$ . One might think that the supplier’s unobservable effort choice would cause information to become asymmetric over time, but in equilibrium the purchasers can infer the supplier’s effort choice and so the uncertainty remains symmetric throughout the model.

*Timing:* Each period has three events—wage offers by the employers, an effort choice by the supplier, and the realization of the supplier’s total contribution for that period. For simplicity, assume that the purchaser firms are perfectly competitive with each other. (Allowing market power for the purchasers would be straightforward.) As noted above, there are no contracts in this model, so all the purchasers offer a wage equal to the supplier’s expected total contribution for the coming period, conditional on the observed history of total contributions in previous periods:

$$w_t(y_1, \dots, y_{t-1}) = E(y_t | y_1, \dots, y_{t-1}).$$

*Analysis:* To analyze this model, we start with the two-period case. As in most finite-horizon models, we solve this one by working backwards. In the second period, the supplier will choose  $a_2 = 0$ , because the purchasers will already have made their wage offers for that period. Consequently, those wage offers will be

$$w_2(y_1) = E(\theta | y_1).$$

Computing the purchasers’ expectation  $E(\theta | y_1)$  is mostly an exercise for the Normal learning model (DeGroot (1972, chapter 9)), but is complicated by the presence of the supplier’s effort choice  $a_1$  in the signal  $y_1$ . If the purchasers conjectured that the supplier’s first-period effort choice was  $a_1$  then the purchasers would compute

$$E(\theta | y_1) = [h_0 m_0 + h (y_1 - a_1)] / (h_0 + h).$$

Anticipating this computation by the purchasers, the supplier would then choose first-period effort to maximize

$$[w_1 - c(a_1)] + \delta [E\{w_2(y_1) | a_1\} - c(a_2)],$$

where  $w_1$  would already be given and  $a_2$  would be anticipated to be zero. The supplier's optimal first-period effort,  $a_1^*$ , therefore solves

$$c'(a_1) = h / (h_0 + h) .$$

The key feature of this first-order condition for  $a_1^*$  is that it is independent of the purchasers' conjecture  $\hat{a}_1$ . Thus, the only rational conjecture for the purchasers to hold is

$$\hat{a}_1 = a_1^* ,$$

from which it follows that the first-period wage will be

$$w_1 = E(y_1) = m_0 + a_1^* .$$

The equilibrium first-period effort  $a_1^*$  has intuitive comparative statics. First, as the precision of the prior belief about the supplier's ability ( $h_0$ ) increases,  $a_1^*$  falls, because there is less opportunity for the supplier to influence the purchasers' belief about the supplier's ability. Second, as the precision of the noise terms ( $h$ ) increases,  $a_1^*$  increases, because each period's observation of the supplier's total contribution has more effect in updating the purchasers' belief about the supplier's ability. These observations also hold in the equilibrium of the T-period model, where the supplier's optimal effort in period  $t$  satisfies

$$c'(a_t) = \sum_{s=t+1}^T s \cdot h / [h_0 + (s-1)h] .$$

This first-order condition produces Holmstrom's twin conclusions about the time-path of effort: managers typically work too hard in early years (while the market is still assessing the manager's ability and the manager has a long time to reap the rewards of improving the market's belief) and not hard enough in later years (when the market's belief is hard to budge and the manager has little time to go). In particular, it is easy to construct examples where effort is above the first-best level early in the career but below late.

This model of career concerns obviously adds learning to the incentive problem, but the theory also requires a lack of commitment: if wages could be fixed in advance, and if suppliers could commit not to switch to new purchasers offering higher wages, then learning on its own would not create incentive effects. Learning and lack of commitment are also responsible for another kind of incentive effect, which I would be remiss not to at least mention: the ratchet effect.

Lincoln Electric is well-known for its piece rates because it has avoided both the infamous ratchet effect (where the firm reduces the piece rate if it learns that the job can be done more easily than was at first thought) and its repercussion, output restriction (where workers anticipate that the firm will ratchet the rate and so work slowly to prevent the firm from discerning the true pace at which the job could be done). A large body of evidence suggests that Lincoln is a rare exception. Clawson (1980:170) summarizes many case studies and concludes that "employers insisted that they would never cut a price once it was set, yet every employer did cut prices," Mathewson (1931)

offers a huge collection of vignettes concerning output restriction, and Roy (1952) gives detailed evidence of carefully restricted productivity in a machine shop.

To explore these qualitative accounts analytically, it seems natural to consider an environment in which the workers have private information about the job's difficulty and a worker's effort cannot be monitored. In a two-period model of such an environment, Gibbons (1987) shows that if neither the firm nor the worker can commit across periods (*i.e.*, the firm cannot commit in the first period to the second-period piece rate and the worker cannot commit in the first period to remain with the firm for the second period) then both the ratchet effect and output restriction are unavoidable: there is no equilibrium compensation scheme that induces efficient effort.

Kanemoto and MacLeod (1991) and Carmichael and MacLeod (1993) study environments that allow firms to avoid the ratchet effect (and so avoid output restriction as well). Kanemoto and MacLeod analyze a two-period model where the worker's output conveys information about the worker's ability, rather than about the job's difficulty as in Gibbons. In the Kanemoto-MacLeod model, competition for the worker's services from the market of prospective employers gives the worker confidence that the firm will match the market in the second period. Carmichael and MacLeod explore a repeated-game (rather than a two-period) model in which the future value of the ongoing relationship may allow the firm to commit. For example, they describe mutual-monitoring arrangements among British cotton spinners in the 19th century in which piece rates were publicly posted and sanctions were administered (by other firms!) against firms that cut rates.

#### **IV. New Applications to Supply Transactions**

In this section I revisit the new foundations described in Section II but now ask how we know whether a given incentive contract is between firms or within (*i.e.*, between divisions) and whether this distinction matters. This is of course the Coase (1937) / Williamson (1975, 1985, 1996) problem of the boundary of the firm, applied to incentive contracts. A few recent papers address this issue. They show that it does indeed matter whether an incentive contract is within or between firms. The key accomplishment of these models is that for a given supply transaction they derive the optimal incentive contract under both integration and non-integration (that is, within and between firms) and then compare the social surplus produced by each. In this sense, vertical integration becomes an instrument in the incentive problem.

To analyze incentive contracts both within and between firms, I consider an economic environment that includes an upstream party, a downstream party, and an asset. The upstream party uses the asset to produce a good that may be used as an input in the downstream party's production process. Following Grossman and Hart (1986), Hart and Moore (1990), and Hart (1995) [hereafter collectively GHM], asset ownership determines whether a particular transaction is within or between firms. That is, if the upstream party owns the asset then the transaction non-integrated (the upstream

party is an independent contractor, working with an asset she owns), whereas if the downstream party owns the asset then the transaction is integrated (the upstream party is an employee, working with an asset owned by the firm).

The GHM model emphasizes a new source of incentives: hold-up (*i.e.*, post-contractual renegotiation). In particular, in our upstream-downstream model of a supply transaction, ownership of the asset conveys ownership of the good. Therefore, if the upstream party owns the asset then upstream can hold-up downstream by threatening to put the good to its alternative use. Parallel to the idea from Section II.A that an objective performance measure ( $p$ ) could be well or poorly aligned with the Agent's total contribution ( $y$ ), we will see in Section IV.A that hold-up can create useful or distortionary incentives. Section IV.B then reconsiders this issue in the presence of relational contracts within and between firms.

#### A. *Hold-up (May Be Your Friend)*

This section builds on early models of asset ownership and court-enforceable contracts by Holmstrom and Milgrom (1991, 1994), Holmstrom and Tirole (1991), and Holmstrom (1999) [hereafter collectively HMT]. In particular, I recite the static model from Baker, Gibbons, and Murphy (2001b), which blends the HMT approach to court-enforceable contracts with the hold-up incentives that arise in the GHM approach to asset ownership.

In Baker, Gibbons, and Murphy (2001b), we analyze two sources of incentives in spot supply transactions: formal (*i.e.*, solely reliant on court-enforceable contracts) and informal (*i.e.*, involving non-contractual mechanisms such as bargaining or hold-up). To do so, we assume that the transaction includes both contractible and non-contractible components. The non-contractible component of the good has value  $Q$  to the downstream party but also has value  $P$  in the alternative use. The asset is specific, in the sense that  $Q > P$ . Finally,  $Q$  and  $P$  are not contractible but are observable to the upstream and downstream parties and so can influence bargaining between the parties.

To model the contractible component of the good, we assume that the good has characteristics  $X$  that are contractible and so can be the basis of a court-enforceable contract  $w(X)$  even in a one-shot transaction. It would be natural to assume that the good's contractible characteristics have some value to the downstream party, and perhaps also some value in an alternative use. For simplicity, however, we assume that these characteristics have no value in and of themselves, to the downstream party or anyone else. Thus, the parties' only interest in the contractible component of the good is that the actions taken to achieve  $X$  may also affect the likelihoods of  $Q$  and  $P$  (and/or the costs of other actions that affect these likelihoods). In particular, the upstream party chooses a vector of actions  $\mathbf{a}=(a_1, a_2, \dots, a_N)$  that stochastically affect  $Q$ ,  $P$ , and  $X$ . These actions cost the upstream party  $c(\mathbf{a})$ . Given the actions,  $Q$ ,  $P$ , and  $X$  are conditionally

independent. The possible values of  $Q$ ,  $P$ , and  $X$  are finite:  $Q \in \{Q_1, \dots, Q_K\}$ ,  $P \in \{P_1, \dots, P_J\}$ , and  $X \in \{X_1, \dots, X_M\}$ . Our assumption that  $Q > P$  thus requires  $Q_1 > \dots > Q_K > P_1 > \dots > P_J$ .

We assume that ownership of the asset conveys ownership of (both components of) the good produced using the asset. But we also assume that trade in the contractible component is contractible, whereas trade in the non-contractible component is not contractible. For example, if the downstream party owns the asset then he could simply take both components of the good, paying only the contractible fee  $w(X)$  but refusing to pay the upstream party anything further. Alternatively, if the upstream party owns the asset then she could collect the contractible fee  $w(X)$  but deliver only the contractible component, threatening to consign the non-contractible component of the good to its alternative use. Note that we use the same notation  $w(X)$  to mean slightly different contracts in these two cases: if downstream owns the asset then  $w(X)$  is simply an agency contract specifying payment  $w$  for results  $X$ ; if upstream owns the asset then  $w(X)$  is the payment specified if upstream transfers ownership of a contractible component with characteristics  $X$ .

As quick examples of our two-component framework, consider R&D, software development, and consulting. In such settings, an outside expert may be hired to conduct a project for a client. If the expert meets the specifications written in the contract then the expert must be paid. But suppose that in meeting the contract specifications the expert also develops an unanticipated by-product that would be valuable to the client. If this by-product is not covered in the original contract then the outside expert is free not to sell it to the client, and may be free to sell it to other users. If the expert were an internal employee, on the other hand, then the client would own everything produced by the expert, whether covered in the original contract or not.

Our assumption that trade in the non-contractible component is non-contractible is necessary to incorporate *ex post* bargaining in the spirit of GHM. In contrast, consider the familiar quantity-quality distinction, where quality is non-contractible, but quantity as well as trade in the quantity-quality bundle is contractible. The upstream party can surely shirk on quality, but cannot produce high quality and then threaten not to deliver it. Absent this threat, there is no *ex post* bargaining over non-contractible product quality.

In this sub-section we analyze the static version of this model; in Section IV.B we enrich the model to allow relational contracts. In both sub-sections, we call the integrated case “employment,” and the non-integrated case “outsourcing.” Thus, this sub-section analyzes “spot employment” and “spot outsourcing” and the next analyzes “relational employment” and “relational outsourcing.”

Under spot employment the downstream party owns the asset and so can simply take both components of the good while paying the upstream party only the contractible payment  $w(X)$ . Incentives in this case are thus only formal, akin to the original multi-task agency problem analyzed

by Holmstrom and Milgrom (1991). Because the downstream party owns the asset, the upstream party will choose actions  $\mathbf{a}^{SE}(w(\cdot))$  to solve

$$\max_{\mathbf{a}} E_X[w(X(\mathbf{a}))] - c(\mathbf{a}) \quad U^{SE}(w(\cdot)),$$

where  $X(\mathbf{a})$  denotes the random variable  $X$  given  $\mathbf{a}$  (and likewise for  $Q(\mathbf{a})$  and  $P(\mathbf{a})$  below). The actions  $\mathbf{a}^{SE}(w(\cdot))$  yield downstream surplus  $D^{SE}(w(\cdot)) = E_{Q,X}[Q(\mathbf{a}) - w(X(\mathbf{a})) \mid \mathbf{a}^{SE}(w(\cdot))]$ . The parties choose the court-enforceable contract  $w(\cdot)$  to maximize  $U^{SE}(w(\cdot)) + D^{SE}(w(\cdot))$ . Denote the efficient contract by  $w^{SE}(\cdot)$ , the induced actions by  $\mathbf{a}^{SE}$ , and the resulting efficient total surplus by  $S^{SE} = D^{SE} + U^{SE}$ .

Under spot outsourcing, in contrast, the upstream party owns the asset so there is ex post bargaining, similar to non-integration in a GHM model but with the addition of court-enforceable contracts. Incentives in this case are thus both formal and informal. The upstream party owns the asset and so can collect the contractible fee  $w(X)$  but deliver only the contractible component of the good, threatening to consign the non-contractible component to its alternative use. Although upstream and downstream cannot contract on the realized values of  $Q$  and  $P$ , they can negotiate ex post over the price of the non-contractible component. We use the Nash bargaining solution (with bargaining power  $\beta$  for the upstream party) to arrive at this price: downstream will pay upstream the alternative-use value,  $P$ , plus  $0 < \beta < 1$  of the surplus from use by the downstream party,  $Q - P$ , so the price is  $Q + (1 - \beta)P$ .

The upstream party's payoff under spot outsourcing is thus  $w(X)$  plus the bargained price  $Q + (1 - \beta)P$  less the cost of actions  $c(\mathbf{a})$ , so upstream chooses actions  $\mathbf{a}^{SO}(w(\cdot))$  to solve

$$\max_{\mathbf{a}} E_{X,Q,P}[w(X(\mathbf{a})) + \beta(Q(\mathbf{a}) - P(\mathbf{a})) - c(\mathbf{a})] \quad U^{SO}(w(\cdot)).$$

After trade occurs, the downstream party's payoff is  $Q - w - \beta(Q - (1 - \beta)P)$ . Define  $D^{SO}(w(\cdot)) = E_{Q,P,X}[(1 - \beta)\{Q(\mathbf{a}) - P(\mathbf{a})\} - w(X(\mathbf{a})) \mid \mathbf{a}^{SO}(w(\cdot))]$ . As above, the parties choose the court-enforceable contract  $w(\cdot)$  to maximize  $U^{SO}(w(\cdot)) + D^{SO}(w(\cdot))$ . Denote the efficient contract by  $w^{SO}(\cdot)$ , the induced actions by  $\mathbf{a}^{SO}$ , and the resulting efficient total surplus by  $S^{SO} = D^{SO} + U^{SO}$ .

In the spirit of HMT, it is easy to construct examples in which spot employment is more efficient than spot outsourcing ( $S^{SE} > S^{SO}$ ) because integration eliminates market incentives. For instance, imagine that  $X = a_1$ ,  $Q = a_1 + ka_2$ , and  $P = a_3$ , where  $k$  is sufficiently small. Then under non-integration the upstream party will pursue the socially wasteful activity  $a_3$ , whereas under integration there will be contractual incentives for  $a_1$  only. For  $k$  sufficiently small, the omission of incentives for  $a_2$  under integration is preferable to the inclusion of incentives for  $a_3$  under non-integration.

In the spirit of GHM, it is easy to construct examples in which spot outsourcing is more efficient than spot employment ( $S^{SE} < S^{SO}$ ) because the market provides useful informal incentives

via ex post bargaining. For instance, imagine that  $X = a_1$ ,  $Q = a_1 + a_2$ , and  $P = a_2$ . Then under integration there will be contractual incentives for  $a_1$  but no incentives for  $a_2$ , whereas non-integration can achieve the first best: the upstream party will have incentives to pursue both  $a_2$  (anticipating bargaining) and  $a_1$  (because of bargaining and a contract).

More generally, this analysis illustrates that incentive contracts are not the only source of incentives. In particular, GHM have emphasized the incentives created by asset ownership (usually in the absence of incentive contracts), and HMT pushed this perspective into new domains: integration, job design, product design, organizational design, and a host of other decisions that allocate control rights may be at least as important as contracts in structuring incentives within and between firms, and these multiple instruments need to be chosen as a system rather than in isolation.

### B. *Relational Contracts Between or Within Firms*

The foregoing model illustrates how the integration decision can become an instrument in the incentive problem, but the analysis includes only asset ownership and court-enforceable contracts. It is therefore natural to ask what happens if one also includes the other foundational model from Section II, relational contracts. Baker, Gibbons, and Murphy (2001b) make a little progress on combining asset ownership, court-enforceable contracts, and relational contracts, but for ease of exposition in this sub-section I will retreat to an analysis of only asset ownership and relational contracts, omitting court-enforceable contracts. Specifically, I will eliminate the contractible variable  $X$  in the foregoing model, leaving only the non-contractible variables  $Q$  and  $P$  as potential arguments of relational contracts. (In particular, I will refer to the total surpluses  $S^{SE}$  and  $S^{SO}$  as though they had been derived above in the absence of the contractible variable  $X$ .) I also set the bargaining parameter  $\beta = 1/2$ . The resulting model then mimics Baker, Gibbons, and Murphy (2001a).

Consider the relational contract  $(s, b(Q, P))$ , where salary  $s$  is paid by downstream to upstream at the beginning of each period and subjective bonus  $b_{ij} = b(Q_i, P_j)$  is supposed to be paid when  $Q=Q_i$  and  $P=P_j$ . For the moment, suppose that the downstream party will indeed pay  $b_{ij}$  as promised (and that the upstream party will make any promised payments if  $b_{ij} < 0$ ). Then the upstream party will choose a vector of actions  $\mathbf{a}^R(b(\cdot))$  to solve

$$\max_{\mathbf{a}} s + E_{Q,P}[b(Q(\mathbf{a}), P(\mathbf{a}))] - c(\mathbf{a}) \quad U^R(s, b(\cdot)).$$

The actions  $\mathbf{a}^R(b(\cdot))$  yield downstream surplus  $D^R(s, b(\cdot)) = E_{Q,P}[Q(\mathbf{a}) - s - b(Q(\mathbf{a}), P(\mathbf{a})) | \mathbf{a}^R(b(\cdot))]$  and total surplus  $S^R = D^R + U^R$ .

We first examine whether the relational contract  $(s, b(Q, P))$  is self-enforcing when the downstream party owns the asset (relational employment). For concreteness, suppose that  $S^{SO} > S^{SE}$ ; see Baker, Gibbons, and Murphy (2001a) for the parallel analysis when  $S^{SE} > S^{SO}$ .

When the downstream party owns the asset, downstream reneges by refusing to pay the promised bonus  $b_{ij}$  to the upstream party, instead simply taking the good and paying nothing. Because  $S^{SO} > S^{SE}$ , efficiency dictates that after reneging the downstream party sells the asset for some price and then earns  $D^{SO}$  in perpetuity. The downstream party therefore will honor the relational contract as long as the present value of honoring the contract,  $-b_{ij} + D^R/r$ , exceeds the present value of reneging,  $+ D^{SO}/r$ , or

$$(1) \quad b_{ij} \geq -\frac{1}{r} (D^R - D^{SO}).$$

When the downstream party owns the asset, the upstream party reneges on the relational contract by refusing to accept a promised payment  $b_{ij}$  when it is offered (or by refusing to make a promised payment if  $b_{ij} < 0$ ). Because  $S^{SO} > S^{SE}$ , if reneging occurs then the upstream party buys the asset for price and then earns  $U^{SO}$  in perpetuity. Thus, the upstream party will honor rather than renege on the relational contract when

$$(2) \quad b_{ij} \leq -\frac{1}{r} (U^R - U^{SO}).$$

If (1) holds for all  $i$  and  $j$  then it must hold for the largest  $b_{ij}$ , while if (2) holds for all  $i$  and  $j$  then it must hold for the smallest  $b_{ij}$ . Combining these two extreme versions of (1) and (2) yields a necessary condition for the relational contract  $(s, b(Q, P))$  to be self-enforcing under relational employment:

$$(3) \quad \max(b_{ij}) - \min(b_{ij}) \leq \frac{1}{r} (S^R - S^{SO}).$$

In fact, (3) is sufficient as well as necessary, because for any  $\max(b_{ij})$  and  $\min(b_{ij})$  satisfying (3), a fixed payment,  $s$ , can always be chosen that satisfies (1) and (2).

We next examine whether the relational contract  $(s, b(Q, P))$  is self-enforcing when the upstream party owns the asset (relational outsourcing), still under the assumption that  $S^{SO} > S^{SE}$ . We find that the good's value in its alternative use,  $P_j$ , affects the reneging decision under relational outsourcing but not under relational employment. As a result, a given relational contract can be feasible under relational employment but not under relational outsourcing, or the reverse.

When the upstream party owns the asset, the downstream party reneges on the relational contract by negotiating to buy the good for the spot-outsourcing price of  $(Q_i + P_j)/2$  instead of for  $b_{ij}$ , thereby realizing a current payoff of  $Q_i - (Q_i + P_j)/2$ . Because  $S^{SO} > S^{SE}$ , the upstream party retains ownership of the asset after reneging occurs, so the downstream party's discounted future



payoff after reneging is  $D^{SO}/r$ . The downstream party therefore will honor rather than renege on the relational contract when

$$(4) \quad Q_i - b_{ij} + \frac{1}{r} D^R \geq \frac{1}{2} (Q_i - P_j) + \frac{1}{r} D^{SO}, \text{ or}$$

$$b_{ij} - \frac{1}{2} (Q_i + P_j) \leq \frac{1}{r} (D^R - D^{SO}).$$

When the upstream party owns the asset, upstream reneges on the relational contract by negotiating to sell the good for the spot-outsourcing price of  $(Q_i + P_j)/2$  instead of for  $b_{ij}$ . Because  $S^{SO} > S^{SE}$ , the upstream party retains ownership of the asset after reneging occurs, so the upstream party's discounted future payoff after reneging is  $U^{SO}/r$ . The upstream party therefore will honor rather than renege on the relational-outsourcing contract when

$$(5) \quad b_{ij} - \frac{1}{2} (Q_i + P_j) \leq \frac{1}{r} (U^{SO} - U^R).$$

If (4) holds for all  $i$  and  $j$  then it must hold for the largest value of  $b_{ij} - (Q_i + P_j)/2$ , while if (5) holds for all  $i$  and  $j$  it must hold for the smallest value of  $b_{ij} - (Q_i + P_j)/2$ . Combining these two extreme versions of (4) and (5) yields a single necessary condition for the relational contract  $(s, b(Q, P))$  to be self-enforcing under relational outsourcing:

$$(6) \quad \max(b_{ij} - \frac{1}{2} (Q_i + P_j)) - \min(b_{ij} - \frac{1}{2} (Q_i + P_j)) \leq \frac{1}{r} (S^R - S^{SO}).$$

As in the analysis of relational employment, this necessary condition is also sufficient.

We have now established the main proposition from Baker, Gibbons, and Murphy (2001a): asset ownership affects the parties' temptations to renege on a relational contract, and hence affects whether a given relational contract is feasible. Formally, the relational contract  $(s, b(Q, P))$  produces the same actions and surplus regardless of asset ownership, but the maximum total reneging temptation is  $\max(b_{ij}) - \min(b_{ij})$  in (3) under relational employment versus  $\max(b_{ij} - (Q_i + P_j)/2) - \min(b_{ij} - (Q_i + P_j)/2)$  in (6) under relational outsourcing. In some situations, the reneging temptation is lower between integrated parties; in others, the reneging temptation is lower between non-integrated parties. This result motivates an under-explored perspective on vertical integration: a major factor in the vertical-integration decision is whether integration or non-integration facilitates the superior relational contract.

## 8. Conclusion

I hope to have established three things. First, the theory of incentive contracts needed but has received new foundations. The tradeoff between incentives and insurance is an important issue in some settings, but "get what you pay for" problems with objective performance measures and reneging/commitment problems with subjective performance assessments are at least as important,

especially in incentive contracts between firms (where risk-aversion seems less important than for individuals).

Second, new directions in incentive theory teach us that incentive contracts are not the only source of incentives. In the (simplest) career-concerns model, for example, there are no contracts whatsoever, but market competition creates incentives. Similarly, a promise of promotion may involve a contract, but not an incentive contract as analyzed in Section II. Instead, a promise of promotion uses a contract that bases the Agent's reward on the Principal's decision, so as to influence how the Principal's decision responds to the Agent's outcome, and thereby influence the Agent's incentives.

Finally, the integration decision is an instrument in the incentive problem. I think this result has both theoretical and practical implications. Theoretically, we now see that models of incentives between firms are incomplete unless they include something like asset ownership. Such incompleteness may not always be a fatal flaw, but there is the possibility that an elegant analysis of a transaction that is asserted to be between firms could be irrelevant because that transaction would always occur within rather than between firms. And practically, this result continues the theme that incentive contracts are not the only source of incentives: integration and a host of other decisions that allocate control rights may be at least as important as contracts in structuring incentives within and between firms.

## References

- Aghion, Philippe, Mathias Dewatripont and Patrick Rey. 1994. "Renegotiation Design with Unverifiable Information." *Econometrica* 62: 257-282.
- Asch, Beth. 1990. "Do Incentives Matter? The Case of Navy Recruiters." *Industrial and Labor Relations Review* 43:89-106.
- Baker, George. 2000. "Distortion and Risk in Optimal Incentive Contracts." Unpublished manuscript, Harvard Business School.
- \_\_\_\_\_, Robert Gibbons, and Kevin J. Murphy. 1994. "Subjective Performance Measures in Optimal Incentive Contracts." *Quarterly Journal of Economics* 109:1125-56.
- \_\_\_\_\_, \_\_\_\_\_, and \_\_\_\_\_. 2001a. "Relational Contracts and the Theory of the Firm." Forthcoming in *Quarterly Journal of Economics*.
- \_\_\_\_\_, \_\_\_\_\_, and \_\_\_\_\_. 2001b. "Bringing the Market Inside the Firm?" Forthcoming in *American Economic Review*.
- Banker, Rajiv, and Srikant Datar. 1989. "Sensitivity, Precision, and Linear Aggregation of Signals for Performance Evaluation." *Journal of Accounting Research* xx: yy-zz.
- Becker, Gary. 1962. "Investment in Human Capital: A Theoretical Analysis." *Journal of Political Economy* 70:9-49.
- Brown, Keith, W. Harlow, and Laura Starks. 1996. "Of Tournaments and Temptations: An Analysis of Managerial Incentives in the Mutual Fund Industry." *Journal of Finance* 51:85-110.
- Bull, Clive. 1987. "The Existence of Self-Enforcing Relational contracts," *Quarterly Journal of Economics* 102:147-59.
- Chevalier, Judith, and Glen Ellison. 1997. "Risk Taking by Mutual Funds as a Response to Incentives." *Journal of Political Economy* 105:1167-1200.
- \_\_\_\_\_, and \_\_\_\_\_. 1999. "Career Concerns of Mutual Fund Managers." *Quarterly Journal of Economics* 114: 389-432.
- Coase, Ronald. 1937. "The Nature of the Firm." *Economica* 4:386-405.
- Compte, Olivier. 1998. "Communication in Repeated Games with Imperfect Private Monitoring." *Econometrica* 66:597-626.
- Eccles, Robert and Dwight Crane. 1988. *Doing Deals: Investment Banks at Work*. Boston: Harvard Business School Press.
- Fast, Norman and Norman Berg. 1975. "The Lincoln Electric Company." Harvard Business School Case #376-028.

- Feltham, Gerald, and Jim Xie. 1994. "Performance Measure Congruity and Diversity in Multi-Task Principal/Agent Relations." *The Accounting Review* xx: yy-zz.
- Gibbons, Robert. 1987. "Piece-Rate Incentive Schemes." *Journal of Labor Economics* 5:413-29.
- \_\_\_\_\_. 1997. "Incentives and Careers in Organizations." Chapter 1 in Volume II of D. Kreps and K. Wallis (eds.), *Advances in economics and econometrics: theory and applications*, Cambridge University Press.
- \_\_\_\_\_. 1998. "Incentives in Organizations." *Journal of Economic Perspectives* 12: 115-32.
- \_\_\_\_\_ and Kevin J. Murphy. 1992. "Optimal Incentive Contracts in the Presence of Career Concerns: Theory and Evidence." *Journal of Political Economy* 100:468-505.
- Grossman, Sanford and Oliver Hart. 1986. "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Ownership." *Journal of Political Economy* 94:481-510.
- Hart, Oliver. 1995. Firms, Contracts, and Financial Structure. Oxford: Clarendon Press.
- \_\_\_\_\_ and John Moore. 1990. "Property Rights and the Nature of the Firm." *Journal of Political Economy* 98: 1119-58.
- Hayes, Rachel and Scott Schaefer. 2000. "Implicit contracts and the Explanatory Power of Top Executive Compensation for Future Performance." *RAND Journal of Economics* 31: 273-93.
- Healy, Paul. 1985. "The Effect of Bonus Schemes on Accounting Decisions," *Journal of Accounting and Economics* 7:85-107.
- Holmstrom, Bengt. 1979. "Moral Hazard and Observability." *Bell Journal of Economics* 10:74-91.
- \_\_\_\_\_. 1982. "Managerial Incentive Problems—A Dynamic Perspective." In *Essays in Economics and Management in Honor of Lars Wahlbeck*. Helsinki: Swedish School of Economics. Republished in *Review of Economic Studies* 66 (1999): 169-82.
- \_\_\_\_\_. 1999. "The Firm as a Subeconomy." *Journal of Law, Economics, and Organization* 15: 74-102.
- \_\_\_\_\_ and Paul Milgrom. 1987. "Aggregation and Linearity in the Provision of Intertemporal Incentives." *Econometrica*, 55:303-328.
- \_\_\_\_\_ and \_\_\_\_\_. 1991. "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design." *Journal of Law, Economics, and Organization* 7:24-52.
- \_\_\_\_\_ and \_\_\_\_\_. 1994. "The Firm as an Incentive System." *American Economic Review* 84: 972-91.
- \_\_\_\_\_ and Jean Tirole. 1991. "Transfer Pricing and Organizational Form." *Journal of Law, Economics, and Organization* 7:201-28.
- Kahn, Charles and Gur Huberman. 1988. "Two-sided Uncertainty and 'Up-or-Out' Contracts." *Journal of Labor Economics*. 6:423-44.

- Kandori, Michihiro and Hitoshi Matsushima. 1998. "Private Observation, Communication and Collusion." *Econometrica* 66:627-52.
- Kanemoto, Yoshitsugu and Bentley MacLeod. 1991. "The Ratchet Effect and the Market for Secondhand Workers." *Journal of Labor Economics* 10:85-98.
- Kerr, Steven. 1975. "On the Folly of Rewarding A, While Hoping for B." *Academy of Management Journal* 18:769-83.
- Kulp, Susan, Srikant Datar, and Richard Lambert. 1999. "Balancing Performance Measures." Unpublished manuscript, Stanford University.
- Laffont, Jean-Jacques and Jean Tirole. 1986. "Using Cost Observations to Regulate Firms." *Journal of Political Economy* 94: 614-41.
- Lazear, Edward. 2000. "Performance Pay and Productivity." *American Economic Review* 90: 1346-61.
- Levin, Jonathan. 2000. "Relational Incentive Contracts." Unpublished manuscript, Stanford University.
- MacLeod, Bentley and James Malcomson. 1989. "Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment." *Econometrica* 57: 447-80.
- \_\_\_\_\_. and \_\_\_\_\_. 1998. "Motivation and Markets." *American Economic Review* 88:388-411.
- Mathewson, Stanley. [1931] 1969. *Restriction of Output Among Unorganized Workers*. Carbondale, IL: Southern Illinois University Press.
- McMillan, John. 1990. "Managing Workers: Incentive Systems in Japanese and U.S. Industry." *California Management Review* 32:38-55.
- Mirrlees, James. 1974. "Notes on Welfare Economics, Information, and Uncertainty." In M. Balch, D. McFadden, and S. Wu (eds.) *Essays on Economic Behavior Under Uncertainty*, Amsterdam: North-Holland.
- Moore, John and Rafael Repullo. 1988. "Subgame-Perfect Implementation." *Econometrica* 56: 1191-1220.
- Oyer, Paul. 1998. "Fiscal Year Ends and Nonlinear Incentive Contracts: The Effect on Business Seasonality." *Quarterly Journal of Economics* 113:149-85.
- Prendergast, Canice. 1993. "The Role of Promotion in Inducing Specific Human Capital Acquisition." *Quarterly Journal of Economics* 108:523-34.
- \_\_\_\_\_. 1999. "The Provision of Incentives in Firms." Forthcoming in *Journal of Economic Literature* 37: 7-63.
- Roy, Donald. 1952. "Quota Restriction and Goldbricking in a Machine Shop," *American Journal of Sociology* 57:427-42.
- Shapiro, Carl and Joseph Stiglitz. 1984. "Equilibrium Unemployment as a Discipline Device." *American Economic Review* 74:433-44.

Stewart, James. 1993. "Taking The Dare." *The New Yorker*, July 26, pp. 34-39.

Williamson, Oliver. 1975. *Markets and Hierarchies*. New York: The Free Press.

\_\_\_\_\_. 1985. *The Economic Institutions of Capitalism*. New York: The Free Press.

\_\_\_\_\_. 1996. *The Mechanisms of Governance*. New York: Oxford University Press.