The Provision of Incentives in Firms

CANICE PRENDERGAST[1]

1. *Introduction*

INCENTIVES ARE the essence of economics. Despite many wide-ranging claims about their supposed importance, there has been little empirical assessment of incentive provision for workers. The purpose of this paper is to critically overview existing work on the provision of incentives. Since the interests of workers and their employers are not always aligned, a large theoretical literature has emphasized how firms design compensation contracts to induce employees to operate in the firm's interest. This literature has reached into many areas of compensation and has pointed to a multitude of different mechanisms that can be used to induce workers to act in the interests of their employers. These include piece rates, options, discretionary bonuses, promotions, profit sharing, efficiency wages, deferred compensation, and so on. My objective here is to evaluate this literature in the light of a growing empirical literature on compensation. Where possible, I will address the

literature from two perspectives. First, an underlying assumption of this literature is that individuals respond to contracts that reward performance. Accordingly, I consider whether agents behave in this way, and whether these responses are always in the firm's interest. Second, I address whether firms write contracts with these responses in mind. In other words, do contracts look like the predictions of the theory?

Incentives are provided to workers through the compensation practices of firms, encompassing monitoring, evaluation, and contracting, and firms use many different mechanisms to align interests. Some workers, such as salesforce employees, are predominantly rewarded for their efforts through explicit contracts that relate pay to observed measures of performance. Others are rewarded not on individual measures of performance but on more aggregate measures, such as profit-sharing arrangements. However, many employers eschew the use of explicit contracts, preferring to reward individuals based on a discretionary subjective measure of performance. Finally, some employers prefer to avoid pay-for-performance altogether. The objective here is to understand these diverse means of compensation and their implications for the provision of incentives for agents to exert effort and allocate their time in the appropriate way. Of course, the answer to how contracts affect performance is

not going to be universal; instead, the purpose will be to illustrate situations where it appears to improve incentives but also to point out the pitfalls of such a reliance on contractual outcomes. Such prescriptions will be tempered by the nature of the job carried out by workers, the extent to which they have discretion in their jobs, and the extent to which the measures used to pay workers truly reflect the inputs of effort.

The paper is organized along the central themes of the literature. Section 2 considers static contracts, in which incentives are offered in a single-shot setting. I begin in Section 2.1 by setting up the basic theoretical apparatus that will be used throughout the paper. A single model is provided at this stage which can encompass the main themes of the literature, though I initially address the trade-off of risk and incentives. Here the provision of incentives is aided by the use of pay-for-performance, but the primary constraint on incentives is that their provision imposes additional risk on workers, which is costly to firms through higher wages. From this perspective, pay-for-performance is constrained by the noisiness of the measures used to reward agents, and the ability of agents to handle risk.

There is a substantial empirical literature testing the trade-off between risk and incentives. The premise of this literature is that relating pay to performance increases output, but at the cost of imposing risk on workers, which is reflected in higher wages. Two basic themes have been taken. First, a series of papers considers "Do Incentives Matter?"; in other words, do employees perform better when placed on compensation schemes where pay is more closely related to performance? Recent evidence suggests that there are strong responses of output to the use of pay-for-performance contracts. The second

approach, which takes the answer to the first question as given, is to identify whether observed contracts vary in the way that the theories suggest they should. For instance, if risk is a constraint to offering incentives, does the strength of the relationship between pay and performance fall as the measures on which contracts are conditioned become more noisy? This is a truer test of recent contributions to agency theory, which largely hold that contracts are designed with the responses of agents in mind. Here the evidence is more mixed, with some work finding evidence in favor of the theories, while others find little.

An alternative reason why it may be difficult to provide incentives is that contracts cannot completely specify all relevant aspects of worker behavior. As a result, contracts offering incentives can give rise to *dysfunctional behavioral responses*, where agents emphasize only those aspects of performance that are rewarded. For example, consider a baseball player who receives a contract with a reward for hitting home runs. The danger here is that the player will attempt to hit home runs even in situations where it is not warranted. Or teachers who are rewarded on test scores may teach "for the test." Such behavioral responses arise because contracts often cannot rely upon a holistic measure of the worker's contribution at every moment in time. Because of this, agents can "game" the compensation system when they have multiple instruments at their control. Following Bengt Holmstrom and Paul Milgrom (1990) and George Baker (1992), this incentive problem has become known as multitasking, where compensation on any subset of tasks will result in a reallocation of activities toward those that are directly compensated and away from the uncompensated activities. Recent

empirical work has illustrated such behavioral responses to incentive contracts. As a result of these concerns, it is predicted that in those positions where there are significant opportunities for reallocation of activities, there will be an absence of pay-for-performance; in essence, complex jobs will typically not be evaluated through explicit contracts.

Since it is difficult to specify all aspects of workers' jobs in an explicit contract, a common way of providing incentives is to use *subjective performance evaluation*, perhaps in addition to some objective assessments. The typical worker operates in a setting where efforts are exerted in the hope of a promotion, salary revision, or bonus, which are typically at the discretion of superiors. Such subjective assessments have the benefit that they can be a more fully rounded measure of performance; for instance, the baseball player could be rewarded for hitting a home run only if attempting to do so was warranted at the time. However, there is considerable evidence that subjective assessments also give rise to biases. For instance, when evaluations are subjective, workers are likely to waste valuable resources (work time, for example) currying favor with their bosses. In addition to the incentive to engage in such activities, a number of other problems have been highlighted, ranging from "leniency bias," where supervisors are reluctant to give bad ratings to workers, and "centrality bias," where supervisors compress ratings around some norm rather than truly distinguishing good from bad performance. Section 2.2 considers the provision of incentives as a trade-off of the distortions implicit in evaluation; explicit contracts result in agents optimizing relative to the contract ("you get what you pay for"), while subjectively assessed benefits may be tainted by supervisor bias or workers currying favor.

Perhaps the most common means of rewarding white-collar workers for effort is by promotion. Workers are primarily allocated to positions in firms on the basis of their talents. However, an implication of job allocation and promotion is that a job hierarchy provides incentives to workers. Section 2.3 assesses the contribution of *tournament theory*, where promotions are modeled as a tournament among a group of agents competing for a fixed set of prizes, rather like a sports setting. Tournament theory makes a number of predictions, which have found support in empirical work. First, larger prizes should result in more effort. Second, in a contest where there is a single winner, the prize should be increasing in the number of contestants; finally, workers who fall behind in contests should be likely to take risky strategies to "catch up." An additional implication is that when dysfunctional behavioral responses arise, firms may use bureaucratic rules to allocate promotions. For instance, many firms heavily weigh seniority in promotion decisions after controlling for productivity, where sometimes the "wrong" worker is promoted on seniority grounds. This section shows how such bureaucratic rules are the optimal response to incentives for workers to engage in dysfunctional responses to evaluations procedures.

Many workers are employed in settings where output measures are not the outcome of the inputs of a single individual, but rather derive from the joint contributions of many individuals. Such *team production* concerns are the subject of Section 2.4, where the focus is on two incentive aspects of team production. First, there is considerable evidence of free riding in teams. Despite this, the data suggest productivity

improvements of the order of 4 to 5 percent from the introduction of company-wide profit-sharing schemes, where the benefits of increased effort are shared with often thousands of others. Since this constitutes an apparent violation of standard agency theory (why exert effort if I gain only $\frac{1}{1,000}$ of the benefits?), I address this issue in some depth. Second, one of the reasons often suggested for the success of such team compensation schemes is that it gives workers an incentive to monitor one another via peer pressure. Available evidence, though scant, suggests outcomes rather different from those predicted by the theoretical literature.

Efficiency wage theory argues that the provision of incentives causes workers to receive rents above their market wages. In effect, workers are offered rents because they are less likely to shirk if their jobs are valuable; hence high wages induce effort. I provide a brief review of this literature in Section 2.5, and argue that some of the tests used here are of lower power than one would like.

All the effects described above ignore the fact that workers remain with employers and in the labor market for long periods of time. The fact that workers have careers (rather than one-time relationships) allows workers and firms discretion over pay, which results in some different implications from the static models described above. Section 3 considers such dynamic linkages in the provision of incentives, where the optimal contract offered today depends on either the contract offered yesterday or behavior yesterday. Two aspects are emphasized. Section 3.1 considers *deferred compensation*, where firms systematically "overpay" older workers and "underpay" their younger counterparts. Then part of the return to exerting

effort as a younger worker is not just the contemporaneous return but the prospect of receiving the returns of an older worker in the future. Considerable empirical evidence suggests that firms do indeed follow such compensation practices, though there are often other plausible interpretations of these results. This section also considers how the return to promotion changes as a worker ascends a firm's hierarchy.

Another important feature of dynamic agency contracts is that they can be renegotiated over time based on previous performance. Such renegotiation opportunities have been termed *career concerns,* which are the topic of Section 3.2. Once again, consider the baseball player example, but where the player is offered a fixed salary in a given season. Despite the fact that there is no immediate relation between pay and performance, he is likely to have incentives to exert effort because good performance will improve future contracts. In other words, the market "settles up." Such reputational concerns imply that effort exertion can occur without explicit pay-for-performance contracts, though rarely at the efficient level. This career concerns model has testable implications for the behavior of workers and contracts which have been borne out in the small existing literature on this subject.

A final role for repeated relationships in this environment is that they allow for honest behavior in settings where cheating would occur in a static relationship. The ability of workers to punish firms that renege on their obligations implies that repetition of the relation can imply better outcomes when performance measures are subjectively determined. In addition, repeated observations on the performance of workers can allow more precise inferences on their performance, thus

providing another role for repeated relationships. Such dynamic contracting issues are dealt with in Section 3.3.

The literature has provided an important organizing tool for understanding the wage policies of firms, and enormous advances have been made in some areas. Most of this paper is designed to illuminate such advances. However, the paper concludes with three observations in areas where the literature has not been so successful. First, there is considerable evidence that incentives matter; paying individuals to do X causes them to do X. However, it is much less clear that the theories that assume such responses have great predictive power. Second, the literature has suffered from identification difficulties. Such identification concerns arise from two separate sources. First, researchers must overcome the standard empirical identification problem (which realizes that the choice of contracts is endogenous). However, even in situations where there is evidence consistent with agency theory, the literature has been plagued with a second (theoretical) identification problem where outcomes are often equally consistent with other plausible theories.

Third, I believe that there has been insufficient focus on workers whose outputs are hard to observe, in particular those where subjective assessments are used. Instead, the understandable focus of the literature has been on occupations (such as CEOs, mutual fund managers, professional golfers, etc.) for which measures of output are available. However, the majority of workers do not satisfy these criteria. Instead, most workers are evaluated on subjective criteria, where firms choose how to evaluate and how to pay based on those evaluations. Consequently, it seems to me that a critical avenue for future research should be to better understand the evaluation and compensation of those with noncontracted output.

Before beginning the substantive parts of the paper, let me make clear what this overview does not cover. First, as mentioned above, its purpose is not to be a comprehensive overview of the large theoretical literature on compensation. Instead, a theoretical apparatus is provided as a guiding device for understanding the main influences in the theoretical literature. References to the relevant theoretical work are provided for the interested reader. Second, a large literature on compensation considers aspects other than incentive provision, since obviously compensation reflects other considerations, such as human capital acquisition and learning. This literature is not considered here; the emphasis is solely on the provision of incentives. Another important restriction is that I do not address how ownership of assets alleviates incentive and bargaining concerns.[2]

Perhaps the most important omission from the paper is that it concerns only how effort can be induced by incentive contracts. As a result, I do not consider such issues as job enrichment, as in J. Richard Hackman and Greg Oldham (1980). More generally, I do not address in much depth the noneconomic literature on the effect of compensation schemes on performance, though where the predictions of the economic models appear particularly at odds with those of other disciplines, I briefly consider the noneconomic evidence.

Finally, it should be noted that this is not the only overview of the issues covered in this paper; for alternative

---

[2] Following Sanford Grossman and Oliver Hart (1986) and Hart and John Moore (1989), there is now a large literature on the theory of the firm taking the perspective of incomplete contracting, where hold-up or bargaining problems constrain efficiency, which has implications for vertical integration or, more generally, control rights.

perspectives on some similar issues, see Robert Gibbons (1996), Gibbons and Michael Waldman (forthcoming), Edward Lazear (1995), and James Malcomson (forthcoming).

2. *Static Contracts*

The first part of the paper considers the use of static contracts, where the relationship between the worker (agent) and firm (principal) is onetime. I begin by setting out the basic theoretical structure that will be used throughout the paper, though in this section the primary focus is on the trade-off of risk and incentives.

2.1 *The Basic Theory and the Trade-off of Risk and Incentives*

The premise of agency theory is that a principal designs contracts in order to guide appropriate actions by an agent. The agent is assumed to take some action $e \geq 0$, which is unobserved by the principal. Throughout the paper I will refer to e as effort, though there are other plausible interpretations as will become clear below. The agent is effort averse. The purpose of this survey is to identify the major themes of the theoretical literature in a simple way. To do so, I choose a simple parameterization of the agent's utility function, where the agent cares about wages w and effort e; I assume that the agent has exponential utility

$$V = -exp[-r(w-C(e))], \qquad (1)$$

where w is the worker's wage, $r \geq 0$ is the constant rate of absolute risk aversion, and the worker's cost of supplying effort is $C(e)$. Purely for tractability, the cost function is assumed to be quadratic, where $C(e) = \frac{ce^2}{2}$. The principal is assumed to be risk neutral, and the worker has reservation utility $U*$.[3]

Although the principal cannot observe the actions of the agent, she can potentially condition payments on a set of signals that are correlated with the agent's actions. For illustrative purposes, I consider two such signals, an objective measure of performance, y, and a subjective measure of performance, s. Objective measures are characterized by the fact that they can be verified for contractual purposes, while a subjective measure is anything that is not verifiable to a third party. I assume that the principal maximizes expected profits (output minus wage costs), where expected output is given by the sum of the effort of the worker, e, and his ability, α. I assume that the signals y and s are characterized by

$$y = e + \alpha + \varepsilon_y, \qquad (2)$$

and

$$s = e + \alpha + \varepsilon_s, \qquad (3)$$

where $\varepsilon_i \sim N(0,\sigma_i^2)$. Thus, σ_y^2 is the measurement error of the objective signal and σ_s^2 is its counterpart on the subjective signal. Although s is subjectively determined, in this section I assume that contracts can credibly be written on that measure. Problems associated with subjectivity are addressed in more detail below. The term α refers to the ability of the agent and for the moment is symmetrically unknown to all agents (the case where it is privately known is considered below). I assume that $\alpha \sim N(0,\sigma_\alpha^2)$. All random variables are uncorrelated with each other.

Perhaps the most important observation of the early contributions to agency theory (Holmstrom 1979) is what has become known as the Informativeness Principle, which (loosely) implies that any measure of performance that (on the margin) reveals information on the

---

[3] This reservation utility is simply the utility arising from the best outside option available to the worker. This could be a position in another firm or the value of leisure if the agent chooses not to work.

effort level chosen by the agent should be included in the compensation contract. In this context, this implies that both the objective and subjective measures should be used to reward the worker. The attraction of the exponential utility function with normally distributed errors is that, following Holmstrom and Milgrom (1987), the optimal contract relating wages to these observed signals is linear (which makes comparative statics simple to illustrate).[4] Thus, the optimal means of compensation is given by

$$w = \beta_0 + \beta_y y + \beta_s s, \qquad (4)$$

where β_0 is the worker's salary and β_i is the "piece rate" on signal i. (The salary plays little role here, being chosen simply to ensure that the worker earns his reservation utility. As a result, unless relevant it is ignored in what follows.) The worker chooses effort to maximize V, which implies that optimal effort e^* is given by $e^* = \frac{\beta_y + \beta_s}{c}$. Note that the first-best level of effort is $\frac{1}{c}$, which only occurs if $\beta_y + \beta_s = 1$. Optimizing over the choice of compensation contract, it is simple to show that the firm chooses piece rates of

$$\beta_y^* = \frac{\sigma_s^2}{\sigma_s^2 + \sigma_y^2 + rc((\sigma_s^2 + \sigma_y^2)\sigma_\alpha^2 + \sigma_s^2\sigma_y^2)}, \quad (5)$$

and

$$\beta_s^* = \frac{\sigma_y^2}{\sigma_s^2 + \sigma_y^2 + rc((\sigma_s^2 + \sigma_y^2)\sigma_\alpha^2 + \sigma_s^2\sigma_y^2)}. \quad (6)$$

This simple model illustrates the tradeoff between risk and incentives that has been a central early theme of the literature. If the worker is risk neutral ($r = 0$),

then $\beta_y + \beta_s = 1$ so that the first best level of effort is exerted. However, if $r > 0$ and there is measurement error on both performance indicators, effort is always below the first best level ($\beta_y + \beta_s < 1$). A particular measure's weight is decreasing in the variance of the measure's signal, so noisy measures should be less used. However, the weight attached to any measure is increasing in the noisiness of the other measure, though total incentives, $\beta_y^* + \beta_s^*$, are decreasing in the noisiness of any measure. Similarly, a high degree of risk aversion mutes incentives, and incentives rise as the return to effort ($\frac{1}{c}$) rises.

2.1.1 *An Alternative Interpretation*

A useful way of interpreting this result on incentives is that the principal follows a two-step procedure for providing incentives, where the principal separates performance evaluation from the provision of incentives. First, she optimally aggregates information on the worker's effort, assuming a diffuse prior. This estimate, which minimizes the mean-square error, is given by $z = \frac{\sigma_s^2 y + \sigma_y^2 s}{\sigma_s^2 + \sigma_y^2}$. The firm then optimally compensates on this performance measure by discounting the piece rate from 1 for risk sharing reasons, by rewarding via $w = \beta_0 + \beta z$, where

$$\beta = \frac{1}{1 + rc\sigma_z^2}. \qquad (7)$$

where $\sigma_z^2 = \sigma_\alpha^2 + \frac{\sigma_s^2\sigma_y^2}{\sigma_s^2 + \sigma_y^2}$ is the regression error. Note that β tends to 1 as r tends to zero, so piece rates on this optimal aggregation converge to unity as the worker becomes more able to handle risk.

The importance of this interpretation is that in the standard model of incentives, the firm optimally aggregates information on performance and then

---

[4] Holmstrom and Milgrom illustrate this point in a dynamic environment, where agents control efforts continuously where in any period a binary realization of success is achieved. In that setting, compensation that is linear in the score (number of successes) is optimal.

discounts this optimal aggregator for risk sharing reasons. One point that will become clear below is that this rule of optimal aggregation will not be followed when agents take distortionary actions in response to contracts.

2.1.2 *Implications of the Basic Theory*

Relative Performance Evaluation. The most important implication of the analysis above is that errors in measuring performance constrain the provision of incentives. As a result, any signal that is informative about performance should be used in compensation packages (the Informativeness Principle). The most common example of the use of this principle has been application of Relative Performance Evaluation, where the performance of one agent is compared to another when choosing compensation.

Relative performance evaluation is used as a means of filtering out common risk from compensation packages. To give a concrete example, consider two salesforce workers who carry out similar jobs. Demand for the products in the area in which they both work varies for common reasons beyond their control. If agents are compensated solely on their own productivity, they are exposed to the risk inherent in the common fluctuations in demand. A solution to this problem is to (at least partially) reward the workers on how well they do relative to each other; in this way they are not penalized so much for marketwide changes in demand.[5] Rela-

tive performance evaluation has two testable implications. First, in environments where there are common factors affecting compensation, agents should be partially rewarded on how well they do relative to others, and second, the degree of relative performance evaluation should increase in the correlation between the two signals.

The Selection Effects of Contracts. The second outcome of this simple model is that compensation contracts have selection effects, with higher piece rates being relatively more attractive to better workers, as in Lazear (1986). An implication of this is that firms now design contracts not only to induce effort but also to affect the type of workers that they hire. To see this, adapt the basic set-up above by assuming that workers privately know their own ability, α, where for simplicity I assume that the reservation utility of the worker does not depend on α. Let M be the monetary certainty equivalent of the reservation utility U^*. Since the optimal contract will then reflect selection concerns, it will differ from (5) and (6) above. Then for any linear contract w, as in (4), only those workers whose ability exceeds α^* will choose to work for the firm, where

$$\beta_0 + (\beta_y + \beta_s)[\alpha^* + e^*] \\ - \frac{r}{2}(\beta_s^2\sigma_s^2 + \beta_{\tilde{y}}^2\sigma_{\tilde{y}}^2 + \beta_s\beta_y\sigma_s\sigma_y) - C(e^*) = M. \quad (8)$$

By substitution, this implies that

$$\alpha^* = \frac{\frac{r}{2}(\beta_s^2\sigma_s^2 + \beta_{\tilde{y}}^2\sigma_{\tilde{y}}^2 + \beta_s\beta_y\sigma_s\sigma_y) - \beta_0 + M}{\beta_y + \beta_s} \\ - \frac{\beta_y + \beta_s}{2c}. \quad (9)$$

Then it is straightforward to show that the optimal contract for the worker is $w = \beta_0 + \beta_y y - \tilde{\beta}\tilde{y}$, where

$$\beta_y = \frac{1}{1 + rc\sigma_y^2(1 - \rho^2)} \text{ and } \tilde{\beta} = \rho\beta_y.$$

Thus, one worker is penalized when the other does better, all other things equal.

---

[5] To see a role for relative performance evaluation, consider the model above where (i) there are no subjective signals ($\sigma_s^2 = \infty$), (ii) ability is unimportant ($\sigma_\alpha^2 = 0$), but (iii) there is another worker whose performance is correlated with that of the agent. Let that worker's output be given by $\tilde{y} = \tilde{e} + \tilde{\varepsilon}_y$, where the "tilde" refers to the other agent, and the error terms are distributed as above. Assume that there are some common shocks that hit both agents so that $cov(\varepsilon, \tilde{\varepsilon}) = \rho$.

Then any case where a firm shifts from a fixed wage scheme to one where piece rates are used will increase the average quality of worker.

The primary focus of the agency literature has been on how contracts induce certain behaviors from agents. However, compensation contracts also play a central role in recruiting workers to firms.[6] By offering greater pay-for-performance, firms may hire a better distribution of workers, since the more able will benefit more from these contracts than will be the case for the less able. Of course, this changes the design of optimal contracts, since compensation contracts now fulfill a dual role of both inducing effort and aiding the selection of appropriate workers. As a result, firms may distort the effort decisions from the choices in (5) and (6) in order to select certain types of workers.

The Shape of Compensation Contracts. The specification of preferences and measurement errors above is not innocuous. First, effort is one-dimensional. A more general setting would allow the agent to carry out multiple activities, a point that is returned to below. Second, the actions of the agent can affect only the mean of the distribution of output; no actions can affect the higher moments of the distribution. For instance, agents cannot vary the riskiness of the performance measures. Finally, the exponential specification of preferences ignores income effects. The combination of normal errors and absence of income effects yields the linear contracts above. However, the reader should be aware that this is a special case and that in general the sharing rule will not be linear. See Holmstrom (1979) and Grossman and Hart (1983) for details. The attraction of the structure above is therefore its tractability. Yet many observed contracts are nonlinear, where, for example, discrete bonuses are offered for exceeding some performance threshold. For instance, Kevin J. Murphy (1998) highlights the importance of such bonus contracts for executives. Perhaps the most important form of nonlinearity concerns the threat of being fired, where wages vary little with performance but where poor performance is punished by dismissal. Rather remarkably, the theoretical literature has made little progress in understanding the *observed* (nonlinear) shape of compensation contracts, despite costs associated with nonlinearities, which are described below.

2.1.3 *Empirical Tests of the Basic Agency Model*

Empirical work has taken two conceptual approaches. First, a body of work considers "Do Incentives Matter?" In other words, does worker performance improve when pay is more sensitive to performance, and if so, by how much? It should be remembered that this is not a test of the agency theory above. Instead, it tests a necessary ingredient for the theory, namely, that agents respond to incentives, not necessarily that contracts are designed to reflect the trade-off above. Consequently, a second approach considers the contracts offered to agents to identify whether the concerns mentioned above (risk, the return to effort, etc.) are reflected in observed compensation practices. I address each in turn.

---

[6] There are many other aspects of compensation that have selection effects. One mechanism which has been emphasized (George Akerlof, 1976) is where agents reveal their willingness to work by working harder than is efficient. Law partnerships are the typical example used, where associates "burn the midnight oil," largely to appear motivated, though the productivity effect of the final hours is low. Renee Landers, James Rebitzer, and Lowell Taylor (1996) empirically address this issue using survey evidence from law partnerships, which suggests a preference for hours reductions if no inferences are drawn from doing so.

Do Incentives Matter? Until recently, there was remarkably little work in economics documenting the effect of compensation policies on performance.[7] The paucity of such work probably arose from the absence of the necessary information. An adequate test of the effect of pay on performance needs data on contracts offered to workers, measures of performance, and an understanding of why the contracts vary across workers. Despite these constraints, a number of recent papers have illustrated quite substantial incentive and worker selection effects. It is important to bear in mind here in evaluating these studies that in each of the cases documented below, the nature of the job carried out by the workers is "simple," in the sense that an aggregate measure of the worker's performance is easily available.

Lazear (1996) considers the impact of piece rates on the performance of workers who install auto windshields. Management changes provided the impetus for changes in compensation from fixed salaries to piece rates, and Lazear illustrates that productivity rose by approximately 35 percent from this change, with wages increasing by 12 percent. Lazear also uses turnover data to illustrate the selection effects above, where approximately a third of the improved performance can be attributed to selection effects; the less able left the firm and more talented workers replaced them. Similar evidence is presented from a study of Canadian tree planters

by Harry Paarsch and Bruce Shearer (1996). In this case, climatic and soil conditions determine the use of piece rates or salaries. Their data are less extensive than Lazear's, so their estimates are less precisely measured. They carry out a number of useful bounds tests that constrain the effects of pay on productivity. First, wages rise by 6 percent when workers operate under piece rates relative to salaries. This constitutes a lower bound on the effect of pay on productivity; otherwise the firms would prefer to simply retain workers on fixed salaries. A plausible upper bound on the effect of pay on performance is the raw productivity difference, which is 35 percent.[8] The authors use a structural form of estimation to control for contract selection effects and estimate that the incentives from piece rates for a given worker are about 10 percent.[9]

The attraction of these two pieces of work is that both have individual data on performance and contracts. A series of other papers has been more constrained by data limitations, but nonetheless has provided useful information on the effect of compensation policies on performance. First, Rajiv Banker, Seok-Young Lee, and Gordon Potter (1996) consider the effect of piece rates on sales in retail department stores. Data are collected at the store level rather than for individuals, and they show that store productivity rises by between 9 and 14 percent from the change, though the authors cannot distinguish between true incentive effects and worker selection. Sue Fernie and David Metcalf (1996) address the compensation of British jockeys, where some jockeys are employed on fixed

---

[7] At a general level, this section is concerned with understanding the effect of prices on the market for leisure; when the price of on-the-job leisure rises, do agents consume less of it? The premise of this section is that the alternative to exerting effort is laziness, but a little-understood aspect of this literature concerns quite how agents' incentives differ from those of the principal. For recent work pointing out incentives for agents other than to be "lazy," see James Heckman, Jeff Smith, and Chris Taber (1996).

[8] This constitutes an upper bound because piece rates were used in favorable conditions.

[9] See Chris Ferrall and Shearer (1998) for another structural approach to identifying the parameters of the agency problem.

retainers and others are offered prizes for winning races. Their results again suggest significant incentive effects, though their sample size is small. Finally, John McMillan, John Whalley, and Lijing Zhu (1989) and Theodore Groves et al. (1994) address how Chinese economic reforms have affected performance levels through changed compensation practices. McMillan, Whalley, and Zhu estimate a production function for Chinese agriculture using aggregate data, and show that perhaps 75 percent of the increases in agricultural productivity from 1978 to 1984 can be attributed to the introduction of the responsibility system, which allows local communes to retain a share of their profits. Groves et al. use survey data on 800 enterprises in the Chinese industrial sector, where information reported by managers suggests a strong link between industrial performance and the use of bonuses and contract labor.[10]

Larry Kahn and Peter Sherer (1990) use the personnel files of a large company to identify the effects of subjective performance evaluation on the performance of white-collar office workers. They show that better evaluations were achieved by those employees who have a steeper relation between evaluations and pay, once again suggesting the efficiency of relating pay to measures of performance. Finally, each of the papers mentioned above considers the effect of contracts on outputs rather than measuring effort itself. An exception to this is Andrew Foster and Mark Rosenzweig (1994), who collect data on effort exerted by agricultural workers in the Philippines. They do so by examining weight changes for workers on piece rates and salaries (time wages), with the inference being that weight reduction reflects greater effort. They note, first, that conditional on calorie intake, those on piece rates lose more weight than those on fixed salaries. This suggests that the exertion of effort under piece rates causes weight loss. Second, without conditioning on calorie intake, weight gain is higher for those on piece rates than salaries, illustrating that those who operate on piece rates ultimately put on more weight, i.e., surplus created from the use of piece rates.

In summary, this new literature points to considerable effects of compensation on performance. Studies that allowed the effects of incentives to be separated from worker selection issues suggest that perhaps one-third of the increase in performance arises from attracting better workers. It is worth emphasizing two points here. First, in each of the cases considered above, workers carry out "simple" jobs, in the sense that aggregate measures of performance are available; it is for these jobs that piece rates are most likely to work. Second, while it is important to show that incentives matter, these studies are not truly a test of agency theory. They are merely a test of an input to the theory, where a more precise test is to address whether contracts are structured as predicted by the theory. I now turn to this issue.

[10] At a more aggregate level, Louis Putterman (1990) illustrates large increases in the performance of Chinese township and village enterprises when they were allowed to keep larger shares of output. By contrast, Nicholas Barberis, Maxim Boycko, and Andrei Shleifer (1996) show little incentive effects of equity holdings in the Russian retail sector, instead emphasizing human capital aspects of success. See also John Abowd (1990), who uses event study methodology to determine whether the stock market responds favorably to the introduction of sensitive pay-for-performance schemes. He finds evidence in favor of this, though, as he admits, the results are not conclusive. Note, however, that the interpretation of these results depends on why firms introduced more sensitive compensation schemes; only if they were originally "too low" should we expect this response.

Before doing so, however, I should mention a noneconomic literature that holds that offering explicit incentives can reduce productivity by eliminating the *intrinsic* desire to carry out some activity. In other words, pay-for-performance harms incentives, unlike the suggestions above. The premise of this review is that "effort" refers to some activity that the individual would rather avoid. Yet sociologists and psychologists take the perspective that individuals often have pride in their work and enjoy carrying out required tasks. This, of course, is not a problem for the theory above as long as such intrinsic interest is not adversely affected by pay-for-performance. Yet it is sometimes argued that such a link exists, so that paying people on the margin to carry out some activity reduces their intrinsic enjoyment of the task.[11]

While this idea holds some intuitive appeal, it should be noted that there is little conclusive empirical evidence (particularly in workplace settings) of these influences.[12] See Edward Deci

(1971) and Mark Lepper, David Greene, and Robert Nisbett (1973) for example, and Barry Staw (1989) for other interpretations of these findings. Perhaps the most cautious caveat that we can apply to the results above based on these findings is that they may be most plausible for activities where little intrinsic motivation is evident without explicit incentives. See David Kreps (1997) for more observations on this issue.

Do Contracts Reflect Agency Concerns? The theoretical apparatus set up above suggests not only that compensation should change with measures of performance, but also that the size of this relationship depends on such factors as the noisiness of these measures, the marginal return to effort, and the risk tolerance of the agents. Accordingly, a second theme of the literature has concerned identifying the relationship between compensation schemes and proxies of these measures.

Perhaps the most celebrated example of empirically estimating compensation schemes has been a series of papers that estimate pay-for-performance for executives and, particularly, chief executive officers. More specifically, the β coefficients above are estimated for a series of performance measures. Here the typical paper has estimated the relationship between performance (stock price return, earnings, etc.) and some measure of the agent's welfare (pay, propensity to be fired, etc.). See Murphy (1985), Michael Jensen and Murphy (1990), and Stephan Kaplan (1992), for example.[13] Using data for U.S. chief

[11] One version of this is that when an individual performs an act, he must justify the action. If he is not directly paid for the act, he will rationalize his efforts by perceiving that he enjoyed the task. By contrast, if he is rewarded for carrying out the task, this rationalization is no longer necessary and will attribute the reason for doing the task to the monetary rewards, which will lead him to dislike the activity. This dislike could result in worse performance under piece rates.

[12] The methodology typically used in this literature is to consider two groups carrying out some interesting activity. For instance, some experiments have allowed children to draw pictures or play with toys. One group is placed on pay-for-performance while the other is not. Intrinsic motivation is then tested by considering the behavior of the individuals after the supposed period of the experiment is over. If those who are on pay-for-performance are less willing to continue the activity than those who are not on such schemes, it is argued that intrinsic motivation falls from the use of explicit incentives. While this logic may indeed be correct, an alternative which seems plausible is that if those who operate on piece rates perform better during the experiment period, they are simply more tired of carrying out that activity than

those who have operated at a more leisurely pace without pay for performance. Thus they may be less likely to continue the activity for reasons other than intrinsic motivation; instead, diminishing marginal returns to the activity will suffice.

[13] Also see Richard Lambert and David Larker (1987), Ann Coughlin and Ronald Schmidt (1985), and Martin Conyon and Simon Peck (1996) for

executive officers, Jensen and Murphy estimate that a $1,000 increase in the value of a typical large U.S. company increases pay by approximately $3.25, most of this coming from stock ownership. On the basis of this number, the authors argue that too few incentives are provided to executives. While the conclusions taken from this literature could be correct, this seems a poor method of testing agency theory.[14] This is because many of the factors relevant for choosing the *level* of compensation are unobserved; the optimal piece rate depends on risk aversion and the returns to effort, both of which are unknown to the econometrician.[15] As a result, it is difficult to determine whether compensation schemes are set optimally, or to claim that the relationship between pay and performance is too low or too high.[16] It is a little like claiming that prices are too high without knowing costs.

A second approach to understanding the impact of agency theory is not to consider the level of pay-for-performance but to address how coefficients vary with relevant parameters. In other words, are signals used less heavily when they are noisier, or when agents are less able to handle risk? An early attempt to test the predictions of agency theory is Seiichi Kawasaki and McMillan (1987), who are concerned with the relationship between Japanese firms and their subcontractors. Since firms would like subcontractors to constrain costs, they are reluctant to write cost-plus contracts. Instead, a sharing rule is specified, where a fraction of costs can be passed on. On average, about 60 percent of cost overruns are passed on, but this figure varies with the environment. First, subcontractors who face very volatile costs can pass on more costs than those for whom there is little volatility. Second, smaller subcontractors, who are less able to handle risk, can pass on more costs. Finally, the authors use a measure of the marginal product of the effort of the subcontractors (whether they order and negotiate over the prices of their materials) to show that when the marginal product of effort is higher, fewer costs can be passed on. These predictions are supportive of the agency concerns above.[17]

Many of the other studies testing for the trade-off between risk and incentives in contracts concern executive compensation, largely because of data availability.[18] The results here are rather mixed. First, John Garen (1994) finds little evidence that the noisiness of performance measures has any effect on contracts, though Rajesh Aggerwal and Andrew Samwick (1998) find evidence for such a

---

other examples. Another strand of this literature addresses the monitoring power of boards of directors. See Michael Weisbach (1988) and Ben Hermalin and Weisbach (1988) for example.

[14] One interesting conclusion of this paper is that the flow of payments of stock and options in a given year is largely independent of the level that a CEO has inherited entering that year. The fact that the inherited level does not crowd out new issues of stock suggests that there is no common level of desired incentives.

[15] See Jean-Jacques Laffont and Mohamed Matoussi (1995) and Chris Ferrall and Anthony Smith (1997) for a structural approach to identifying these unknowns.

[16] Some simple calculations show that a 1 percent change in the stock price of a large U.S. company changes pay by over a quarter of a million dollars; although $3.25 per $1,000 may seem small, Fortune 500 firms are so large that this translates into large dollar sums. I have no way of evaluating whether this is a large sum of money for a CEO relative to the private benefits they get from "shirking." See Brian Hall and Jeff Leibman (1996).

[17] Banri Asanuma and Tetsuya Kikutani (1992) carry out a similar study on the Japanese auto industry, and also find results supporting the trade-off of risk and incentives.

[18] Though see Lee Alston (1981) and Alston and Robert Higgs (1982) for data on share-cropping contracts, and Charlie Brown (1990) for more aggregate data.

trade-off. Lambert and Larker (1987), Robert Bushman, Raffi Indejikian, and Abbie Smith (1996), and Chris Ittner, David Larker, and Madhav Rajan (1996) test whether the weights placed on objective and subjective measures respond to the noisiness of the objective measures. Straightforward manipulation of (5) and (6) illustrates that the weight placed on subjective measures of performance should increase in the noisiness of the objective measures, while the weight on objective measures obviously falls.[19] Lambert and Larker (1987) and Ittner, Larker, and Rajan (1977) find evidence in favor of this, though the results are rarely resounding. For instance, Ittner, Larker, and Rajan find that the ratio $\frac{\beta_s}{\beta_y}$ is significantly increasing in σ_y^2, but β_s is not.[20] By contrast, Bushman, Indijikian, and Smith (1996) find little effect of variability of objective measures on subjective contracts. Testing across many occupations, Brown (1990) also finds little relation between the existence of piece-rate compensation schemes and the noisiness of those measures.[21]

Available evidence on relative performance evaluation has also focussed on the compensation of executives.[22] First, Richard Antle and Abbie Smith (1986) find weak evidence that the compensation of executives falls as other firms do better, holding own performance fixed, although their data set is small. Using a more comprehensive survey of firms, Gibbons and Murphy (1990) find that executives are indeed penalized when a competitor group fares better, as predicted by the theory. However, somewhat surprisingly, the relevant peer group seems to be the entire stock market rather than companies in the same industry. (One would imagine that there would be more correlation in shocks within the same industry.) Finally, they illustrate that the degree of correlation between the market and the firms (i.e., the extent to which there is a common shock) predicts the use of relative performance evaluation. Murphy (1998) also notes that direct observation of contracts illustrates more extensive use of such evaluation than when inferring contracts as above.[23]

---

[19] Marianne Bertrand and Sendhil Mullanaithan (1997) also consider how various means of incentive provision can act as substitutes for one another by showing that direct contractual incentives for CEOs are increased in situations where takeovers are less likely. In particular, they use state-level variation in takeover laws to show that when states pass legislation that makes hostile takeovers more difficult, firms respond by making their executives more financially liable for the returns of the firm.

[20] It should not be surprising that the authors are more likely to find a stronger relationship from the ratio of the levels than with the level of a single measure. However, the absence of a statistically significant β_s suggests that the size of the effect of noise on incentives is not huge.

[21] One prediction of agency theory which is borne out in the data is that those workers on piece rates will typically earn more than those on fixed wages. Agency theory would predict this as a return to risk (or rents to ability in the case of worker selection). John Pencavel (1977), Trond Peterson (1992), and Daniel Parent (1998) illustrate such differences. Also see Scott Shaefer

(1994) and Rachel Hayes and Shaefer (1997). Shaefer illustrates how pay-for-performance varies with firm size, where larger firms have lower β coefficients due to risk aversion or liquidity constraints. Hayes and Shaefer provide a useful contribution to understanding the effect of subjective performance evaluation by showing that future performance measures (such as earnings) can be predicted by previous discretionary compensation changes to chief executives. Their interpretation of this is that these agents are rewarded for taking the "right" actions even in settings where the immediate objective returns do not arise.

[22] Though see Edward Fee and Charles Hadlock (1997), who note that managers in major newspapers are more likely to be replaced when competitor newspapers increase circulation. This can clearly be interpreted as relative performance evaluation, though since papers are substitutes in the product market, the circulation of other papers may simply be another measure of the poor performance of the newspaper manager.

[23] It should not be assumed that there is universal agreement on the frequency of relative perfor-

To summarize this section, there is some evidence that contracts are designed to optimally trade off risk against incentives. However, the evidence is hardly overwhelming, with some studies showing the effect of noise on piece rates while others show little. Thus, while agents do appear to respond to incentives, it would not appear that on the margin, the risk measures that have been considered are the true constraining factors on the provision of incentives. At one level, of course, risk aversion (or at least liquidity constraints) must play a role, since otherwise agents could be offered piece rates of 100 percent (i.e., sell the firm to the agent), but on the relevant margin, the effect of risk appears to be limited. Instead, perhaps the most striking aspect of observed contracts is that the Informativeness Principle, i.e., that all factors correlated with performance should be included in a compensation contract, seems to be *violated* in many occupations. For instance, there are many measures of the productivity of an academic (such as publications, teaching ratings, etc.) or a baseball player (batting average, home runs, etc.), yet explicit contracts are rarely written on those measures. The reason for this is not because these measures are infinitely risky (as the previous section would require), but rather that contracts can typically be written only on a subset of activities carried out by an agent, and rewarding agents on a subset of all things that they do can cause dysfunctional behavioral responses, to which I now turn.

mance evaluation. See Aggarwal and Samwick (1998); Jason Barro and Robert Barro (1990); David Blackwell, James Brickley, and Michael Weisbach (1994); and Janakiraman, Lambert, and Larker (1992) for empirical work finding little evidence on this incentive device.

2.2 *Other Behavioral Responses to Compensation Schemes*

So far, we have considered only an agent's incentive to exert "effort." However, compensation schemes often have unintended consequences caused by agents changing their activities in other ways that are beneficial to them but not to their employer. Therefore, a potential cost of pay-for-performance schemes is not only that they impose risk on agents, but also that the agents can "game" the evaluation procedure to their advantage. This arises because many jobs are complex, in the sense that many aspects of those jobs are hard to contract over. As a result, the use of explicit contracts could cause agents to focus too much on those aspects of the job included in the contract to the detriment of those that are excluded. A couple of examples should illustrate the nature of the problem. Consider the contract offered to Ken O'Brien, a football quarterback, in the mid-1980s. Early in his career, he had a tendency to throw interceptions. As a result, he received a contract that penalized him every time he threw the ball to a member of the opposition. However, while it was the case that he subsequently threw fewer interceptions, this was largely because he refused to throw the ball, even in cases where he should have done so. As Joe Namath put it, "I see him hold onto the ball more than he should . . . I don't like incentive contracts that pertain to numbers" (quoted in Brown 1990). Or the practice used at AT&T, where computer programmers were rewarded on the number of lines of code that they produced in their programs. Not surprisingly, this resulted in longer programs than was necessary. These examples have the same conceptual feature; agents can change the nature of their activities in response to *objective*

contracts in a way that is privately beneficial to the agent but harmful to his employer. Following Holmstrom and Milgrom (1991), this distortion has become known as multi-tasking.

As a result of the danger of agents overemphasizing objective criteria, it is typically argued that firms should not pay based on objectively measured criteria, but instead should use *subjective performance evaluation*. The attraction of subjectively determined measures of performance is that they allow a more holistic picture of performance to be attained, not possible with objective contracts. For instance, in the AT&T example, a subjective assessment could reward for long programs only in cases where those programs are warranted. As a result, for jobs without clear aggregate measures of performance, rewards tend to be allocated in a discretionary fashion.

Two examples are apposite here. First consider the case of a baseball player. It is difficult to imagine an occupation for which there are more measures of performance. Despite this, it is not common for players to have contracts where pay is directly related to *specific* performance measures. Part of the reason for this is that teams are reluctant to offer a contract that rewards a player for home runs, say, because the player may have an incentive to hit home runs even when it is not in the interest of the team for him to do so. By contrast, the more common cases where players are offered explicit bonuses are for *aggregate* measures of performance, such as making the All Star Team or being the league's Most Valuable Player. Since these are more holistic measures of performance, they suffer less from the multi-tasking dilemma. The second relevant example concerns chief executive officers. No one could claim that their jobs are not complicated; clearly

their jobs are multi-dimensional, and opportunities for reallocation of tasks to increase rewards are certainly possible. Despite this, most incentives for these jobs are provided by explicit incentives (primarily through stock holdings). The reason is that aggregate measures of performance are available through, say, the stock price return, which is relatively exempt from multi-tasking concerns. In situations where executives are assessed on non-holistic measures, I provide evidence below that they also behave in ways that are privately beneficial.

Subjective assessments, however, also induce inefficient behavioral responses. The literature in both economics and more particularly in human resources management has emphasized how incentives provided through subjective assessments cause agents to change their behavior, and cause supervisors to distort their reports, in such a way that efficiency is harmed. The purpose of this section is to address how objective and subjective signals should be used in situations where both potentially induce inefficient responses in behavior. In order to highlight the distinctive features of this section, I restrict attention to the case of risk-neutral agents, so that any effects that arise are due to behavioral responses. The effect of risk aversion in this setting is largely additive, in the sense that higher risk aversion reduces incentives; since there are no interesting interactions between behavioral responses and risk aversion, risk neutrality is assumed.

2.2.1 *Multi-Tasking*

The essence of this section is that at times agents will take actions other than those the principal would like to induce. Since contracts are an imperfect representation of the worker's

contribution to the firm, workers can "game" the compensation scheme to their benefit. Holmstrom and Milgrom (1991) address this issue by assuming that agents carry out multiple activities, and choose the allocation of activities based on offered contracts. Consider a simple setting where an agent chooses between activities *a* and *b*, where the cost of effort for the worker is such that he is indifferent to how they are allocated between tasks (i.e., only the sum of effort matters).[24] Then with the linear contract above, if the principal offers a contract with a higher return to one activity than the other, the agent allocates all his effort to one task, even if the principal would prefer the agent to allocate his time to both tasks. In order to induce the agent to allocate time to both tasks, the same incentives must be offered on both. But there need be no reason why this is otherwise optimal; for instance, the measurement error on one (e.g., quantity produced) may be much lower than for the other (e.g., quality produced). As a result, multi-tasking imposes constraints on the trade-off between risk and incentives.[25]

Rather than model this approach to "gaming," I use a simpler structure that makes comparison to the basic model easier and derives from Baker (1992), where an agent chooses only a single effort decision, but there is a divergence between the privately and socially optimal effort level. For example, an agent who is rewarded on quantity produced may know that his effort is worthless if there is no demand for the goods, though it is in his private interest to produce. To model this, assume that the expected marginal product of the agent is, as above, $e + \alpha$, but the objective measure on which the worker is rewarded, $\tilde{y}$, is given by

$$\tilde{y} = \mu e + \alpha + \varepsilon_y, \qquad (10)$$

where all variables are distributed as in Section 2.1. Throughout the paper a "tilde" over a variable refers to a corrupted version of the appropriate signal. The only difference from the basic setting is through μ. Here the marginal effect of effort on the indicator depends on μ, while true productivity is independent of that measure. Assume that μ is privately known by the agent so that the marginal return to effort on surplus is unity, but the expected private return depends on μ because contracts are written on $\tilde{y}$. Assume that $\mu \sim N(1, \sigma_\mu^2)$, where σ_μ^2 is a direct measure of the extent to which the agent can "game" the compensation system. (In the previous section, $\sigma_\mu^2 = 0$.)[26]

---

[24] An important implication of this theory has been implications for ownership of assets. To take a simple example, consider a worker who is employed on a piece rate. If he is employed by a firm, he may have an incentive not to take due care of his machinery, as his incentives are simply to produce as much as possible over a short period of time. By contrast, an agent who owns his machinery has better incentives to exercise due care. As a result, ownership can solve some multi-tasking problems, and this result is also consistent with the observation that the self-employed have high piece rates, while those employed by firms tend to have compensation that is less sensitive to performance. This observation is based on Holmstrom (1996), and empirical work by Eric Andersen and David Schmittlein (1984) is consistent with this. Recent empirical work on franchising, such as Francine Lafontaine (1992) and Margaret Slade (1996), also emphasizes this approach.

[25] Agents reallocate efforts between activities when efforts on tasks are substitutes or complements; for instance, if time is limited, more time

on one task likely leads to less on others. Such substitution opportunities constrain the ability of the firm to offer piece rate contracts.

[26] This section is not meant to be a detailed description of the theoretical implications of dysfunctional behavioral responses; instead the results are meant to be illustrative. For example, in this hidden action game, the firm could design revelation mechanisms where the agent reveals μ to the firm, rather than restrict attention to linear sharing rules. As in Baker (1992), it is assumed that the firm cannot design such revelation mechanisms.
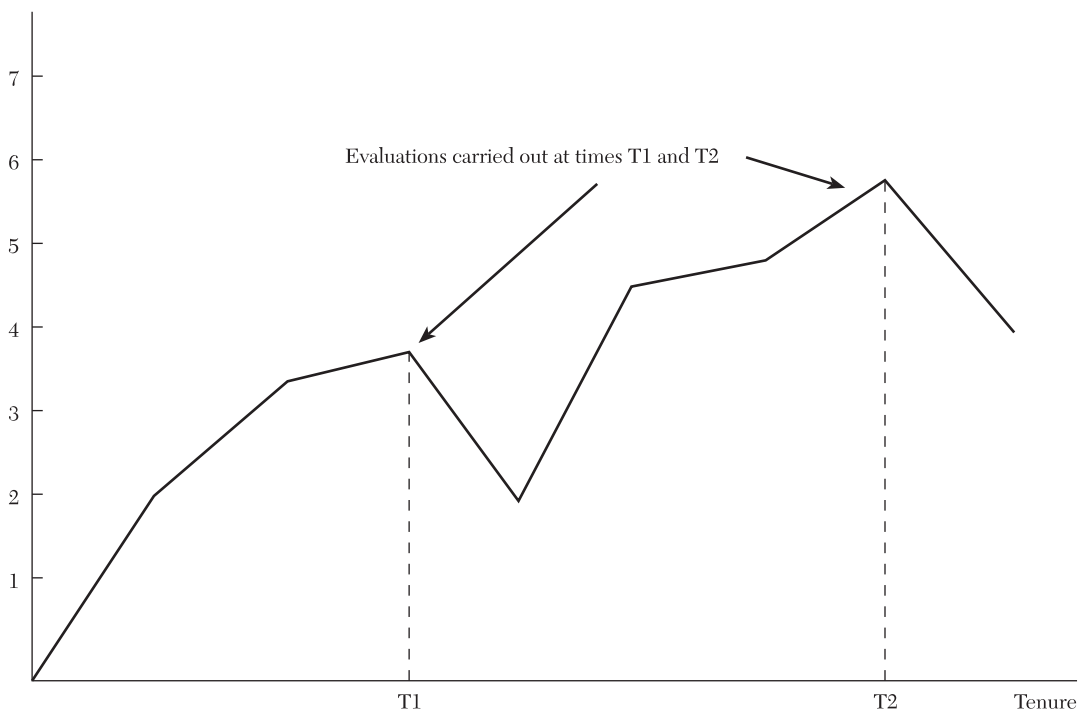
Number of Recruits Per Month

Figure 1. The Behavior of Navy Recruiters (Asch 1990).

The problem with basing compensation on $\tilde{y}$ is that the marginal return to effort depends on μ. One solution is to condition compensation on a subjective measure of performance, which is exempt from this problem. We defer this possibility until the next subsection by assuming that $\sigma_{\tilde{s}}^2 = \infty$, so that only objective measures are used. Compensation is based on a linear signal of the performance measure,

$$w = \beta_0 + \beta_{\tilde{y}}\,\tilde{y}. \qquad (11)$$

The agent optimally chooses effort equal to $e^* = \dfrac{\mu\beta_{\tilde{y}}}{c}$. (If $\sigma_{\mu}^2 = 0$, this becomes $e^* = \dfrac{\beta_{\tilde{y}}}{c}$, as in the previous section.) If $\sigma_{\mu}^2 > 0$ note that the agent bases effort on a measure uncorrelated with social surplus (which is costly as effort costs are

convex). The principal responds to this by muting incentives. Given this incentive, the firm's expected surplus maximization problem is equivalent to maximizing

$$\beta_{\tilde{y}} - \frac{\beta_{\tilde{y}}^2(1 + \sigma_{\mu}^2)}{2}.$$

The optimal piece rate is then trivially given by

$$\beta_{\tilde{y}}^* = \frac{1}{1 + \sigma_{\mu}^2} < 1 \qquad (12)$$

if $\sigma_{\mu}^2 > 0$. Thus, even with risk neutral agents, incentives are below unity in order to constrain inefficient behavioral responses.[27] In other words, firms mute

[27] This simple set-up implies that firms should always offer some incentives to workers. However, in the multi-task setting in Holmstrom and Milgrom it is straightforward to show situations where the firm is better off offering no incentives. Essentially, this requires that agents be willing to supply

incentives when workers can game the system.[28]

2.2.2 *Empirical Evidence on Dysfunctional Responses to Compensation Schemes*

Many of the empirical studies illustrating inefficient behavioral responses to objective compensation schemes arise from the fact that many compensation schemes are nonlinear in performance measures or use aggregates over long time periods, such as a year. Paul Healy (1985), Beth Asch (1990), Paul Oyer (1998), and Pascal Courty and Jerry Marschke (1996) consider various forms of agency behavior under such compensation systems. Consider a situation where an agent gets a discrete prize for reaching a quota by some date, say December 31. The agent must choose effort throughout the year based on the probability of reaching that quota by the end of the year. The social surplus of his efforts is independent of (i) when he carries out the activity and

(ii) how close he is to the quota; this is given by $e + \alpha$ above. However, private and social returns are unlikely to coincide with a quota system of compensation, where agents are rewarded periodically. A number of distortions can arise. First, if individuals are rewarded on aggregates over long time periods, they may have an incentive to wait "until the last minute" simply for discounting reasons. Asch (1990) considers such incentives for Navy recruiters to reallocate effort over time. The recruiters are offered pay-for-performance (based on number of recruits) through the prospect of either speedier promotion or increased probability of further tours of duty. Her primary interest is in identifying how the recruiters allocate their efforts across time in response to this incentive scheme. Figure 1 illustrates the average number of recruits over the evaluation period.

Incentives are provided at the pre-specified dates T_1 and T_2. Prizes vary by the number of recruits, but contract termination dates are specified in advance so that the recruiters know by when they must reach their targets. Figure 1 shows performance gradually increasing until the evaluation date and then discretely falling after the evaluation period. This suggests intertemporal effort reallocation in response to the compensation scheme. Oyer (1998) provides similar evidence on the effect of intermittent rewards based on quotas on the average levels of effort for salesforce workers.[29]

Similar effort reallocation would, of course, occur in settings where there are linear contracts with infrequent evaluation; people simply prefer to wait

---

some effort even when there are no explicit contracts.

[28] Firms must typically determine not only how to pay their employees but also what they let them do, i.e., the must choose task assignments. Holmstrom and Milgrom (1991) use multi-task logic to identify the optimal allocation of workers to safe and risky jobs. In order to reduce the possibility of the worker misallocating activities across tasks, Holmstrom and Milgrom show that the optimal solution to this problem may be to allocate only risky tasks to some agents and only safe tasks to others. Another aspect of job design concerns the amount of authority offered to workers. Philippe Aghion and Jean Tirole (1997) address this by noting that incentive concerns may induce "excessive" delegation of authority to workers. Consider the problem of a firm that would like its subordinates to exert effort, but where rewards from doing so are not through wages, but through doing desirable tasks. Workers may fear that their decisions will be overruled by a superior, so that the agent does not get the returns to exerting effort. A solution to this problem is to give the worker authority over the decisions made, so that he realizes that if he exerts effort, he is likely to see the return from doing so.

[29] Oyer also illustrates that the "threshold" effect of rewards (i.e., that they are nonlinear) can either increase or decrease average effort levels over the evaluation period, depending on the difficulty of attaining the quota.
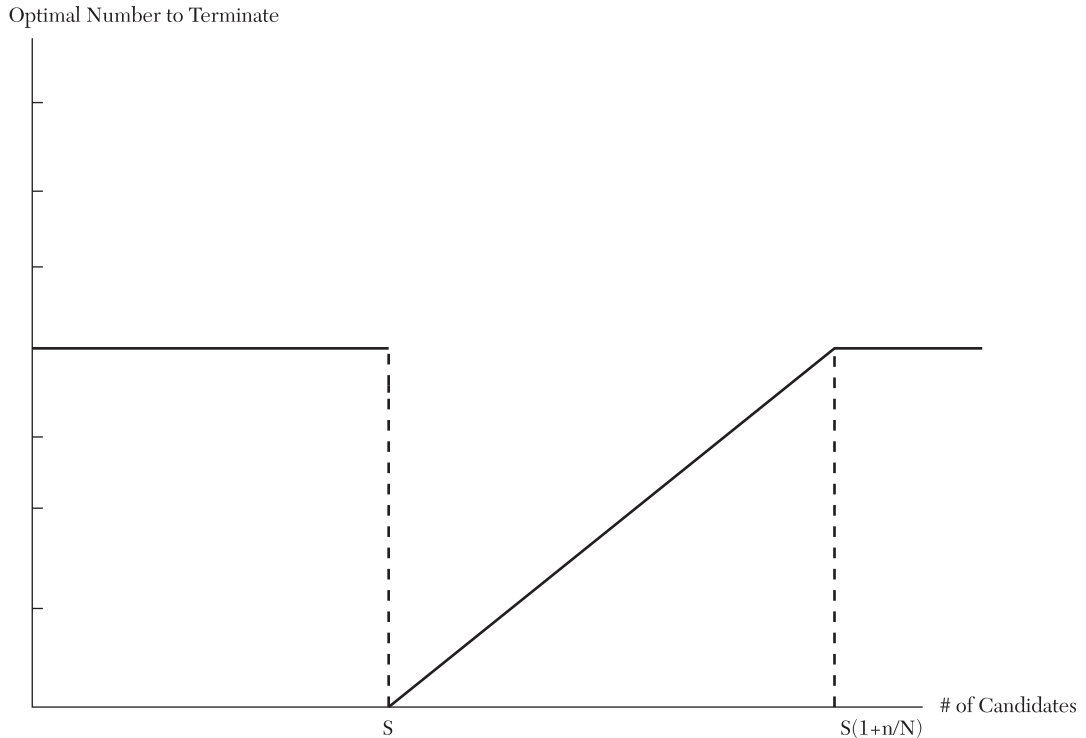
Optimal Number to Terminate

Figure 2. Terminating Trainees (Courty and Marschke 1996)

to exert effort. However, another characteristic of quota systems is that incentives vary by whether the agent is close to the evaluation quota. In particular, an agent who is close to winning the prize will have greater incentive (μ high in the terminology of the previous model) than one who has either exceeded the quota or is unlikely to reach that quota (μ low in both cases). Evidence on this is provided in Healy (1985), Oyer (1997), Courty and Marschke (1997), and Andrew Leventis (1997). Courty and Marschke consider the effect of incentive contracts offered to agencies that provide job training for individuals on welfare. Job training for welfare recipients is typically carried out by private agencies, which are offered incentives for desirable outcomes. In particular, these agencies are offered

a bonus if they attain certain standards by June 1 of each year. For example, the agency could be rewarded if 40 percent of its trainees attain jobs. Critically, the Department of Labor (which administers this system) offers these incentives as a function of "graduated" employees (i.e., those clients who have finished the training program). But the agencies can decide when to graduate them. As a result, the agencies have an incentive to strategically graduate employees. Consider a case where at June 1, the agency must decide how many of n unemployed candidates to graduate where it has already graduated N during the year, and must place a percentage of s in employment in order to achieve a bonus. Figure 2 illustrates the strategic incentives.

   If the agency has not reached its

target or has exceeded $1 + \frac{n}{N}$ of its target, it will graduate all n to maximize returns (assuming that these are less employable than new trainees, as seems reasonable). However, if the agency slightly exceeds its standards, it will be more reluctant to graduate, since graduating all no longer means getting the bonus. Thus graduations close to June 1 will be lower for success rates close to the target, as illustrated in Figure 2. This plots the non-monotonic relationship between privately optimal number of graduations and candidates. Courty and Marschke empirically illustrate that the agencies do exhibit this behavior.

Healy (1985) illustrates that executives exhibit similar strategic behavior in reporting earnings when they are rewarded as a nonlinear function of earnings. Consider compensation schemes that reward for earnings only over some range, where there is a floor below which compensation cannot fall and a ceiling above which it cannot rise. Healy shows that executives who have already reached their ceiling do not report all their earnings. Similarly, they do not report all earnings if they have earnings so low that they are unlikely to reach the region where they earn positive marginal returns. Leventis (1997) considers the response of surgeons to incentives. In New York, surgeons are penalized if their mortality rates exceed a threshold. They respond by taking less risky cases as they approach that threshold. Finally, Oyer (1997) finds empirical evidence on another implication of quota compensation schemes; namely, that there will be greater variability in sales at the end of the financial year than at other times.

Behavioral responses are not specific to quota systems. A number of other authors have directly considered the incentive for agents to reallocate their efforts (or activities) in response to the compensation scheme that they face. Keith Brown, W. Harlow, and Laura Starks, (1996) and Judith Chevalier and Glenn Ellison (1997a) consider the likelihood that agents will change the riskiness of their activities in response to incentives. They consider inefficient risk taking by mutual fund managers.[30] Mutual fund managers are typically rewarded as a linear function of the assets that they control. However, the flow of funds into mutual funds is not a linear function of performance (particularly for younger funds, shown here). Instead, as illustrated in Chevalier and Ellison, the flow of funds is given in Figure 3.

The important aspect of this figure for our purposes is its nonlinearity. Over some ranges, the shape of the returns function is concave, while over others it is convex. These data reflect the propensity for individuals to place funds in mutual funds based on previous performance. In those regions where the payoffs are concave, there is an incentive to take inefficiently few risks, while convexity implies an incentive to take too much risk. For example, for those funds that perform 20 percent worse than the risk adjusted market return, there are incentives to avoid (local) risk, while for those performing 10 to 15 percent better than the market, there is an incentive to increase risk. Chevalier and Ellison estimate the risk taking of mutual fund managers at the end of the year (as measured by the standard deviation of the difference between fund return and the market) as a function of performance until

---

[30] Note that allowing agents to choose the riskiness of their outcomes violates the assumptions of the model above. In the context of risk taking, nonlinear incentives are generally optimal. See Holmstrom and Joan Ricart i Costa (1986) for a model on contracting in the context of risk taking.
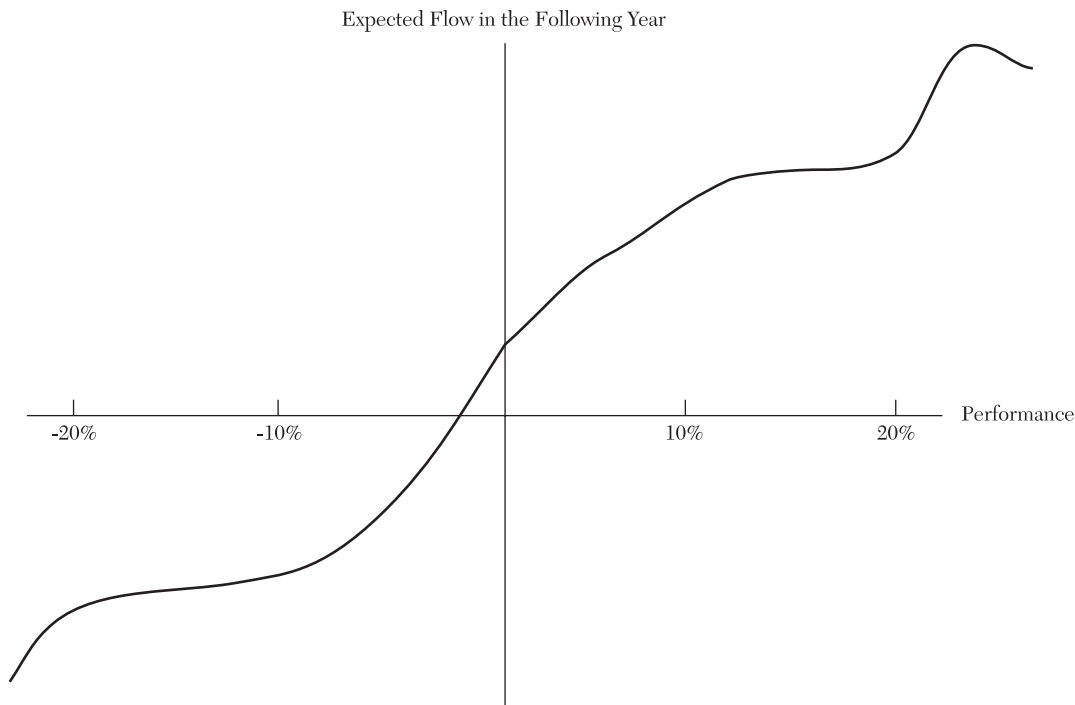
Expected Flow in the Following Year

Figure 3. The Flow of Funds into Mutual Funds (Chevalier and Ellison 1997).

September. For example, those funds that are 20 percent worse than the January–September risk-adjusted market return are predicted to take too little risk in the October–December period, while those that perform 15 percent better have an incentive to take excessive risks. They show that the agents do allocate assets in this way.[31] This risk taking occurs at the cost of lower risk-adjusted returns, suggesting a divergence between social and private returns.

Other work has suggested different dimensions on which agents respond to compensation schemes. For instance, in the context of the job training setting described above, Anderson, Burkhauser, and Raymond (1993), Heckman, Heinrich, and Smith (1997), and Courty and Marschke (1997) show that when the training agencies are rewarded on their success in placing trainees in jobs, they "cream skim," i.e., they recruit only the most qualified candidates rather than the most needy. Marschke (1996) additionally shows that when these agencies are rewarded on certain criteria, they focus more on the types of training that induce these outcomes, though at the cost of other types of desired training. In a sports setting, Brian Becker and Mark Huselid (1992) show that increases in prize money among professional auto drivers result in more risky driving, as witnessed by more caution flags. Finally, Robert Drago and Gerald Garvey (1997) use Australian survey data to illustrate that when

[31] Excessive risk taking by the high performers is statistically significant only using portfolio time series data. When examining the set of funds for which portfolios themselves can be observed, this effect becomes insignificant.

agents are placed on individual pay-for-performance schemes, they are less likely to help their coworkers.

In summary, this work suggests that firms get what they pay for; by emphasizing certain outcomes in pay, they make those outcomes more likely to occur. Under some circumstances, this increases productivity (as the studies in Section 2.1 would suggest), yet in more complex settings, this can result in a reallocation of activities that is not obviously efficient.

2.2.3 *Subjective Performance Evaluation*

Because of such multi-tasking concerns, firms are often reluctant to offer rewards based on objective measures of performance. Instead, they prefer to use subjective measures, where pay is at the discretion of the impressions of a superior. The attraction of such means of payment is that they offer a more holistic view of performance; the agent can be rewarded for a particular activity only if that activity was warranted at the time. For example, in the case of the football quarterback who stopped throwing the ball when confronted with a payment scheme that penalized him for doing so, suitable adjustments could be made with a subjectively determined performance measure. However, the essential feature of subjective assessments is that they cannot be verified by outsiders (or at least it is costly for third parties to determine performance), which gives rise to the possibility that performance measures will be manipulated or distorted from their true values. Such distortions can arise for a number of different reasons, which I now describe.

Theft. One danger of assessments that are subject to manipulation is that a principal will underreport performance in order to save on wages. If a supervisor is also residual claimant on profits, any wages offered to the agent come from his pocket. Thus, even though an agent exerts effort and performs well, the supervisor may claim otherwise to keep costs down.[32] An example where measures may be manipulated in order to reduce costs is the movie industry, where actors are sometimes paid on the "net profits" of a film. As a result, there have been numerous court cases regarding "creative accounting" designed to keep net profits low even for apparently successful films.[33] See Carole Cheatham, Dorothy Davis, and Leo Cheatham (1996) for more details on this. Yet theft is not solely the privilege of principals. Consider the case of a cab driver who leases his cab

[32] The theoretical literature has suggested a number of solutions to the danger that a principal will underreport to save on wage costs. First, Clive Bull (1987), Bentley MacLeod and Malcomson (1989), and Baker et al. (1994) consider a role for repeated interaction between the principal and agent as a means of reducing incentives to renege. Here the principal remains honest because a failure to do so involves retaliatory action through a future failure to provide effort. Second, in a static setting, firms can at times be induced to act honestly by imposing other contractual costs on them from claiming performance is poor. Charles Kahn and Gur Huberman (1988) show how the use of up-or-out contracts, where a worker is either retained at a high wage or fired, can solve this problem in settings where efforts involve training. In that case, if the principal claims poor performance, she must fire the agent, which may be more costly than retaining him at the high wage. (See Dominique Demougin and Aloysious Siow (1994) for another interpretation of up-or-out rules.) Canice Prendergast (1993) makes a similar point in the context of promotions. Finally, Jan Zabojnik (1997) argues that firms can mitigate the incentive of supervisors to steal deserved wage payments by designing contracts that emphasize revenues rather than profits.

[33] For instance, "Forrest Gump" supposedly earned negative net profits despite being the fourth highest grossing movie of all time. It is claimed that manipulation of the net profit figures occurred because the writer had been promised 3 percent of net profits. James Garner successfully sued Universal Studios when his 37.5 percent of net profits from "The Rockford Files" turned out to be worth zero. These cases are described in Cheatham, Davis, and Cheatham (1996).

on a daily basis. In most US cities, the cab driver has a piece rate of 100 percent; he pays a fixed fee for the cab and keeps all revenues. This compensation scheme is not used because driving a cab has little risk; the demand for cabs clearly depends on such variable factors as weather. Instead, cab drivers typically keep all their revenues because they can manipulate output, as true output cannot be observed. More specifically, they can turn off the meter and negotiate a fare with the passenger, as occurred in situations where piece rates less than unity were used. The most efficient (static) solution to this problem is simply to let the driver keep all revenues, as he no longer has an incentive to privately contract.

Compression of Ratings. There is considerable evidence in the personnel literature that supervisors distort subjective performance ratings by not sufficiently differentiating good from bad performance in their ratings. In this scenario, the supervisors are themselves agents, who have incentives to treat workers in ways not desirable to the principal when offering evaluations. Two relevant forms of compression are noted in this literature: "centrality bias" and "leniency bias." Centrality bias refers to a practice where supervisors offer all workers ratings that differ little from a norm. Leniency bias implies that supervisors simply overstate the performance of the poor performers.[34] Such compression is well documented in the personnel literature, where Frank

Landy and J. Farr (1980), A. Mohrman and Edward Lawler (1983), Kevin R. Murphy and Jeannette Cleveland (1991), and Patrick Larkey and Jonathon Caulkins (1992) document negligible difference in ratings and compensation across workers.[35] This reduces the value of subjective assessments as a means of providing incentives, since the relationship between effort and pay is clouded by other influences.

This literature also points out that such compression is more severe in situations where ratings are important for pay setting: supervisors are reluctant to impart bad news to workers if it means salary adjustments. Ironically, an implication of this is that many firms now explicitly separate pay setting from subjective evaluations. According to George Milkovich and Alexandra Wigdor (1991, p.109), "A traditional rule of thumb among managers has also suggested the wisdom of decoupling the appraisal process from merit pay . . . [The] concern has been that managers will deliberately inflate performance appraisal rating to distribute merit pay, thus decreasing the chances that employees with real training needs will be identified or increasing the chances that overrated employees will be promoted beyond their capacities."[36] From

---

[34] An obvious reason for this is that it is simply unpleasant for supervisors to offer poor ratings to workers, so they avoid this pain. It is also worth pointing out that such compression need not be inefficient in a dynamic setting. For instance, suppose that a worker performs poorly. Telling the worker that their performance was poor can easily result in discouragement, say because they feel that their promotion prospects are low. As a result, firms may prefer to reveal little information.

[35] Direct evidence on leniency in ratings is provided in H. F. Rothe (1949) and E.A. Rundquist and R. H. Bittner (1948), while Leonard Ferguson (1949) and Lee Stockford and H.W. Bissel (1949) illustrate that such leniency is exacerbated when the supervisor knows the subordinate for a long time.

[36] There is almost no empirical work in the economics literature on this topic (though see James Medoff and Katherine Abraham 1981 and Baker, Gibbs, and Holmstrom 1994a,b for indicative evidence). The only example I know of is a Harvard Business School case by Kevin J. Murphy on the compensation practices of Merck in the mid-1980s. During this time at Merck, supervisors were required to rate on a 1–5 scale, yet 97 percent of workers were offered ratings of 3 or 4,

this perspective, appraisals are largely used for training purposes rather than to allocate rewards to those exerting most effort.

Rent-Seeking Activities. In most large organizations, supervisors are not residual claimants on output, and so the incentive to underreport performance to keep costs low may not be critical. A more pertinent problem with subjective assessments in large firms may be the danger of "rent-seeking activities," which refers to any actions that agents carry out that are designed to increase the likelihood of better ratings from supervisors, but that have less value on surplus than some other activity that they could carry out. This has been the primary focus of the economics literature on subjective performance evaluation. The relevant theoretical work on this issue includes Holmstrom (1982), Milgrom (1988), Milgrom and Roberts (1988), Tirole (1992), and Franklin Allen and David Gale (1992). At this level of generality, there is nothing that makes influence activities specific to situations where evaluations are subjective.[37] However, a central theme of this literature, which arises with subjectively evaluated schemes, is that supervisors will misreport when evaluating workers. For instance, workers may curry favor in such a way that supervisors exaggerate performance, or, as in Tirole (1992), simply bribe their superior for a better evaluation. As a result, supervisors exhibit favoritism to those who spend the most time currying favor. For instance, D. Bjerke et al. (1987) note

how Navy supervisors admitted distorting performance ratings in order to increase the prospects of their preferred subordinates. Such activities have two possible distortions. First, agents devote time and energy to "sucking up" that would be better spent on productive tasks. Second, information may be inefficiently collected on individuals, so it may be difficult to determine whether good performance ratings derive from favoritism or from genuinely good performance.[38] As a result, firms may not know who to promote or reward.[39]

Contracts under Subjective Performance Evaluation. In this section, I consider the rent-seeking effects of

---

[37] For instance, a worker could exert effort on visible tasks (Jonathon Paul 1992) at the expense of those that are truly productive. Or he could work "too hard" in order to make a good impression (Holmstrom 1982).

where there were negligible differences between the two groups. Management felt that this discretion offered to supervisors offered few incentives to their workers.

[38] In many situations there are simply no objective measures of output that can used to reward the agents and supervisors. Other problems arise here. Consider the case of a figure skating judge. How would one give incentives for a judge to offer honest assessments of the performance of a skater? There are essentially no objective measures of output; instead, the only available measures are the subjective assessments of other judges. Yet agency concerns abound in these settings, and a solution used to evaluate judges is to compare assessments with the assessments of other judges. A judge is then penalized if she repeatedly differs from the average assessments of other judges. This is not an isolated case; similar problems arise with evaluating academic candidates, art and film critics, and so on. This gives rise to the problem of "yes men," whose objective is simply not to look different from anyone else, leading to less efficient evaluations. See Prendergast (1993) for details. Another example of the absence of output measures concerns situations where diagnosis is important, as in Curtis Taylor (1995). Consider the market for auto repair. There are few measures of output in an auto repair shop if the car is not broken; all that is known is that the car leaves in working order, not that the marginal product of the auto shop was high. As a result, an inefficiency can arise where the auto shop will claim repairs are needed even in cases where they are not.

[39] It is not necessarily the case that trades between supervisors and employees are inefficient. For instance, Prendergast and Robert Topel (1996) consider a model of favoritism that illustrates the costs considered above; also, these trades have benefits in that they offers power to supervisors, which they value. See Holmstrom and Milgrom (1990) and Hideshi Itoh (1992, 1993) for other work on efficient side-trades.

subjectivity. Suppose that in addition to the corrupted objective measure $\tilde{y}$, agents can also be rewarded on a subjective assessment $\tilde{s}$ made by a supervisor. In keeping with the rent-seeking literature, it is assumed that $\tilde{s}$ need not equal s, the true evaluation. Instead, the agent can ingratiate himself to his supervisors by carrying out a bias activity b, where he can induce the supervisor to make a report which is b higher than s at a personal cost of $K(b) = \frac{\kappa b^2}{2}$.[40] Therefore, on the margin "sucking up" has a payoff, though the firm realizes that the agent is carrying out such activities in equilibrium. The report made by the supervisor on the worker is $\tilde{s} = s + b$, so $\tilde{s}$ is the corrupted version of s. Let the compensation contract offered to the worker be given by

$$w = \beta_0 + \beta_{\tilde{y}}\,\tilde{y} + \beta_{\tilde{s}}\,\tilde{s}. \qquad (13)$$

The firm now must choose compensation weights for an objective measure that is subject to gaming opportunities and a subjective measure subject to influence activities. The worker now makes two choices, e and b, both of which depend on the contract offered. He optimally chooses

$$e^*(\mu,\beta_{\tilde{y}},\beta_{\tilde{s}}) = \frac{\mu\beta_{\tilde{y}} + \beta_{\tilde{s}}}{c} \text{ and } b^*(\mu,\beta_{\tilde{y}},\beta_{\tilde{s}}) = \frac{\beta_{\tilde{s}}}{\kappa}.$$

Therefore, increases in κ make influence more costly and hence less prevalent. The returns to the supervisor from the rent-seeking are assumed to be negligible, and routine calculations show that the optimal contract is characterized by the "piece rates"

$$\beta_{\tilde{y}}^* = \frac{\frac{c}{\kappa}}{(1 + \sigma_\mu^2)(1 + \frac{c}{\kappa}) - 1} \qquad (14)$$

and

$$\beta_{\tilde{s}}^* = \frac{\sigma_\mu^2}{(1 + \sigma_\mu^2)(1 + \frac{c}{\kappa}) - 1}. \qquad (15)$$

These piece rates illustrate the trade-off of gaming and influence. Here risk neutrality no longer guarantees efficient effort unless either the agent cannot exert influence ($\kappa = \infty$) or there is no incentive to "game" the objective scheme ($\sigma_\mu^2 = 0$). Subjective assessments rise with κ; i.e., as the cost of influence activities increases, firms will rely more on the subjective measures. As above, the objective measure's use falls with σ_μ^2. In the case where there are no objective signals that can be used ($\sigma_\mu^2 = \infty$), the optimal choice of $\beta_{\tilde{s}}^* = \frac{1}{1 + \frac{c}{\kappa}}$. This view of compensation contracting shows how pay-for-performance is constrained not by the risk-sharing considerations of Section 2.1, but rather by the behavioral responses of agents. The use of objective measures has the drawback that agents allocate their efforts at the wrong time (i.e., based on μ), while subjective assessments waste resources on ingratiation.[41]

Empirical Evidence on Subjective Contracts. A primary focus of the personnel literature is on the design and implementation of contracts for workers whose output is not easily observed. The issues that arise in this empirical literature concern optimal discretion offered to supervisors, the use of bureaucratic rules (such as maximum pay increases allowable within job

---

[40] The costs of effort, e, and bias, b, are independent purely for simplicity.

[41] This view of influence incentives considers how changes in effective piece rates on subjective measures induce influence. However, other variables may play a similar role. Margaret Meyer, Milgrom, and Roberts (1992) and Shaefer (1994) argue that the financial performance of firms may also induce changes in influence activities. For example, firms in decline may be laying off workers, so that the returns to influence could rise if there are considerable rents from retaining a job.

grade), and the costs and benefits of different evaluation schemes. A disappointment of the economics literature has been the paucity of information collected on the evaluation of workers with poorly measured output. Despite the fact that most workers in the economy are evaluated subjectively, the economics literature has largely focused on the aggregation of observed objective signals.[42] While we have learned much from this literature, the set of workers with easily observed output is a small fraction of the population.

Of course, many of the insights that govern compensation under objective measures also hold when workers are subjectively assessed. Despite this, the observations above make clear that there are many other influences at play with subjective assessments. For instance, if supervisors refuse to differentiate between the good and bad performers, the insights of Section 2.1 regarding trading off risk against incentives are probably of less importance than might be the case, as there is little variance in performance measures anyway. There is a need to collect more information about the evaluation and compensation of the worker for whom output measures are hard to obtain. Remarkably, the only studies in economics that I am aware of that address pay and evaluation for such workers are Brown (1990) and MacLeod and Daniel Parent (1998). Brown considers the determinants of standard rate pay, subjective merit pay, and piece rates for a large sample of workers. His most robust finding is that a greater diversity of du-

ties carried out by the agent reduces the likelihood of piece rate pay, consistent with the theory above. MacLeod and Parent also show that piece rate jobs typically involve few tasks, while those with many tasks are characterized by subjective assessments of performance.

2.3 *Tournaments*

So far, attention has been restricted to individual compensation schemes where the level of pay varies with the performance of the agent. However, in many situations agents exert effort in order to get promoted to a better paid position, where the reward associated with that position is fixed and where there is competition between agents for those positions. For instance, Gibbs and Hendricks (1996) use the personnel files of a large firm to illustrate little variation in pay within job grades based on performance; instead, most pay increases occur through job (or at least job title) changes. Since a large proportion of wage changes are associated with such promotions, a central theme of the economics literature, following Lazear and Sherwin Rosen (1981), has been to examine the incentive effects of promotion schemes via tournament theory. This section briefly describes the main themes of this literature.

Promotions are used for many different reasons, perhaps the most important of which is to sort workers on the basis of their talents. For instance, Rosen (1982) illustrates how a competitive labor market allocates workers to different positions based on their talents, and rewards them accordingly. An implication of such allocation decisions is that promotions also have incentive effects, and a common theme of the literature has been to address the incentive effects of promotions, through the lens of tournament theory. Tournament

---

[42] This is not to say we know little about how wages change within firms. There is considerable work using personnel files (such as Baker, Gibbs, and Holmstrom 1994a,b; Michael Gibbs 1995; and Gibbs and Wallace Hendricks 1996) illustrating how job tenure and promotions affect wages. What is not known is how these compensation decisions are made.

theory considers a group of agents competing for a fixed set of prizes. The prizes are specified in advance and agents exert effort to increase the likelihood of winning a better prize. Rather like a sports game, all that matters for winning is not the absolute level of performance, but how well one does relative to others.

I begin by considering the simple analytics of tournament theory. To do so, consider two agents 1 and 2, who exert efforts e^1 and e^2 respectively under exactly the same circumstances as in the section on risk sharing, where signals y and s are observed on each agent, which I call y^i and s^i for agent i. (None of the distortions associated with multi-tasking or subjective performance evaluation are initially considered.) They compete for two fixed prizes. To simplify further, I assume that the two signals are equally valuable so that $\sigma_s^2 = \sigma_y^2 = \sigma^2$. The principal designs a tournament in this setting by choosing (i) a prize to be given to the winner, W, (ii) a prize given to the loser, L, and (iii) a rule that determines who the winner should be. Since both signals are equally valuable, the optimal rule for determining who wins the prize is simply that agent 1 wins if

$$z_1 = \frac{y^1 + s^1}{2} \geq \frac{y^2 + s^2}{2} = z_2. \quad (16)$$

Otherwise, agent 2 should be awarded the winner's prize. As in the section on risk-sharing, this rule is nothing more than optimally aggregating information on performance and then awarding the prize to the worker who has highest expected effort.[43] While this may appear

obvious, this aggregation rule turns out not to be efficient when dysfunctional behavioral responses arise; this will become clear below.

All that matters for rewards and hence effort decisions is relative performance. Accordingly, note that the distribution of $z_1 - z_2 \sim N(e^1 - e^2, \sigma^2 + 2\sigma_\alpha^2)$. Assume that the agents are risk neutral. Then each agent exerts equilibrium effort until

$$e^{i*} = \frac{W - L}{c} \frac{\partial [z_i \geq z_j | e_1, e_2]}{\partial e^i}. \quad (17)$$

Since each is perceived to be equally likely to win, the marginal change in the probability of winning is the density of the distribution of $z_1 - z_2$ evaluated at zero. This implies equilibrium effort of

$$e^{i*} = \frac{W - L}{c\sqrt{2\pi(\sigma^2 + 2\sigma_\alpha^2)}}. \quad (18)$$

Therefore, the agent's effort is increasing in the size of the prize and in the efficiency of monitoring. Because the optimal level of effort is $\frac{1}{c}$, the firm sets the optimal prize $W^* - L^* = \sqrt{2\pi(\sigma^2 + 2\sigma_\alpha^2)}$ to induce the first best level of effort. Thus, as illustrated in Lazear and Rosen (1982), the principal has induced the first best level of effort through the use of a tournament.

Empirical Tests of Tournament Theory. Tournament theory offers a number of testable implications. First, greater prizes lead to more effort. A number of authors have verified this prediction, typically from the sporting arena. First, Ron Ehrenberg and Michael Bognanno (1990) illustrate that professional golfers on the European circuit have lower scores when the prize money for which they compete increases. They illustrate this both by

[43] Where workers are risk neutral and there are no allocation effects of promotion, it actually doesn't matter which (symmetric non-degenerate) aggregation rule is used, as the wage spread can be changed to counter any inefficiencies in the aggregation rule at no cost to wages. But this result is special; it occurs only in this case. If either the

agents are risk averse, or the firm is allocating the most able workers to more responsible jobs, the firm strictly prefers this aggregation rule to any other.

looking across tournaments (where different tournaments have different prize money) and by observing the incentives of players who start in different positions beginning the final round.[44] Becker and Huselid (1992) show that higher prizes result in faster (though riskier) driving by professional NASCAR drivers. One of the few tests of tournament theory outside a sports setting is Charles Knoeber (1989) and Knoeber and Walter Thurman (1994), who study the broiler chicken industry. Large broiler companies reward farmers on a relative performance metric rather like tournaments, in order to filter out important common risk due to such factors as disease. As predicted by the theory, higher prizes result in better performance, here measured by the weight of chickens.

Of course, these tests are simply more evidence on whether incentives matter, not whether contracts are designed with these responses in mind. However, a prediction of tournament theory is that in settings where there is a single winning prize,[45] the prize for victory is increasing in the number of competitors.[46] Empirical tests of this proposition have been carried out on executive data, where the return to becoming CEO is increasing in the number of individuals competing at the next rank below. Brian Main, Charles O'Reilly, and James Wade (1993), Tor Eriksson (1996), and Martin Conyon and Simon Peck (1997) illustrate that more competitors do indeed increase the prize for becoming CEO. Another prediction of agency theory is that biased tournaments (where one agent has a greater chance of winning than the other in a two-person setting) result in lower effort, and for those who fall behind, excessively risky behavior.[47] I know of no empirical work that illustrates that biased tournaments reduce incentives, other than casual observation in sports tournaments of teams "giving up" when they are behind at the end of the game and where both winning and losing teams replace their best players with substitutes. Knoeber and Thurman (1995) illustrate that broiler farmers who are unlikely to win the tournament do indeed take riskier actions in order to improve the prospects of winning.

A problem with tournaments is that since individuals are evaluated on how well they do relative to others, they are unlikely to help their competitors in need. This point is theoretically illustrated in Lazear (1989) and Raffi Rob and Peter Zemsky (1997).[48] Drago and

---

[44] The payoff in such tournaments is convex, where for instance the payoff for coming first rather than second is much higher than from coming 34th rather than 35th. Thus position starting the final round offers different incentives.

[45] This is the case that has been studied empirically, where the interest has been on one agent acceding from senior executive to CEO.

[46] This result relies on the distribution of the measurement errors being single-peaked at zero. The marginal return to effort can be parameterized by the density of the distribution of measurement errors evaluated at the equilibrium probability of promotion. Then as the number of competitors rises, the marginal effect of effort falls, since the density is lower as promotion becomes less likely. As a result, the prize must rise to compensate.

[47] In the context of the model above, suppose that one agent is A more able than the other. Then the marginal incentive to exert effort is $(W - L)f(A) < (W - L)f(0)$, where f is the density of relative luck.

[48] A related point arises in Lorne Carmichael (1988) and Guido Freibel and Michael Raith (1997). They describe a situation where agents worry about hiring good colleagues where there could be competition for available slots. In the absence of some constraint on competition or the allocation of rents, those who choose new workers will be likely to prefer the less able, as they do not constitute such strenuous competition. Carmichael argues that tenure for academics is a solution to this problem, as the positions of the insiders are already guaranteed. Freibel and Raith instead focus on restrictions on the communication of information as the optimal solution. In both cases,

Garvey (1997) find evidence consistent with this using survey data from the Australian manufacturing sector. They show that when agents report promotion incentives to be strong, they are less likely to let others use their equipment, tools, or machinery.

Why Are Tournaments Used? The available evidence suggests that to a large extent, firms primarily provide incentives through the prospect of promotion (Baker, Gibbs, and Holmstrom 1994a,b; and Gibbs and Hendricks 1996), where higher wages can only be attained through changing ranks. Rather surprisingly, there is very little work devoted to understanding why this is the case, i.e., why the optimal means of providing incentives within large firms (at least for white-collar workers) seems to be tournaments rather than the other means suggested in the previous sections.

An important function of promotions is in sorting workers to jobs. Promotion in many firms takes the form of a job change, in the sense that responsibilities increase with ability. While the issue of sorting workers to jobs has been studied at some length (Rosen 1982; Michael Sattinger 1993), the interaction between incentives and sorting remains little understood. At a very general level, it appears that promotion can "kill two birds with one stone," as it both improves the allocation of talented workers to jobs and provides incentives (Baker, Jensen, and Murphy 1988), but the exact mechanics of this remain unclear (though see Prendergast 1993; and Dan Bernhardt 1995). To phrase this another way, we know relatively little about how internal labor markets,

which must assign workers to tasks in firms based on comparative skills, interact with the provision of incentives for workers.

In the context of the standard model with risk aversion in Section 2.1, there is little reason why the firm should pay solely on the basis of relative output, as occurs in tournaments.[49] While agency theory suggests that relative performance should be used in situations where there is common risk, it is only in very special cases that the optimal means of compensation involves *only* relative performance evaluation (Dilip Mookherjee 1984), as occurs in tournaments. Intuitively, there is information on effort from the worker's absolute performance, independent of his rank, which is all that matters for tournaments. Given this, why are they so popular?[50]

A related reason to filtering out common shocks is that evaluators often cannot place a number on the performance of a worker, but are capable of making rank order comparisons. Thus, all that is necessary to carry out evaluations of workers is to determine which worker is better. In addition, since prizes are fixed, it is not necessary to determine how much better one worker is than another; all that is needed is rank order information. While this answer seems to have some plausibility, it is hardly complete. For instance, firms frequently have to make decisions based on the absolute performance of workers: for example, should they respond to a wage

---

these distortions arise because there are restrictions on the ability of agents to make efficient monetary transfers. For instance, if the insiders could sell their positions to the newcomers, efficient allocations would arise.

[49] It is of course true that with risk neutral workers who carry out one activity, the contract above gives rise to the first best. However, so do many other contracts, so why are tournaments typically chosen?

[50] In the description above, tournaments are effectively competitions between agents. However, an equally valid interpretation of promotion has agents competing against a fixed exogenous threshold, such as a tenure standard.

offer from a competing firm? Since they can make such absolute comparisons in this setting, it remains a mystery why they are ignored in the provision of incentives.

Another reason provided by the literature for the use of tournaments, where prize structures are fixed, is that it avoids the possibility of the firm reneging on paying wages. In the previous subsection, I mentioned that when performance is subjectively assessed, it cannot be verified to a third party. As a result, there is a danger of the firm underreporting performance in order to save wages. As pointed out in Carmichael (1983), firms can avoid this by committing to a prize structure, as in the model above, where the prizes W and L are prespecified. The workers realize that the distribution of prizes is fixed, yet retain incentives through the prospect of improving their particular prize. While theoretically correct, it is unclear how important this is in reality for two reasons. First, firms rarely commit to the size of prizes. Second, it would seem that in many cases they can easily renege, by claiming bad business conditions or whatever.[51] Finally, another solution is simply that the firm commits to a wage bill and allows the supervisor to assign wages as he sees fit; there is no need to specify that wages be attached to ranks.

Two more speculative explanations may also be offered for the use of tournaments. First, the psychology literature described above illustrates a marked reluctance of supervisors to distinguish the performance of the able from the less able. Instead, performance ratings are compressed around some norm. One advantage of tournaments is that they force managers to make decisions; they no longer have the "luxury" of paying everyone the same (or close to the same). Second, as mentioned above, promotions are linked to changes in responsibilities. An advantage of tying wages to responsibilities, pointed out in James Fairburn and Malcomson (1996), is that it may cut down on influence activities. Assume that a supervisor could simply allocate a fixed pool of bonus money in any way she chooses, independent of job assignment. Then she is particularly susceptible to rent-seeking activities or outright bribery to obtain those bonuses. By contrast, consider a case where the supervisor can offer a high wage only if the worker is reassigned to a more responsible position, for whose output the supervisor is responsible. In this setting, the supervisor is less likely to respond to the rent-seeking activities of a less able employee because promoting that employee results in lower productivity than if the most qualified person is promoted. Thus, tying wages to job responsibilities can reduce inefficient influence activities.

2.3.1 *Bureaucracy*

A central feature of organizations is the use of bureaucracy, where rules are used to allocate resources rather than allowing individuals discretion over resource allocation. In the context of human resources, many examples come to mind. First, Richard Freeman and Medoff (1984) illustrate the importance of seniority in promotion and layoff decisions, independent of profitability considerations. They note that among nonunion firms, almost 42 percent lay off solely on the basis of seniority

---

[51] Perhaps the solution to this problem is that in situations where labor markets reward talent, promotions become an effective form of commitment for the reason that if a worker is promoted to a more responsible position, his wage will rise simply because the labor market values these skills, as in Waldman 1984. Thus workers may be willing to exert effort in the hope of achieving this endogenously determined prize.

considerations, while a mere 14 percent ignore seniority, only considering profitability considerations. Second, Seymour Spilerman (1986) notes that supervisors are often constrained in the raises that they can offer to their subordinates, as job grades typically carry ranges (minimum and maximum) that cannot be exceeded. This feature is considered at some length in Spilerman and Hiroshi Ishida (1994), Baker, Gibbs, and Holmstrom (1994b), and especially Gibbs and Hendrick (1996), who address the provision of incentives to workers who are "maxed out" (i.e., are at the top of their pay ranges). In each case cited, it appears that these pay restrictions have real effects. Spilerman also notes that positions are often characterized by minimum experience requirements, where workers must stay in a particular position for a certain amount of time before they can be promoted. This occurs independent of the ability level of the agent involved. In each of these cases, discretion is taken from the hands of supervisors.

The essence of bureacratic rules is that resources are allocated in an ex post inefficient fashion. For instance, a worker is promoted on seniority even though a better candidate exists. Recent developments in agency theory, following Milgrom and Roberts (1988) and Tirole (1992), provide a simple reason for such rules: while rules harm ex post efficient allocations, they improve the incentive for agents to allocate their activities correctly, by avoiding influence activities. For instance, although promotion by seniority has allocative inefficiencies, at least there is little lobbying. I address this issue here in the context of the tournament model, because most of the prominent examples of bureaucracy involve such decisions. However, bureaucracy will typically occur in any

setting where agent can respond to compensation schemes in inefficient ways. To illustrate the incentive to act bureaucratically, two ingredients are necessary. First, some measures of performance must be corruptable. I illustrate this by considering a situation where the subjective signal is subject to influence activities. Second, bureaucracy has the connotation that information is not effectively collected. To model this, I consider a situation where promotion involves the allocation of the worker to a new position, where there is a higher return to ability. As a result, the firm would like to aggregate information efficiently to minimize worker misallocation.

In particular, the winner of the tournament is now assigned to another job, which is identical to the previous job except that the (linear) marginal return to ability is higher.[52] Thus the winner of the tournament is reallocated to a new position. Since the winner is assigned to a job with higher return to ability, there is a return to identifying which worker is more talented; this reduces the probability of inefficient allocation. The other distinction from the set-up in the previous subsection is that that the subjective signal can be distorted, as in Section 2.2. Thus, the agent is evaluated on a non-corruptible objective measure y^i as above, but also on $\tilde{s}^i = s^i + b^i$. The cost of bias is as in the previous subsection, $K(b^i) = \dfrac{\kappa b^{i2}}{2}$ and the two noise terms are equal, $\sigma_s^2 = \sigma_y^2 = \sigma^2$.

Consider the ex post optimal allocation rule. Since the productivity of the most talented worker is higher in the promoted position, the ex post optimal rule places the "best" worker in that position. This means that agent 1 should be

---

[52] So expected output is given by $e + \gamma\alpha$, where $\gamma > 1$.

promoted if

$$z_1 = \frac{y^1 + s^1}{2} \geq \frac{y^2 + s^2}{2} = z_2,$$

exactly as in the previous section. Otherwise, agent 2 wins the tournament. This is the *non-bureaucratic* allocation, as decisions are made in an ex post optimal fashion.

Now consider an alternative decision rule, where the measure used to reward and promote is

$$\hat{z}^i(\theta, y^i, \tilde{s}^i) = \frac{y^i + \theta \tilde{s}^i}{1 + \theta},$$

for $0 \leq \theta \leq 1$. Thus, worker i is promoted if $\hat{z}^i(\theta, y^i, \tilde{s}^i) > \hat{z}^j(\theta, y^j, \tilde{s}^j)$, which deviates from the ex post optimal rule if $\theta < 1$. Let $\Omega(\theta)$ be the gain from allocating workers using an aggregation rule with weight θ, for $0 \leq \theta \leq 1$, relative to not using the $\tilde{s}$ signal.[53] Bureaucracy arises if $\theta < 1$. The reason bureaucratic rules are used here is that the rent-seeking activity is increasing not only in the prize for promotion $W - L$, but also in θ, the weight placed on the signal. It is straightforward to show that the level of bias chosen by both agents is given by

$$b^{i*} = \frac{\theta(W - L)}{\sqrt{2\pi} \Sigma(\theta) \kappa}, \qquad (19)$$

where $\Sigma^2(\theta) = V(\hat{z}(\theta))$. Thus agents exert more bias when either the monetary return is high, or when the effect of the signal on decisions is high. Critically, influence activities are increasing in θ. As a result, the firm commits to underweight the corruptable signal. It is simple to show that the optimal choice of θ^* is given by the first order condition

$$\Omega'(\theta^*) = \frac{\theta(W^* - L^*)}{\sqrt{2\pi} \Sigma(\theta)} \frac{db^i}{d\theta} > 0, \qquad (20)$$

where W^* and L^* are the optimally chosen prizes. The key point here is that if

---

[53] It is easy to show that $\Omega(0) = 0, \Omega'(\theta) > 0$ for $\theta < 1$, $\Omega'(1) = 0$, and $\Omega''(\theta) < 0$.

workers exert influence activities

$$(\kappa < \infty \text{ so } \frac{db^i}{d\theta} > 0),$$

the optimal choice of θ^* is less than 1 and bureaucracy is used. In other words, firms commit to bureaucratic rules in order to reduce the incentive to engage in influence activities, even though this sometimes involves the misallocation of resources.[54] Hence the optimal provision of incentives involves bureaucracy.

2.4 *Team Production*

Most workers hold jobs that involve productive interactions with their colleagues, where output reflects the contribution of many individuals. Team production problems potentially arise in situations where individual contributions to output cannot be easily identified and compensation must be based on team production. In that setting, the classic free-riding problem arises, where agents fail to internalize the benefits that accrue to other members of the team when making effort decisions. This effect, which has also been referred to as the $\frac{1}{N}$ problem (since each agent receives this share of output in a partnership with N members), prevails in situations where rewards cannot exceed the revenues of the group (Holmstrom 1982; Kenneth McLaughlin 1994). The available tests of free riding in teams largely come from the observation of partnerships in law firms or medical practices. First, Joseph Newhouse (1973) considers the effect of

---

[54] At an anecdotal level, the economics department of an esteemed U.S. university uses a rule where once a faculty member has published a certain number of papers, he is offered tenure. This rule, which takes away important discretion from the hands of evaluators, was introduced because of previous accusations of favoritism, when deans and senior faculty had discretion over promotion decisions. This is used to reduce the incentives of the junior faculty to curry favor with their senior colleagues, even though it sometimes induces inefficient promotions.
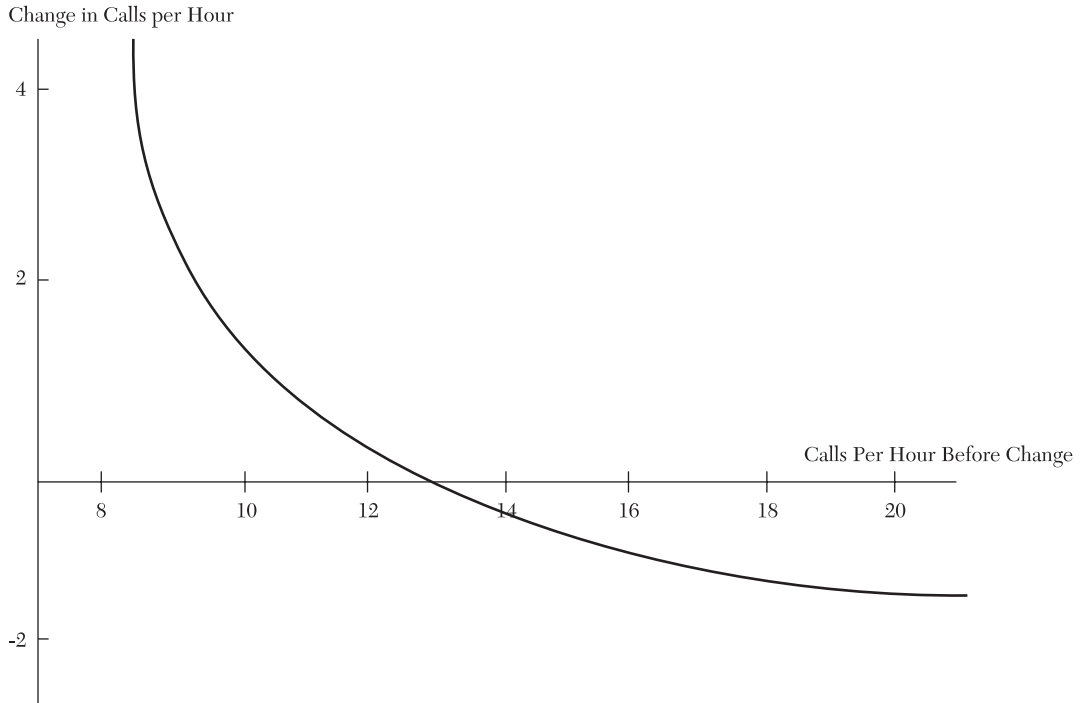
Change in Calls per Hour

Figure 4. Response of Telephone Operators to Team-Based Compensation (Hansen 1997).

group incentives in a medical practice, and notes that when the fraction of revenues that are shared with others rises, (i) overhead costs rise, and (ii) doctors work fewer hours. Richard Bailey (1970) finds qualitatively similar results, while Arlene Leibowitz and Robert Tollison (1980) find that larger law partnerships typically result in worse cost containment. These studies simply compare productivity measures of partnerships on different sharing rules without addressing why contracts vary, and so are subject to obvious selection criticisms. For instance, it could be that the less able work in teams since they have less to share, which could explain the low performance measures, independent of any behavioral effect of teams. Martin Gaynor and Pauly (1990) use survey evidence on medical practices, where reported risk aversion is

a measure used to exogenously identify variation in practice size.[55] They illustrate that poorer measures of performance arise when more revenues are shared with others, once again endorsing the importance of the free-rider problem.[56]

There are many possible solutions to

[55] The idea here is that more risk-averse doctors will operate in larger practices, as they value income-smoothing opportunities more. Once these measures of risk aversion are shown to be independent of other productivity measures (and that the instrument has some explanatory power), this is a legitimate identification strategy.

[56] However, this is not to say that team-based rewards cannot generate incentives relative to no incentive pay, as illustrated for the steel industry by Boning, Ichniowski, and Shaw (1998), who also effectively control for heterogeneity in contract choice through variation in the manufacturing environment. See also Encinosa, Gaynor, and Rebitzer (1997) for an interesting attempt to distinguish between economic and noneconomic notions of team production and compensation.

the team production problem.[57] One solution that has attracted some attention has been the use of peer pressure, where agents can possibly monitor one another and mete out punishments to those who fail to perform adequately. If the cost of such monitoring is sufficiently low, this can negate the free-rider problem, as illustrated in Eugene Kandel and Lazear (1992). However, empirical evidence on peer pressure reveals behavioral responses different from those posited in the theory. Andrew Weiss (1987) and Daniel Hansen (1997) consider the effect of team-based compensation on the *individual* productivity of agents. (In both cases, the employers choose to pay on the productivity of the team despite the availability of measures of individual production.) Weiss studies the productivity of blue-collar workers in a pharmaceutical company, while Hansen addresses the incentives of telephone operators for a large financial company. Both authors illustrate that the use of team-based compensation schemes improves the performance of those who were less productive on individual schemes but *decreases* that of the more productive.

Hansen's results are summarized in Figure 4. He studies the performance of telephone operators, measured by the number of calls they handle in an hour. Figure 4 plots the change in the number of calls dealt with after the introduction of team-based compensation schemes as a function of the number of calls that were handled under fixed wages. The negative slope shows that the more able agents reduce the number of calls made while the less able improve. A related point concerns the selection effect of teams. While one's first impression is that team production is likely to be more attractive to the less

able (with the more able preferring individual based schemes), Weiss identifies a U-shaped relationship between worker turnover and prior productivity. In his sample, the medium-ability workers are more likely to remain than either the more able or the less able when placed on team-based compensation. One interpretation of this is that the most able leave as they prefer individual-based schemes elsewhere, while the least able also leave as the peer pressure makes their jobs too unpleasant. Thus the (admittedly scant) evidence suggests that team-based compensation gives rise to problems when workers vary in their ability.

Most of the work on team compensation concerns profit-sharing schemes, as in Derek Jones and Takeo Kato (1995), Douglas Kruse (1993), and Marc Knez and Duncan Simester (1997). These studies are carried out on large firms, where the wages (or often pensions) of employees are based on the profits of the entire firm, either through ESOPs or bonuses. Standard reasoning of the $\frac{1}{N}$ problem suggests that there should be a negligible response by agents to these incentives, since, for example, a worker who gets to keep $\frac{1}{1,000}$ of the returns to effort in a 1,000-worker firm should have few incentives. Despite this, studies consistently show that the productivity of firms using profit-sharing plans exceeds that of those that do not, with available estimates suggesting improvements in the range of 4–5 percent from these schemes.

Since these results appear to be such a violation of standard agency theory, alternative explanations have been sought. The most popular versions involve some notion of either peer pressure[58] (as described above), mutual

---

[57] See Holmstrom (1982).

[58] It should be emphasized that merely alluding to peer pressure hardly suffices here. For profit sharing to have effects on peer pressure in a large
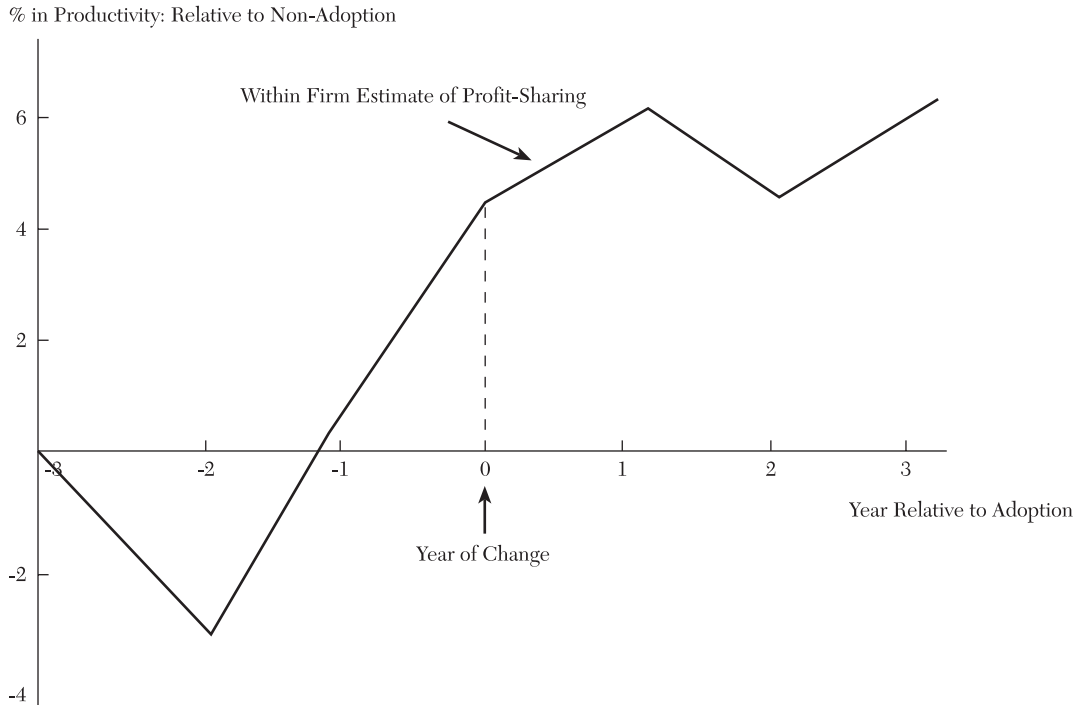
% in Productivity: Relative to Non-Adoption

Figure 5. Profit Sharing and Productivity (Kruse 1993).

monitoring, or "belonging," where the employees feel as if they are "in this together." Without meaning to dismiss these potential motivations, there are a couple of reasons to be skeptical about the validity of these results as a test of team production incentives. These doubts arise because (i) the data may not really illustrate productivity increases due to the compensation scheme, or (ii) the observed increases, though related to the compensation changes, may have little to do with the team production problem. First, the cross-sectional data illustrate that firms that use profit sharing have higher productivity than those that do not. In the

cross section, this could simply reflect the possibility that firms with no profits rarely introduce such schemes, so higher profitability could have little to do with the effect of such schemes. Researchers have solved this by looking "within firm." In other words, does productivity rise in those firms with profit sharing more than in those without such schemes?

Using this methodology, a large-scale study by Kruse (1993) finds that this is the case, where productivity rises by 3 percent more in firms with profit sharing than in those without. While this is an interesting approach to understanding the effect of pay on performance, and a considerable improvement over existing work, it constitutes a legitimate identification strategy only if the *trend* in productivity changes is identical between the two sets of firms.

___

firm, it must be the case that the costs of enforcement through peer pressure (pointing out errors or slacking to the relevant person) must be negligible, since the monitor equally receives only a $\frac{1}{N}$ share of any improvements herself.

Suppose, for example, that some firms have declining productivity and do not use profit sharing, while others have rising productivity and use profit sharing. Then even if there were no effects of the compensation scheme on productivity, this methodology would suggest such a relationship since there are unobserved differences in *changes* in productivity correlated with the introduction of the contracts. Relevant empirical evidence taken from Kruse is presented in Figure 5.

This measures productivity changes (value added per employee) for the adoptors of profit sharing compared to those firms that do not implement profit sharing, measured from three years before (-3) to three years after $(+3)$ the scheme was initiated. From the point of adoption (date 0) until 3 years later, productivity rose by 3 percent, suggesting an effect caused by the compensation scheme. This is the "within firm" estimate used in the literature. However, a comparison of the firms before adoption suggests that productivity may have been rising faster in those firms anyhow.[59] Perhaps the 3 percent is merely a continuation of that trend, so that the productivity effects are not caused by the compensation plans.[60] This is not, of course, conclusive evidence that profit sharing does not work; instead, my objective is simply that one should be wary of simple "fixed effect" estimates as a way of eliminating unobserved heterogeneity.

Assume for the moment that these problems were solved, and that the effects of the compensation schemes on productivity were robust after controlling for this issue. Could we then conclude that agents are willing to exert effort despite the $\frac{1}{N}$ problem, so that the free-rider problem loses some of its potency? A second problem with the literature is that these studies do not generally test for free riding in a team setting, in the sense laid out in the theoretical literature, because the theory considers the effect on incentives *holding utility constant*. But this is not the case with profit sharing; compensation rises in most firms that use profit sharing; Knez and Simester (1997) for one example.[61] Could the empirical results simply reflect the effect of giving workers more money, and not the effect of team production? To take an extreme example, suppose that profit sharing increased the pay of a worker by 25 percent. There are a multitude of reasons to expect that such an increase in wages will improve productivity. An obvious implication is that the firm will attract better workers, or existing workers will work harder as their jobs are more valuable; this is the premise of the efficiency wage literature described below. In either case, productivity will rise in a way that has little to do with the fact that pay depends on profits; instead, incentives and selection effects arise simply from more pay, not pay conditional on firm behavior. This is not a problem, of course, if all the authors are interested in is the effect of

---

[59] This effect is not specific to the use of profit-sharing schemes. Hassan Tehranian and James Waegelein (1985) illustrate that executive compensation plans relating pay to performance are generally introduced after a number of years of abnormally positive stock returns.

[60] This depends on the time-series correlation in productivity changes. If there is no correlation in changes, then the Kruse interpretation is clearly right, as those firms with productivity growth are no more likely to have future productivity growth than those without. However, to the extent that productivity changes are positively correlated, these results overestimate the effect of profit sharing on productivity.

[61] The theory suggests that wages should rise, but only by the increased cost of effort plus any risk premium associate with the variability in wage payments. However, there remains the possibility that profit sharing gives rise to rents earned by workers.

profit sharing; however, it does preclude testing the classical team production problem.

2.5 *Efficiency Wages*

So far, it has been assumed that workers earn their reservation utilities, where a firm does not offer rents to its workers to induce effort exertion. Efficiency wage theory concerns situations where firms offer workers such rents in order to induce effort exertion. In the context of incentive provision, firms overpay workers in order to make their jobs valuable, which makes them less likely to shirk. In this way, the cost of job loss (which ensues if agents are caught shirking) is large, so they exert effort at the efficient level (Carl Shapiro and Joseph Stiglitz 1984; Dan Raff 1992; Daron Acemoglu and Andrew Newman 1997). A simple way to interpret the shirking version of efficiency wage theory is to consider a situation where the agent's wage cannot be reduced below 0, which is assumed to be the reservation utility. In other words, even if the agent is caught shirking, he cannot be penalized by offering him a wage less than the reservation utility. To simplify matters, assume that the effort decision e is binary, set equal to either 0 or 1, so effort of 1 has a marginal cost of $\frac{c}{5}$. Monitoring is such that the worker who shirks is caught with probability p. Since the worker cannot be penalized below 0 for shirking, the firm must offer a wage of at least $w^* = \frac{c}{2p}$ to induce effort exertion, which implies that the worker earns rents of $\frac{(1-p)c}{2p}$ from the relationship. Thus, inefficient monitoring ($p < 1$) yields rents for the worker.

This theory has spawned a large literature, ranging from studies of unemployment to examinations of interindustry wage differentials. A small number of papers have directly tested for the importance of efficiency wages using firm-level data by examining the relationship between supervision and wage rents. A reasonable conjecture is that the probability of being caught shirking is increasing in the supervisor-worker ratio. It immediately follows that firms face an isoquant in (wage rent, supervisor-worker ratio) space, where they can trade off higher wages against more supervisors. Thus, wage rents and supervisors are substitutes. Erica Groshen and Alan Krueger (1990) address this issue using hospital employee data, and find evidence in favor of the theory. By contrast, Derek Neal (1992) uses more aggregate data and finds little relationship between these variables.

It is difficult to test for the existence of efficiency wages, where workers earn rents to induce effort exertion. First, while finding that wages and supervisors are substitutes along an isoquant of fixed effort is consistent with efficiency wages, exactly the same conclusion is true in the basic agency model with no rents.[62] Thus, this is a test of incentive theory, not necessarily a test of workers earning rents. In order to test for rents, one would need to see, for example, whether higher levels of supervision within a job increase worker turnover (since more supervisors reduce wages).

A second possible problem with this methodology concerns the prospect that the observed variation in supervisors

---

[62] To see this, remember that holding effort fixed, wages in the basic model exceed reservation wages both by effort costs and by the riskiness of the evaluation procedure. But if supervisors can be hired to provide more accurate reads of performance, wages fall for a fixed level of effort as the riskiness of the compensation scheme falls. Thus, once again wages and supervisors are substitutes, though without any implications for the existence of rents.

and wages across firms may not involve moving along a common isoquant in wage-supervisor space. More specifically, in available data, would we expect to see wages and supervisors as substitutes or complements? The problem is that either may easily arise in a world of efficiency wages and depends critically on the source of variation across firms. On the one hand, if the source of variation across firms is the cost of supervisors, then the two instruments are likely to be substitutes, where firms substitute away from high-cost supervisors into wages. On the other hand, if the source of variation across firms is in the return to effort (so some firms value effort exertion more than others), those firms that want more effort will use more of both instruments relative to those that do not value such high effort. This effect, which relies on the marginal cost of each instrument to be increas ing in its quantity, implies that supervisors and wages will be complements in the data.[63] As a result, it is hard to see how one can refute the existence of efficiency wages with this methodology.[64]

3. *The Dynamics of Agency Contracts*

The focus of the paper so far has been on static contracts, where the contracts and behavior in one year have no effect on future contracts. Yet employees and firms make matches that typically last for long periods of time, and a considerable literature now exists that addresses intertemporal links in the contracts offered to workers, where the contract offered this year depends on last year's con-

[63] See Susan Athey and Scott Stern (1997).
[64] Groshen and Krueger attempt to control for this by arguing that rates of supervision are largely set by regulation, though it is unclear how regulatory authorities set these levels.

tracts and realizations. This section addresses the primary themes in this literature.

3.1 *Deferred Compensation*

When agents remain with an employer for a long period of time, there is no necessary reason why the employer should pay the worker his expected marginal product in all periods; instead, workers could be paid better in some periods than in others. One aspect of this that has attracted both theoretical and empirical interest has been "deferred compensation," where workers are overpaid when old, at the cost of being underpaid when young. From this perspective, part of the reason why older workers are better paid than younger workers is not that they are more productive, but simply that they have accumulated enough tenure to garner these contractual returns. To make matters clear, the purpose of this subsection is to address whether and why firms use compensation schemes like Figure 6, where wage exceeds productivity for older workers, but is less than productivity for younger workers.

Many explanations have been offered for the use of such compensation profiles. Stephan Salop and Joanne Salop (1976) argue that delayed compensation aids the selection of desirable workers. For example, firms often incur significant turnover costs when workers leave, and one way of attracting those who are less likely to leave is to offer (quasi-)rents only to those who remain at the firm for long periods of time. An alternative possibility is that deferred compensation is useful because there may be a significant delay in observing the effects of efforts. As a result, firms may prefer to wait to generate a better inference on the worker's performance before rewarding
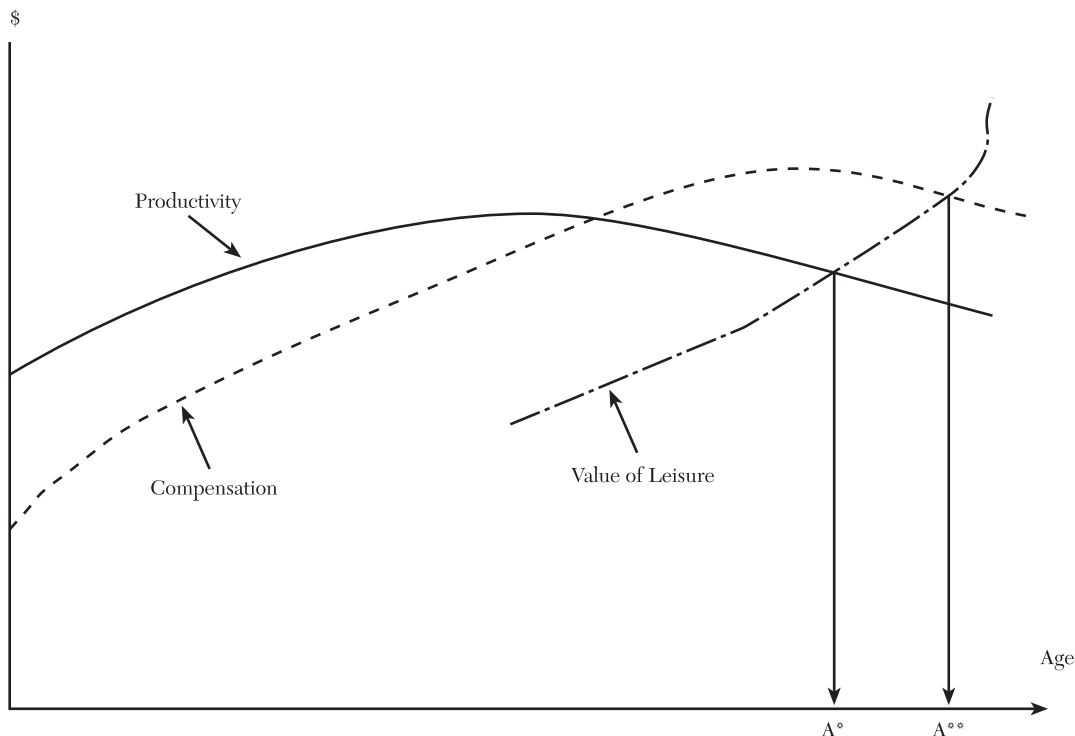
Figure 6. Deferred Compensation.

him.[65] However, the primary focus on deferred compensation has been as a means of providing incentives to workers, as in Lazear (1981). The idea here is simple. Consider a firm that offers rents to its older workers for the efficiency wage reason described above. For large enough rents, older workers are willing to exert effort rather than be fired. But rents to older workers are also attractive to younger workers, because exerting effort increases the likelihood of surviving in

the firm long enough to attain those rents. As a result, younger workers can be offered lower current compensation than older workers (relative to market options), while maintaining incentives for all (Akerlof and Lawrence Katz 1989).

To understand the mechanics of this problem, consider the efficiency wage model above, where there are two periods of the worker's career, "young" and "old." (In this section, I will typically consider two-period settings for simplicity.) In the single-period setting, it was shown that the firm must offer the agent a wage of $w^* = \dfrac{c}{2p}$ to induce effort exertion of $e = 1$. Since "old" workers have only a single period of employment remaining, the firm will offer that wage when workers are old. Remember, how-

---

[65] Finally, it may be that wages are deferred simply because workers have preferences for wages that increase with age. This is interpreted either as a preference for thinking that we are doing better from year to year, or as a means of forced savings, which agents do not trust themselves to do. See George Loewenstein and Nachum Sicherman (1991) and Robert Frank and Robert Hutchens (1993) for empirical evidence on such preferences.

ever, that this entails rents of $\frac{(1-p)c}{2p}$ for the worker at this point in life. The firm can take advantage of this when designing the younger worker's contract, because if the younger worker is caught shirking, he loses not only the rents associated with the "young" wage, w_y, but also the future rents $\frac{(1-p)c}{2p}$ because the firm fires him if he is caught shirking.[66] Let δ be the discount factor between the two periods. Then it is trivial to show that the agent will exert effort of $e = 1$ so long as

$$w_y \geq \frac{c}{2p} - \frac{\delta(1-p)c}{2p}, \qquad (21)$$

which is less than w^* if $p < 1$. In other words, the wage offered to the young is strictly less than that offered to the old. Thus, the firm defers compensation as part of an optimal payment package.[67]

Do Firms Defer Compensation? Determining whether firms defer compensation is conceptually easy; simply compare wages to productivity. However, productivity is typically difficult to measure. But there are some occupations for which productivity is observed.[68]

One possibility is to compare wages to productivity in those occupations. However, it is in those occupations that we expect that deferring compensation is unlikely to be used. For instance, consider the compensation of salesforce workers. Because their productivity is easy to observe, compensation can be based directly on those measures; there is no need to use deferred compensation.[69] For this reason, one may not be able to consider those occupations with easily observed output to discover the extent of deferred compensation.

Despite these constraints, a number of papers have shown wage changes for older workers that appear to have little to do with productivity effects. Each of these studies has some problems, in that there are other interpretations, but the aggregate picture suggests the deferring of compensation. First, Richard Freeman and James Medoff (1984) and Spilerman (1986) illustrate that firms often build seniority provisions into pay, promotion, and retention decisions, even when not warranted by productivity considerations.[70] For example, rules that promote workers on the basis of seniority rather than productivity offer such workers tenure-related advantages. Second, Medoff and Katherine Abraham (1980) illustrate that

---

[66] It should be noted that perhaps the most common way in which firms provide incentives to workers is through the threat of being fired if their performance is not satisfactory. However, the paper has had little to say about this means of incentive provision. Firing can be seen as a form of nonlinear incentive contract where the worker is paid a fixed wage for performance above some critical level, but is terminated otherwise. This simple model provides a reason for such incentive schemes. In particular, the firm uses firing as a way of excluding agents from future benefits which would accrue if they retain their jobs. In the presence of liquidity constraints, this becomes an efficient means of incentive provision.

[67] The central feature of deferred compensation is that good performance in one period yields the opportunity for benefits in future periods. Similar results in the context of tournaments arise in Meyer (1992).

[68] An interesting recent study on this is Leventis (1997b), who considers the productivity and pay of

surgeons. He illustrates that surgeons' pay rises with age, although their performance as a surgeon (risk adjusted mortality rates) becomes *worse*. Of course, it could be that older surgeons have more unmeasured aspects of production, such as training, which is reflected in higher wages.

[69] In fact, one could argue that if we find evidence of backloading for those occupations, it must be that backloading occurs for reasons other than incentives, since incentives can be provided in simpler ways.

[70] Robert Frank and Robert Hutchens (1984) also consider the wage profiles of two occupations (bus drivers and airline pilots) where wages continue to rise with seniority, even though there is little reason to expect productivity to increase after some initial period.
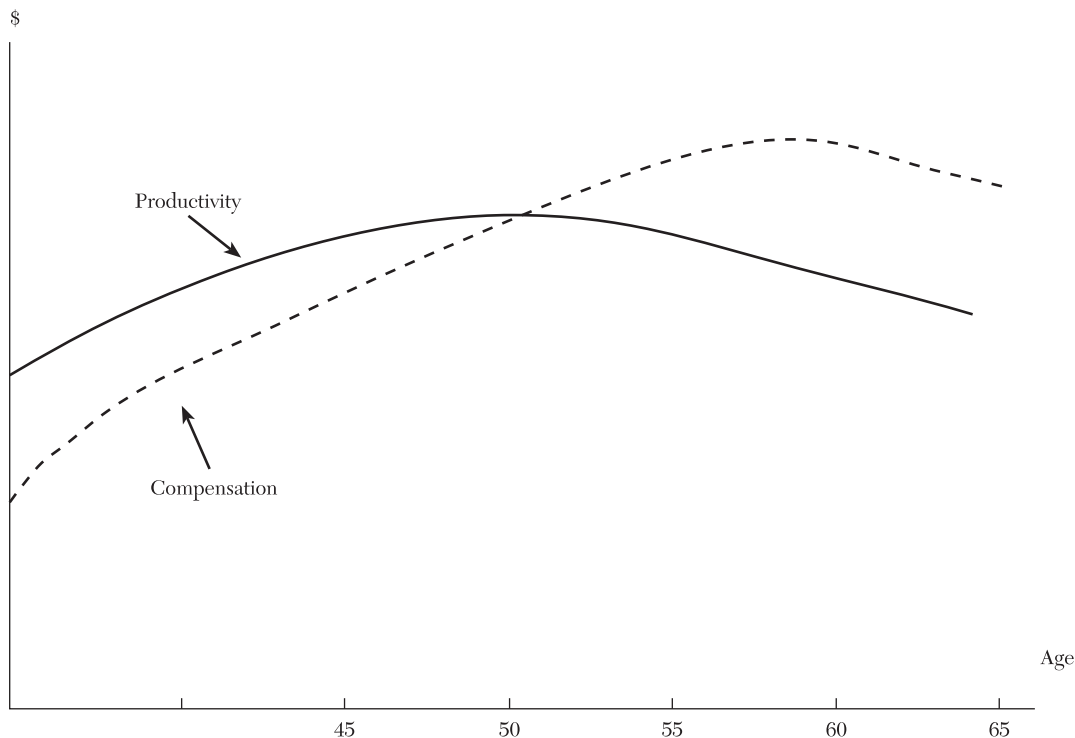
Figure 7. Deferred Compensation for Office Workers (Kotlikoff and Gokhale 1992).

the performance evaluations of senior workers differ little from those of their less senior counterparts, yet their wages are higher. They interpret this as further evidence of the use of deferred compensation.[71]

Another approach to addressing the importance of deferred compensation is to compare the wage profiles of the self-employed to those in similar positions who are employed by firms. Consider two workers who, say, are consultants, where one is self-employed and the other is an employee of a firm. If they both carry out the same job with equal efficiency, the wages of the self-employed consultant should be a good

proxy for the productivity of the employed person, since there is no one to shield the self-employed worker from changes in his productivity. Lazear and Robert Moore (1986) show that the wage profiles of the self-employed are indeed flatter than those of the employed. Hence if the wage profile of the self-employed maps the productivity of the employed, this suggests the "overpayment" of older employed workers.[72]

In a similar vein, Lawrence Kotlikoff and Jagadeesh Gokhale (1992) use the wages of newcomers to a large firm to identify the returns to seniority within

[71] An alternative interpretation of these data is that assessment standards depend on seniority, i.e., workers could be assessed relative to their potential, in which case senior workers could be better despite similar evaluations.

[72] Of course, there are other interpretations. For instance, it could be that more training is provided to employed workers, which they pay for early in their careers, but garner the returns later in life.

firms. Consider two 65-year-old workers in the same position, one of whom has been with the firm since leaving high school and the other who has just joined. The firm has little obligation to the newcomer and will pay him his marginal product. But if specific human capital is of limited importance, newcomers in the same position as those with more tenure should have similar productivity. If so, any differences in wages between the newcomer and the worker with longer tenure could be attributed to a difference between wages and productivity. Thus, they proxy the productivity of workers with long tenure by the wages of newcomers in the same position. Using this methodology, the authors can roll back the productivity profile of a worker based on age of arrival. Their results for two occupations (office workers and salesforce workers) are replicated in Figures 7 and 8 below, where the effect of a pension vesting at 55 is ignored to simplify the picture.

Figure 7 provides evidence on the compensation of office workers; the authors note a discernable difference between pay and productivity, with younger workers being less well paid than their alternatives, and older workers earning more. These data are consistent with a view that for occupations where it is difficult to provide objective measures of performance, as would be the case for these white-collar clerical workers, the optimal means of providing incentives is to offer "carrots" in the future. By contrast, consider the compensation of salesforce workers in Figure 8.

Here there is little difference between wages and imputed productivity. A plausible interpretation of the difference between this and Figure 7 is that for salesforce workers, incentives can be provided by tying pay

to readily available measures of productivity, so that there is little need to use deferred compensation. It should be remembered here, however, that an alternative interpretation of Figure 7 is that specific skills are important, so that newcomers earn less because they are less productive, even within positions.[73]

3.1.1 *The Returns to Promotion in Hierarchies*

The essence of the previous section is that contracts in one period depend on contracts in previous periods. Another application of this idea concerns how the returns to promotion vary as workers progress through a firm's hierarchy. It is well known that the returns to promotion increase at higher ranks in a firm. See Gibbs (1992), Main et al. (1991), Conyon and Peck (1997), and Richard Lambert et al. (1993) for details.[74] Since workers typically progress through the ranks of firms over time, this has an obvious relation to deferred compensation. More generally, this suggests inter-temporal linkages in contracts, where prizes at one level depend on previous prizes. A number of reasons have been proposed for this behavior. First, ignoring incentive issues, Rosen

[73] An auxiliary implication of deferring compensation is that workers will be reluctant to retire, i.e., workers who are paid "too much" will stay too long. Following Lazear (1981), Figure 6 illustrates that workers will retire when their wage equals the value of their leisure, i.e., at time A^{**}. On efficiency grounds, they should retire at A^*, when productivity equals the value of time. Mandatory retirement at A^* is efficient (though there are other possible mechanisms to replicate this outcome, such as the use of pensions). Hutchens (1988) provides some empirical evidence consistent with this model, by noting that US firms whose wage profiles rose rapidly as workers aged were also those which tended to use mandatory retirement.

[74] None of these empirical papers has attempted to distinguish between the competing hypotheses below, though they are often framed in terms of the Rosen (1986) work.
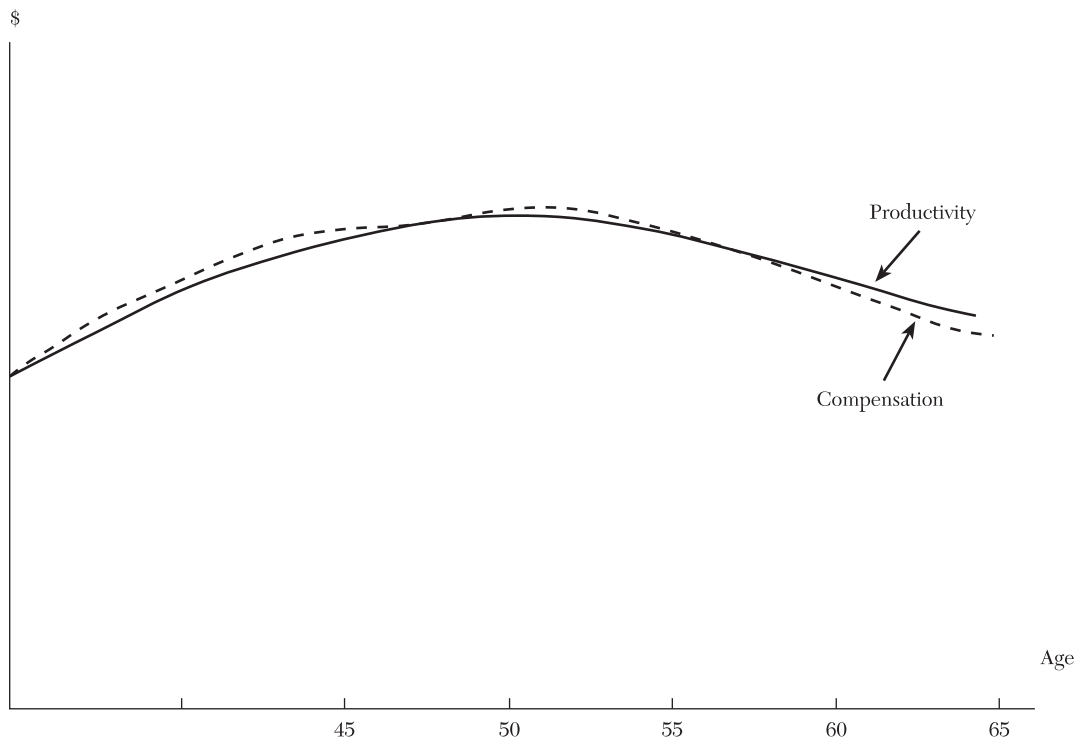
$

Figure 8. Deferred Compensation for Salesforce Workers (Kotlikoff and Gokhale 1992).

(1982) illustrates that in hierarchies where the decisions of superiors have implications for the marginal productivity of those in lower positions, there is a large return to ability. Due to what has become known as the "magnification effect" (where the decisions of senior workers are magnified many times), the returns to ability are convex, so on simple marginal productivity grounds, more able workers will earn many times the wages of their less able counterparts.

In the neoclassical model, the wage earned by a worker is the supply price of labor. Despite this, the wages of senior executives often triple overnight when they accede to the position of CEO, so it is doubtful that this is the only influence generating wages. As a result, it is generally felt that incentives

also play a role. A number of possible explanations for convex wage structures generated by incentive considerations can be imagined. First, income effects may cause wage increases on promotion to rise as workers ascend the hierarchy. Quite simply, it may take more money to induce effort from the rich than from the less well off.[75] Second, raises upon promotion may increase because the optimal level of effort is higher at more elevated ranks, as decisions made at higher ranks have more wide-reaching effects; it is more important for the CEO to work hard than for a shop floor

[75] This relies on the marginal rate of substitution between income and leisure varying with the level of income, unlike the exponential utility function described above. For instance, a utility function of the form $V(w, e) = U(w) - C(e)$, where $U(w)$ has the usual properties of risk aversion, will suffice.

worker to do so. If the marginal return to effort is increasing in rank, convex wage profiles will arise.

Neither of these reasons implies that rents are awarded to workers at higher ranks. However, a final reason proposed by Rosen (1986) argues that such convex wage schedules may offer rents to workers in senior positions as a means of providing incentives to all workers. Rosen points out that part of the return for promotion from rank *a* to *b* is not simply the pay difference between the two ranks, but also the increased prospect of further promotion to ranks *c*, *d*, and so on. A suitable analogy is that part of the return to winning the first round of a tennis tournament is not just the prize money for reaching the second round, but the deferred prospect of the prize money for future rounds. This idea, which has been labeled the "option" value of promotion, implies that wage increases from promotion are decreasing in the prospect of future promotions. Thus the winner of the final round must be offered greater incentives than the winner of the first round, for whom the deferred benefit is larger. This idea is formally similar to the Lazear (1979) model of deferred compensation described above, where workers value the future rents associated with acceptable behavior today when making their effort decisions. As with Lazear's work, a critical component of both models is that rents must be earned in the static agency setting; otherwise there is no need for dynamic links between contracts.[76]

[76] Perhaps the most striking observation in Rosen (1986) is that in an elimination tournament with risk-neutral workers, the optimal wage structure consists of constant prizes for "promotion" until one reaches the last rank, at which point there is a discretely higher prize. This final prize is often interpreted as necessary to give incentives at the end of the competition, since there are no longer incentives generated by possible future

3.2 *Dynamic Renegotiation of Contracts: Career Concerns*

Performance measures reveal information not only on the efforts exerted by an individual but also on his innate ability. Career concerns arise in situations where agents exert effort not just to maximize current pay but also to affect the perceptions of others. Consider the case of a baseball player on a fixed salary. The analysis of the previous sections would suggest that the player would have little incentive to exert effort, as there is no immediate relationship between pay and performance. This, of course, is patently false, since players exert effort in order to affect future contracts, which depend on current performance. In essence, contracts can be renegotiated on the basis of performance as the market settles up. Following Eugene Fama (1980) and Holmstrom (1982), such career concerns have been proposed as a means of

prizes. However, this wage distribution can be seen as the outcome of the optimal static tournament being repeated, with no inter-rank links in the design of the optimal tournament. To see this, assume that the winner of the optimal static tournament gets a prize of P. Then with reservation utilities normalized to zero, the winner of any given round has a net utility increase of $\frac{P}{2}$ and the loser has a net utility change of $-\frac{P}{2} < 0$. Then consider the wage distribution that arises from an elimination tournament, i.e., where one loss excludes one from future tournaments. Consider an agent who loses in the first round; his net utility is $-\frac{P}{2}$. For a loser in the second round, net utility is 0, while a loser in the third round earns net utility of $\frac{P}{2}$, and so on. Thus, the net utility from winning successive rounds is linear, with the gain being $\frac{P}{2}$. But this is not true for the final round, because the winner of the final must win *two* more net tournaments than the loser of the final. As a result, the final jump in utility in the elimination tournament is P, not $\frac{P}{2}$, for the reason that the ultimate winner does not finish the contest with a loss, unlike all other agents. Therefore, the utility (and wage) distribution is linear in rank until the final level, at which point it "jumps." Hence the distribution of wages generated in Rosen could be seen as arising from the arithmetic of repeated static tournaments rather than the generation of an option in repeated tournaments.

inducing efficient effort exertion even in the absence of explicit contracts. One attraction of this literature is that it places the agency issues more firmly in the context of a labor market that values employees and affects the behavior of the firm. From the career concerns perspective, because outside options matter (in the sense that other firms will bid for workers), incentives are provided even in circumstances where explicit pay-for-performance is not offered.

The following changes are made to the basic model outlined above along the lines of Holmstrom (1982) and Gibbons and Murphy (1992). First, assume that the agent works for two periods, $t = 1,2$, rather than the single period of Section 2. Further, assume that the worker gains from being perceived as talented. In particular, let the labor market be competitive, where the worker earns his expected productivity in period 2. The worker is assumed to be evaluated on a common subjective (i.e., non-contractible) signal of $s_t = e_t + \alpha + \varepsilon_{st}$ in period $t, t = 1, 2,$ where all variables are distributed as in Section 2.1 and where the time-subscripted error terms are independently distributed across the two periods. Some of the implication of the career concerns model is on observed contracts. As a result, it is also assumed that the firms can base compensation on a measure $\tilde{s}_t = s_t + b_t,$ where b_t refers to bias activities exerted in period $t,$ which have the same costs $K(b_t) = \frac{\kappa b_t^2}{2}$ as above.

Consider the second period, where the worker receives a contract $w_2 = \beta_{02} + \beta_{\tilde{s}2}\tilde{s}_2.$ Then by analysis identical to that above, the optimal choice of $\beta_{\tilde{s}2}$ is given by $\beta_{\tilde{s}2}^* = \dfrac{1}{1 + \dfrac{c}{\kappa}},$ as in the static setting. Importantly, since wages equal expected productivity, the salary component is

$\beta_{02}^* = (1 - \beta_{\tilde{s}2}^*)[e_2^* + E(\alpha \mid y_1)],$ where e_2^* is the equilibrium level of effort in period 2 and $E(\alpha \mid y_1)$ is the perceived level of ability of the worker.[77] It is through the second-period salary that career concerns arise; the reservation utility of the worker depends on first-period performance. To understand how this affects incentives, note that

$$E(\alpha \mid y_1) = \frac{\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_s^2} (y_1 - e_1^*), \quad (22)$$

where e_1^* is the expected level of first-period effort (thus, the market is not fooled in equilibrium). As in the previous section, let δ be the discount factor between periods 1 and 2. Then for any first period contract, $w_1 = \beta_{01} + \beta_{\tilde{s}1}\tilde{s}_1,$ the agent will exert effort of

$$e_1^* = \frac{\beta_{\tilde{s}1} + \delta(1 - \beta_{\tilde{s}2}^*)\dfrac{\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_s^2}}{c} \quad (23)$$

Consequently, for $\beta_{\tilde{s}2}^* < 1,$ period 1 incentives are greater than in the static setting, because future contracts depend on perceptions. Even in cases where there are no explicit contracts ($\kappa = 0$), the agents will exert effort of $\dfrac{\delta\sigma_\alpha^2}{c(\sigma_\alpha^2 + \sigma_s^2)}$ solely to affect perceptions. In addition however, there are implications for observed compensation contracts. In particular, straightforward calculations show that the optimal choice of $\beta_{\tilde{s}1}^*$ implies less explicit incentives in period one than in the second period, i.e., $\beta_{\tilde{s}1}^* < \beta_{\tilde{s}2}^*,$ which has been tested.

This simple model offers a number of implications of career concerns.[78] First, agents will exert positive levels of effort

[77] The worker earns $\beta_{\tilde{s}2}^*(e_2^* + E(\alpha|y_1))$ in expectation from the piece rate, so this is the salary at which the firm breaks even.

[78] It should be emphasized that some of these results are specific to the case where the speed of learning is independent of effort exerted. See Mathias Dewatripont, Ian Jewitt, and Jean Tirole (1997) for details.

in the absence of explicit contracts. Second, agents will generally exert inefficient levels of effort. In this 2–period setting, in the absence of contracts, there is always underprovision of effort in both periods because the career concerns can never be sufficiently important in the first period as there is only a single period in which to "cash in" from a good reputation. This is not necessarily true in a general T-period model, as Holmstrom (1982) has shown that sometimes agents will exert effort above the efficient level to affect perceptions.[79] A third effect of career concerns models is that effort exertion depends on the length of time "on the job." When workers begin in their positions, little is known about them, so that productivity realizations have significant effects on perceptions of ability. In addition, workers who are young have a long time over which to garner the returns to a good reputation, so that this model predicts high effort among the young (or those with low tenure). This is not true after tenure is accumulated, because much is already known about workers, and there is less time over which to generate a reputation. Thus effort levels fall over time.

A corollary of this is that worker welfare will be more sensitive to early realizations of output than those that arrive for workers who have been "on the job" for a long period of time. A testable implication of this is that current performance should be more predictive of rewards for younger workers (or those with less tenure) than for older workers. To my knowledge, the only test of this prediction has been carried out by Chevalier and Ellison (1997b) who con-

sider the likelihood of mutual fund managers being fired on the basis of current and previous performance. As predicted by the theory, this relationship is greater for younger managers, about whom there is little information, than for older managers. Gibbons and Murphy (1992) consider a final implication of the theory alluded to above; that contracts offered to workers will have more explicit pay-for-performance provisions as workers get closer to retirement. Gibbons and Murphy generalize this simple model to illustrate that optimal sensitivity of contracted pay to performance will monotonically increase as workers accumulate tenure. They test this prediction on US executives and find evidence consistent with the prediction. Paul Gompers and Josh Lerner (1994) illustrate similar effects on the contracts offered to venture capital managers.

Before concluding, it is worth making a couple of other observations on the implications of career concern settings. First, many of the comparative statics alluded to in the static setting need no longer hold. For instance, consider the optimal choice of incentives in the static model of Section 2.1, where noisy measures imply reduced incentives. Unlike the static setting with risk aversion, this model predicts that first period incentives are *increasing* in $\sigma_{\bar{s}}^2$ so that incentives are increasing in the noisiness of the measures of performance. This is because as measures of performance become noisier, career concerns fall in the sense that there is little updating on ability. Thus, explicit measures of performance must rise to compensate, and so the basic trade-off between risk and incentive no longer holds.[80] In a similar vein, Meyer and John Vickers (1997)

[79] The reason for this is that in a T-period model, a reputation gained in period s has value over the remaining $T - s \geq 1$ periods, so that career concerns effects can be large enough to induce more than efficient effort.

[80] This effect arises most clearly in this setting when the worker is risk neutral. The effect would be tempered and possibly reversed in a setting where workers are risk averse.

illustrate that relative performance evaluation may not be desirable when career concerns are present. This arises because reducing measurement error through relative performance evaluation, while good in a static agency model, can be harmful in the context of career concerns, because the more that is known about ability, the less reason to exert effort for career concern reasons.

Second, it should not necessarily be assumed that career concerns always increase effort. A simple reinterpretation of the model, following Gibbons (1987) and Meyer and Vickers (1997), formalizes ratchet effects as a career concern problem. Ratchet effects arise when firms react to information that costs of production are lower by reducing the pay of agents. For example, firms could require workers to produce higher quotas when they illustrate that high performance levels are possible. To formalize this, assume that in the model above α now refers to ability in the firm, which has no value outside, so higher ability means higher productivity. In this case the renegotiation of the contracts imply that better agents receive *lower* salaries, since able agents will earn more from any fixed piece rates. As a result, agents now have an incentive to restrict output (to avoid such downward revision in salaries), so that career concerns can harm incentives.[81]

3.3 *Dynamic Enforcement of Contracts*

The literature has emphasized a couple of other ways in which repetition can improve the agency relationship. First, throughout the paper I have stressed the importance of subjective measures of performance. But if these measures cannot be verified to third parties, why would a principal ever honestly reveal these measures? Here the literature has stressed an important role for repeated relationships. Consider a setting where a principal must choose whether to reward a worker for good performance that is unverifiable. Though it may be part of an optimal contract to reward the worker for good performance, in a static setting the principal will generally renege on the contract in order to save on the extra wage costs. However, standard repeated game logic can imply that the principal will compensate the worker in the appropriate way if the worker can threaten to withhold future effort if he fails to do so. In that setting, Clive Bull (1987) and MacLeod and Malcomson (1989) illustrate that with sufficiently high discount factors, repetition can generate efficient outcomes that would not arise in the static setting. See MacLeod (1993) and Malcomson (1998) for surveys of this literature. Baker, Gibbons, and Murphy (1994) extend this logic to show that such implicit contracts interact with explicit contracts in interesting ways, so that the existence of explicit contracts can either reinforce implicit contracts or crowd them out.

The common feature of these models of incentive provision is that firms

---

[81] In this section, only situations where the worker exerted "effort" were considered. However, career concerns have been shown to affect many dimensions of performance. For example, Holmstrom and Ricart I Costa (1986) and Hermalin (1993) consider career concern problems when agents choose the riskiness of the projects they take. Jeremy Stein (1990) and Paul (1992) address how career concern models can induce myopia, where agents care excessively about short-term returns to projects rather than their net present value. Finally, a series of papers, beginning with David Sharfstein and Stein (1992), have addressed how career concerns can induce agents to become either conservative or impulsive. See Prendergast and Lars Stole (1996), and Jeffrey Zweibel (1995) for details of that literature, and Owen Lamont (1996) for empirical work.

sometimes value the future enough to induce them to act honestly, even in situations where there is an incentive to contemporaneously cheat. An implication of this is that when future rents are high, firms will be less likely to renege on their implicit obligations than in situations where the firms expect to gain little in the future from their reputations. Bertrand (1997) empirically considers this by estimating the likelihood of firms reneging on implicit risk-sharing agreements with their workers. She shows that firms that are subject to more competition (or which are in financial distress) are more likely to renege on these wage contracting agreements than those firms that earn higher rents.

A second use for repeated relationships in agency models is that repetition may allow better inferences to be drawn on performance. For example, there is likely to be considerable period-to-period variation in how well a salesforce worker fares, and rewarding the agent on performance in a single period may expose him to considerable risk. However, the principal can often do better to consider performance over a longer period of time, which may improve inferences on whether the agent has shirked in the past or not. This is surely important in reality, and there has been some work on this issue in the literature, where the principal typically uses law-of-large-number arguments to generate better inferences on the performance of the worker, which allows better risk sharing between the principal and the agent. However, these papers typically use limiting argument to generate the benefits of repetition (as in Roy Radner 1985, for example) where the agent and workers interact for a large number of periods. As a result, this literature has generated few empirically testable predictions.

4. *Conclusion*

This survey has covered a wide range of issues associated with compensation and incentives, ranging from the behavior of professional bowlers to chicken farmers to chief executive officers. Let me conclude here by pointing to what I feel have been the major contributions and drawbacks of the theoretical and empirical literature, which could be used to suggest future research directions.

First, from the evidence collected above, it does appear that agents respond to incentives. Groves et al. (1994), Lazear (1996), Paarsch and Scherer (1996), Banker, Lee, and Potter (1996), and Boning, Ichniowski, and Shaw (1998) all point to strong effects of pay-for-performance on output, admittedly in settings where measures of overall performance were available. Foster and Rosenzweig (1994) provide direct measures of effort costs to suggest that piece rates are associated with trying harder. Similarly, Knoeber (1989), Knoeber and Thurman (1994), and Ehrenberg and Bognanno (1990) point to strong effects of prize money on the behavior of individuals in tournaments. Finally, the empirical evidence on teams (such as Gaynor and Pauly 1990) suggests the importance of free riding in teams, though Weiss (1987) and Hansen (1997) suggest effects of peer pressure that are different from those predicted in the theoretical literature. Yet it should not be implied that such responses to incentives are necessarily beneficial. Evidence from Healy (1985), Asch (1990), Brown, Harlow, and Starks (1996), Chevalier and Ellison (1997), and Courty and Marschke (1995) among others, suggests that agents are also capable of actions that are privately beneficial at the cost of overall efficiency.

The empirical evidence has also pointed to significant selection effects of contracts. Lazear (1996) illustrates positive selection from the use of piece rates; better employees prefer pay for performance. Also interesting is that the selection effects appear to be of roughly equal size to the incentive effects, despite the overwhelming focus on incentive effects in the theoretical literature. Finally, Weiss (1987) has illustrated the attraction of team-based pay, not to the worst workers as predicted by simple theory, but to those of medium ability, where the best and worst find the constraints of team production unattractive.

The available evidence suggests that incentives do matter, for better or for worse. It is much less clear, however, whether the theoretical models based on this premise have been validated in the data. The true test of agency theory is not simply that agents respond to incentives, but that the contracts predicted by the theory are confirmed by observed data. Here the literature has been less successful. The literature on the trade-off between risk and incentives has had mixed results. Some authors, such as Kawasaki and McMillan (1987) and Ittner, Larker, and Rajan (1996), find evidence of such a trade-off, while Garen (1993) and Bushman, Indejikian, and Smith (1996) find little. Even in cases where the effects are present, the results are sometimes brittle or explain very little of the variation in observed contracts. Similarly, there is mixed evidence on the importance of relative performance evaluation. This is not to say that these theories are not correct, merely that the jury is still out. It is difficult to know whether the theoretical predictions on subjective contracts stand up to empirical scrutiny, because there is so little literature on how contracts are designed for workers in complex jobs, a point I will return to below. Finally, the section on deferred compensation seems to suggest that firms do indeed overpay older workers at the expense of their younger counterparts. However, in those cases, there are typically other plausible interpretations of the data.

All in all, the available empirical evidence on contracts does not yet provide a ringing endorsement of the theory. This could be because the tests considered are weak, or because the theory is not capturing all the relevant features of compensation contracts. Many of the constraints on the literature have been imposed by data limitations; there are simply no easily accessible databases with personnel data. Seen in this light, it is unsurprising that much of the work on incentives has been on executives, for whom there are publicly available data. In addition, it seems clear that another limitation of the literature has been the fact that contracts are often unobserved, where they must be inferred from the empirical relationship between pay and performance, which is tainted with many confounding effects. This is not meant as a criticism of the literature; the best work is being done with the available data. But it is not surprising that recent successes in estimating the effect of agency contracts consider settings where data on contracts have been observed, and a critical component of future research will surely be the collection of such data.

A second problem that pervades this literature is identification, which comes in two guises. The first is the standard empirical identification problem, where the researchers need to understand why contracts vary across environments. It is not enough to simply compare the productivity of workers on piece rates to those on salaries to estimate the effect of pay on performance. Various

selection problems have been considered in the literature above, and a central task of recent contributions to the literature has been to explicitly model the source of that variation.[82] The second problem of identification occurs at a theoretical level. The typical theoretical paper addresses how a certain institution may be optimal. Comparative statics, when offered, are usually of the form that institutions or contracts are likely to vary with certain parameters. However, almost no theoretical work has distinguished among plausible theories there. For many observed phenomena, a multiplicity of theories are consistent with the facts.[83] Consequently, a second necessary ingredient for future empirical research is that theoretical work be aimed at better distinguishing among theories.[84]

The literature has been successful in providing an important organizing tool for understanding the compensation practices of firms, and the empirical work has cast light on those aspects that appear to be most important. However, a final problem with the literature thus far has been an excessive focus on the contracts of workers for whom output

measures are easily observed. Largely because of data availability, there has been considerable work done on such occupations as chief executive officers, golfers, mutual fund managers, tree cutters, windshield installers, and so on, for whom it is possible to construct objective measures of output. Work on these occupations has provided important insights into how incentives operate and how they translate into contracts. However, to put it simply, most people don't work in jobs like these. Instead, most workers are evaluated on subjective criteria, where firms choose how to evaluate and how to pay based on those evaluations. The literature on personnel and human resources management has long understood that a difficult aspect of compensation is the evaluation of such workers. For instance, how do firms get supervisors to tell the truth about their subordinates? Contracts surely reflect these concerns, yet the economics literature has had relatively little to say, beyond the observations in Section 2.2. I believe that significant progress could be made by empirically understanding how subjective assessments are made. How are decisions made on performance and how do evaluations translate into pay, training, and promotion decisions? With what factors do such decisions vary? This is a difficult exercise, since there is no obvious taxonomy to categorize types of subjective performance evaluation, but I believe it would be useful.

To conclude, agency theory has provided an important framework for understanding compensation issues. Not surprisingly, there has been a lag in testing some of the empirical predictions, though the last couple of years have seen considerable advances made. The available empirical evidence appears to be supportive of the theory, though not resoundingly so in some

[82] See Prendergast (1995) for a discussion of various identification strategies.

[83] For example, Main, O'Reilly, and Wade (1990), among others, have tested for the importance of tournament theory by considering (i) whether wages rise in a convex fashion as one moves up the hierarchy, and (ii) whether the prize for becoming CEO is increasing in the number of contestants for the job. These outcomes have generally been found to be true, which is consistent with tournament theory. However, it is equally true of a hierarchy as in Rosen (1982), where workers are allocated to jobs on the basis of comparative advantage without incentives being relevant. (Wages rise in a convex fashion due to the magnification effect. Wages increase in the number of competitors as the best of N workers is on average better than the best of N-1 workers.) Consequently, it is difficult to determine which theory best predicts the data.

[84] See Gibbons and Waldman (1998) for one approach to this problem.

settings. As mentioned above, the empirical work has been restricted partly through the unavailability of data on contracts, which is being rectified. Second, empirical and theoretical work needs to continue to address the important identification issues that plague the literature. Finally, there is a lot left to learn about the evaluation of workers whose output is hard to see, where objectives and outcomes are determined by superiors. Since this constitutes most of us, this seems a large hole to fill in the literature.

REFERENCES

Abowd, John. 1990. "Does Performance-Based Managerial Compensation Affect Corporate Performance?" *Ind. Lab. Relat. Rev.*, 43:3, pp. s52–74.

Acemoglu, Daron and Andrew Newman. 1997. "The Labor Market and Corporate Structure," mimeo., MIT.

Aggarwal, Raj and Andrew Samwick. Forthcoming. "The Other Side of the Tradeoff: The Impact of Risk on Executive Compensation," *J. Polit. Econ.*

Aghion, Philippe and Jean Tirole. 1997. "Formal and Real Authority in Organizations," *J. Polit. Econ.*, 105, pp. 1–29.

Akerlof, George. 1976. "The Economics of Caste and the Rat Race and Other Woeful Tales," *Quart. J. Econ.*, 90, pp. 599–617.

Akerlof, George and Lawrence Katz. 1989. "Workers Trust Funds and the Logic of Wage Profiles," *Quart. J. Econ.*, 104, pp. 525–36.

Allen, Franklin and Douglas Gale. 1992. "Measurement Distortion and Missing Contingencies in Optimal Contracts," *Econ. Theory*, 2:1, pp. 1–26.

Alston, Lee. 1981. "Tenure Choice in Southern Agriculture," *Explor. Econ. Hist.*, 18, pp. 211–32.

Alston, Lee and Robert Higgs. 1982. "Contractual Mix in Southern Agriculture since the Civil War: Facts, Hypotheses and Tests," *J. Econ. Hist.*, 42:2, pp. 327–53.

Anderson, Eric and David Schmittlein. 1984. "Integration of the Sales Force: An Empirical Examination," *RAND J. Econ.*, 15, pp. 385–95.

Anderson, Kathryn; Richard Burkhauser, and Jennie Raymond. 1993. "The Effect of Creaming on Placement Rates Under the Job Training Partnership Act," *Ind. Lab. Relat. Rev.*, 46, pp. 613–25.

Antle, Richard and Abbie Smith. 1986. "An Empirical Investigation of the Relative Performance Evaluation of Corporate Executives," *J. Acc. Res.*, 24. pp. 1–39.

Asanuma, Banri and Tatsuya Kikutani. 1992. "Risk Absorption in Japanese Subcontracting: A Microeconomic Study of the Automobile Industry," *J. Japanese Int. Econ.*, 6, pp. 1–29.

Asch, Beth. 1990. "Do Incentives Matter? The Case of Navy Recruiters," *Ind. Lab. Relat. Rev.*, 43, pp. 89–107.

Athey, Susan and Scott Stern. 1997. "An Empirical Framework for Testing Theories About Complementarity in Organizational Design," mimeo., MIT.

Bailey, Richard M. 1970. "Economies of Scale in Medical Practice," in *Empirical Studies in Health Economics*. H. Klarman, ed. Baltimore: Johns Hopkins Press.

Baker, George. 1992. "Incentive Contracts and Performance Measurement," *J. Polit. Econ.*, 100, pp. 598–614.

Baker, George; Robert Gibbons, and Kevin J. Murphy. 1994. "Subjective Performance Measures in Optimal Incentive Contracts," *Quart. J. Econ.*, 109, pp. 1125–56.

Baker, George; Michael Gibbs, and Bengt Holmstrom. 1994a. "The Wage Policy of a Firm," *Quart. J. Econ.*, 109:4, pp. 921–55.

———. 1994b. "The Internal Economics of a Firm: Evidence from Personnel Data," *Quart. J. Econ.*, 109:4 pp. 881–919.

Baker, George and Bengt Holmstrom. 1995. "Internal Labor Markets: Too Many Theories, Too Few Facts," *Amer. Econ. Rev.*, 85:2, pp. 255–59.

Baker, George; Michael Jensen, and Kevin J. Murphy. 1988. "Compensation and Incentives: Practice vs. Theory," *J. Finance*, 43, pp. 593–616.

Banker, Rajiv; Seok-Young Lee, and Gordon Potter. 1996. "A Field Study of the Impact of a Performance-Based Incentive Plan," *J. Acc. Econ.*, 21, pp. 195–226.

Barberis, Nicholas; Maxim Boycko, and Andrei Shleifer. 1996. "How Does Privatization Work? Evidence from Russian Shops," *J. Polit. Econ.*, 104, pp. 764–91.

Barro, Jason and Robert Barro. 1990. "Pay, Performance and the Turnover of Bank CEOs," *J. Labor Econ.*, 8, pp. 448–81.

Becker, Brian and Mark Huselid. 1992. "The Incentive Effects of Tournament Compensation Systems," *Admin. Sci. Quart.*, 37, pp. 336–50.

Bernhardt, Dan. 1995. "Strategic Promotion and Compensation," *Rev. Econ. Stud.*, 62, pp. 315–39.

Bertrand, Marianne. 1997. "From the Invisible Handshake to the Invisible Hand? How Product Market Competition Changes the Employment Relationship," mimeo., Harvard U.

Bertrand, Marianne and Sendhil Mullainathan. 1997. "Executive Compensation and Takeover Legislation," mimeo., Harvard U.

Bishop, John. 1987. "The Recognition and Reward of Employee Performance," *J. Lab. Econ.*, 5, pp. 36–56.

Bjerke, D.; J. Cleveland, R. Morrison, and W. Wil-

son. 1987. "Officer Fitness Report Evaluation Study," Report TR88–4, Navy Personnel Research and Development Center.

Blackwell, David; James Brickley, and Michael Weisbach. 1994. "Incentives of Divisional Managers: Evidence of Texas Banks," *J. Acc. Econ.*, 17, pp. 331–58.

Boning, Brett; Casey Ichniowski, and Kathryn Shaw. 1998. "Incentive Pay for Production Workers: An Empirical Assessment," mimeo., Carnegie Mellon.

Bretz, R. and George Milkovich. 1989. "Performance Appraisals in Large Organizations: Practice and Research Implications," mimeo., Cornell U.

Brown, Charles. 1990. "Firms Choice of Method of Pay," *Ind. Lab. Relat. Rev.*, 43:3, pp. 165–82.

Brown, Keith, W.V. Harlow, and Laura Starks. 1996. "Of Tournaments and Temptations: An Analysis of Managerial Incentives in the Mutual Fund Industry," *J. Finance*, 51, pp. 85–110.

Bull, Clive. 1987. "The Existence of Self-Enforcing Wage Contracts," *Quart. J. Econ.,* 102, pp. 147–59.

Bushman, Robert; Raffi Indejikian, and Abbie Smith. 1996. "CEO Compensation: The Role of Individual Performance Evaluation," *J. Acc. Econ.*, 21, pp. 161–93.

Camerer, Colin and Robert Hogarth. 1997. "The Effects of Financial Incentives in Experiments: A Review and Capital-Labor-Production Theory," mimeo., U. Chicago.

Cappelli, Peter and Keith Chauvin. 1991. "An Interplant Test of the Efficiency Wage Hypothesis," *Quart. J. Econ.*, 106, pp. 769–87.

Carmichael, Lorne. 1983. "Firm Specific Human Capital and Promotion Ladders," *Bell J. Econ.*, 14, pp. 251–58.

————. 1988. "Incentives in Academics: Why is There Tenure?" *J. Polit. Econ.*, 96, pp. 453–72.

Cheatham, Carole; Dorothy Davis, and Leo Cheatham. 1996. "Hollywood Profits: Gone With the Wind?" *CPA J.*, pp. 32–34.

Chevalier, Judith and Glenn Ellison. 1997a. "Risk Taking by Mutual Funds as a Response to Incentives," *J. Polit. Econ.*, 105:6, pp. 1167–200.

————. 1997b. "Are Some Mutual Fund Managers Better than Others? Cross-Sectional Patterns in Behavior and Performance," mimeo., U. Chicago.

Conyon, Martin and Simon Peck. 1997. "Corporate Tournaments and Executive Compensation: UK Evidence," mimeo., U. Warwick.

Coughlin, Ann and Chakravarthi Narasimhan. 1992. "An Empirical Analysis of Sales-Force Compensation Plans," *J. Bus.*, 65, pp. 93–123.

Coughlin, Ann and Ronald Schmidt. 1985. "Executive Compensation, Management Turnover and Firm Performance," *J. Acc. Econ.*, 7, pp. 43–66.

Courty, Pascal and Jerry Marschke. 1995. "Moral Hazard Under Incentive Systems: The Case of a Federal Bureaucracy," mimeo., U. Chicago.

————. 1997. "Measuring Government Performance: Lessons from a Federal Job-Training Program," *Amer. Econ. Rev.*, 87, pp. 383–88.

Deci, Edward. 1971. "The Effects of Externally Mediated Rewards on Intrinsic Motivation," *J. Personality Soc. Psych.*, 18, pp. 105–15.

Demougin, Dominique and Aloysious Siow. 1994. "Careers in Ongoing Hierarchies," *Amer. Econ. Rev.*, 84, pp. 1261–77.

Dewatripont, Mathias; Ian Jewitt, and Jean Tirole. 1997. "Information Structures and Career Concerns," mimeo., ECARE.

Drago, Robert and Gerald Garvey. 1998. "Incentives for Helping on the Job: Theory and Evidence," *J. Labor Econ.*, 16:1, pp. 1–25.

Ehrenberg, Ron and Michael Bognanno. 1990. "The Incentive Effects of Tournaments Revisited: Evidence from the European PGA Tour," *Ind. Lab. Relat. Rev.*, 43, pp. 74–89.

Encinosa, William; Martin Gaynor, and Jim Rebitzer. 1997. "The Sociology of Groups and the Economics of Incentives," NBER Working Paper 5933.

Erikkson, Tor. 1996. "Executive Compensation and Tournament Theory: Empirical Tests on Danish Data," mimeo., Aarhus U.

Fairburn, James and James Malcomson. 1996. "Performance, Promotion, and the Peter Principle," mimeo., U. Sussex.

Fama, Eugene. 1980. "Agency Problems and the Theory of the Firm," *J. Polit. Econ.*, 88, pp. 288–307.

Fee, Edward and Charles Hadlock. 1997. "Management Turnover and Product Market Competition: Empirical Evidence from the US Newspaper Industry," mimeo., U. Florida.

Ferguson, Leonard. 1949. "The Value of Acquaintance Ratings in Criterion Research," *Personnel Psych.*, 2, pp. 93–102.

Fernie, Sue and David Metcalf. 1996. "It's Not What You Pay It's the Way That You Pay It And That's What Gets Results: Jockeys Pay and Performance," London School Econ. Discus. Paper 295.

Ferrall, Chris and Bruce Shearer. Forthcoming. "Incentives and Transactions Costs Within the Firm: Estimating an Agency Model Using Payroll Records," *Rev. Econ. Stud.*

Ferrall, Chris and Anthony Smith. 1997. "A Sequential Game Model of Sports Championship Series: Theory and Estimation," mimeo., Queens U.

Foster, Andrew and Mark Rosenzweig. 1994. "A Test for Moral Hazard in the Labor Market: Contractual Arrangements, Effort, and Health," *Rev. Econ. Statist.*, 76, pp. 213–27.

Freeman, Richard and James Medoff. 1984. *What Do Unions Do?* New York: Basic Books.

Friebel, Guido and Michael Raith. 1997. "Strategic Recruiting and Chains of Command," mimeo., U. Chicago.

Garen, John. 1994. "Executive Compensation and Principal-Agent Theory," *J. Polit. Econ.*, 102, pp. 1175–99.

Gaynor, Martin and M. Pauly. 1990. "Compensa-

tion and Productive Efficiency in Partnerships; Evidence from Medical Group Practice," *J. Polit. Econ.*, 98, pp. 544–74.

Gibbons, Robert. 1987. "Piece-Rate Incentive Schemes," *J. Labor Econ.*, 5, pp. 413–29.

———. 1996. "Incentives and Careers in Organizations," in *Advances in Economics and Econometrics: Theory and Applications*. David Kreps and Ken Wallis, eds. Cambridge: Cambridge U. Press.

Gibbons, Robert and Kevin J. Murphy. 1990. "Relative Performance Evaluation for Chief Executive Officers," *Ind. Lab. Relat. Rev.*, 43, pp. 30–52.

———. 1992. "Optimal Incentive Contracts in the Presence of Career Concerns," *J. Polit. Econ.*, 100, pp. 468–506.

Gibbons, Robert and Michael Waldman. 1998. "A Theory of Wages and Promotion Dynamics Inside a Firm," mimeo., Cornell U.

———. Forthcoming. "Careers in Organizations: Theory and Evidence," in *Handbook of Labor Economics*.

Gibbs, Michael. 1995. "Incentive Compensation in a Corporate Hierarchy," *J. Acc. Econ.*, 19, pp. 247–77.

Gibbs, Michael and Wallace Hendricks. 1996. "Are Administrative Pay Systems a Veil?" mimeo., U. Illinois.

Gompers, Paul and Josh Lerner. 1994. "An Analysis of Compensation in the US Venture Capital Partnership," mimeo., Harvard Business School.

Green, Jerry and Nancy Stokey. 1983. "A Comparison of Tournaments and Contracts," *J. Polit. Econ.*, 91, pp. 349–65.

Groshen, Erica and Alan Krueger. 1990. "The Structure of Supervision and Pay in Hospitals," *Ind. Lab. Relat. Rev.*, 43, pp. S134–47.

Grossman, Sanford and Oliver Hart. 1983. "An Analysis of the Principal-Agent Problem," *Econometrica*, 51, pp. 7–45.

———. 1986. "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Organization," *J. Polit. Econ.*, 94, pp. 691–719.

Groves, Theodore; Yongmiao Hong, John McMillan, and Barry Naughton. 1994. "Autonomy and Incentives in Chinese State Enterprises," *Quart. J. Econ.*, 109:1, pp. 183–209.

Hall, Brian and Jeff Liebman. 1997. "Are CEOs Really Paid Like Bureaucrats?" NBER Working Paper 6213.

Hansen, Daniel. 1997. "Worker Performance and Group Incentives: A Case Study," *Ind. Lab. Relat. Rev.*, 51, pp. 37–49.

Hart, Oliver and Bengt Holmstrom. 1987. "The Theory of Contracts," in *Advances in Economic Theory*. T. Bewley, ed. Cambridge: Cambridge U. Press.

Hart, Oliver and John Moore. 1989. "Property Rights and the Nature of the Firm," *J. Polit. Econ.*, 98, pp. 1119–58.

Hayes, Rachel and Scott Schaefer. 1997. "Implicit Contracts and the Explanatory Power of Top

Executive Compensation for Future Performance," mimeo., Northwestern U.

Hackman, J. Richard and Greg Oldham. 1980. *Work Redesign*. Reading, MA: Addison Wesley.

Healy, Paul M. 1985. "The Effect of Bonus Schemes on Accounting Decisions," *J. Acc. Econ.*, 7:1–3, pp. 85–107.

Heckman, James; Carolyn Heinrich, and Jeff Smith. 1997. "Assessing the Performance of Performance Standards in Public Bureaucracies," *Amer. Econ. Rev.*, 87, pp. 389–96.

Heckman, James; Jeff Smith, and Chris Taber. 1996. "What do Bureaucrats Do? The Effects of Performance Standards and Bureaucratic Preferences on Acceptance into the JTPA Program," in *Reinventing Government and the Problem of Bureaucracy*. G. Libecap, ed. Stamford, CT: JAI Press.

Hermalin, Ben. 1993. "Managerial Preferences Towards Risky Behavior," *J. Law, Econ., Organ.*, 9, pp. 127–36.

Hermalin, Ben and Michael Weisbach. 1988. "The Determinants of Board Composition," *RAND J. Econ.*, 15, pp. 589–606.

Holmstrom, Bengt. 1979. "Moral Hazard and Observability," *Bell J. Econ.*, 10, pp. 74–91.

———. 1982a. "Managerial Incentive Problems: A Dynamic Perspective," mimeo., Kellogg School of Management.

———. 1982b. "Moral Hazard in Teams," *Bell J. Econ.*, 13, pp. 324–40.

———. 1996. "The Firm as a Subeconomy," mimeo., MIT.

Holmstrom, Bengt and Paul Milgrom. 1990. "Regulating Trade Among Agents," *J. Inst. Theor. Econ.*, 146:1, pp. 85–105.

———. 1991. "Multitask Principal Agent Analyses: Incentive Contracts, Asset Ownership and Job Design," *J. Law, Econ., Organ.*, 7:special issue, pp. 24–52.

Holmstrom, Bengt and Joan Ricart I Costa. 1986. "Managerial Incentives and Capital Management," *Quart. J. Econ.*, 101:4, pp. 835–60.

Holmstrom, Bengt and Jean Tirole. 1989. "The Theory of the Firm," in *Handbook of Industrial Organization*. Vol. 1. Richard Schmalansee and Robert Willig, eds. North-Holland, pp. 61–133.

Hutchens, Robert. 1986. "Delayed Payment Contracts and a Firm's Propensity to Hire Older Workers," *J. Lab. Econ.*, 4, pp. 439–57.

Itoh, Hideshi. 1992. "Incentives to Help in Multi-Agent Situations," *Econometrica*, 59, pp. 611–36.

———. 1993. "Coalitions, Incentives and Risk Sharing," *J. Econ. Theory*, 60, pp. 410–27.

Ittner, Christopher; David Larker, and Madhav Rajan. 1997. "The Choice of Performance Measures in Annual Bonus Contracts," *Account. Rev.*, 72:2, pp. 231–55.

Janakiraman, Surya; Richard Lambert, and David Larker. 1992. "An Empirical Investigation of the Relative Performance Evaluation Hypothesis," *J. Acc. Res.*, 30, pp. 53–69.

Jensen, Michael and Kevin J. Murphy. 1990. "Per-

formance Pay and Top Management Incentives," *J. Polit. Econ.*, 98:2, pp. 225–64.

John, George and Bart Weitz. 1989. "Salesforce Compensation: An Empirical Investigation of Factors Related to the Use of Salary versus Incentive Compensation," *J. Market. Res.*, 26, pp. 1–14.

Jones, Derek and Takeo Kato. 1995. "The Productivity Effects of Employee Stock-Ownership Plans and Bonuses," *Amer. Econ. Rev.*, 85:3, pp. 391–415.

Kahn, Charles and Gur Huberman. 1988. "Two-Sided Uncertainty and Up-or-Out Contracts," *J. Lab. Econ.*, 6, pp. 423–44.

Kahn, Lawrence and Peter Sherer. 1990. "Contingent Pay and Managerial Performance," *Ind. Lab. Relat. Rev.*, 43, pp. 107–21.

Kandel, Eugene and Edward Lazear. 1992. "Peer Pressure and Partnerships," *J. Polit. Econ.*, 100:4, pp. 801–17.

Kaplan, Steven. 1994. "Top Executive Rewards and Firm Performance: A Comparison of Japan and the United States," *J. Polit. Econ.*, 102:3, pp. 510–46.

Kawasaki, Seiichi and John McMillan. 1987. "The Design of Contracts: Evidence from Japanese Subcontracting," *J. Japanese Int. Econ.*, 1:3, pp. 327–49.

Knez, Marc and Duncan Simester. 1992. "Firm Wide Incentives and Mutual Monitoring at Continental Airlines," mimeo., U. Chicago.

Knoeber, Charles. 1989. "A Real Game of Chicken: Contracts, Tournaments and the Production of Broilers," *J. Law, Econ., Organ.*, 5:2, pp. 271–92.

Knoeber, Charles and Walter Thurman. 1994. "Testing the Theory of Tournaments: An Empirical Analysis of Broiler Production," *J. Lab. Econ.*, 12:2, pp. 155–79.

Kotlikoff, Lawrence and Jagadeesh Gokhale. 1992. "Estimating a Firm's Age-Productivity Profile Using the Present Value of a Worker's Earnings," *Quart. J. Econ.*, 107, pp. 1215–42.

Kreps, David. 1997. "Intrinsic Motivation and Extrinsic Incentives," *Amer. Econ. Rev.*, 87, pp. 359–65.

Kruse, Douglas. 1993. "Does Profit Sharing Affect Productivity?" mimeo., Rutgers U.

Laffont, Jean-Jacques and Mohamed Matoussi. 1995. "Moral Hazard, Financial Constraints, and Sharecropping in El Oulja," *Rev. Econ. Stud.*, 62, pp. 381–99.

Lafontaine, Francine. 1992. "Agency Theory and Franchising: Some Empirical Results," *RAND J. Econ.*, 23:2, pp. 263–83.

Lambert, Richard and David Larker. 1987. "An Analysis of the Use of Accounting and Market Measures of Performance in Executive Compensation Contracts," *J. Acc. Res.*, 25, pp. 85–125.

Lambert, Richard, David Larker, and Keith Weigelt. 1993. "The Structure of Organizational Incentives," *Admin. Sci. Quart.*, 38, pp. 438–61.

Lamont, Owen. 1995. "Macroeconomic Forecasts and Microeconomic Forecasters," NBER Working Paper 5284.

Landers, Renee, James Rebitzer, and Lowell Taylor. 1996. "Rat Race Redux: Adverse Selection in the Determination of Work Hours in Law Firms," *Amer. Econ. Rev.*, 86:3, pp. 329–49.

Landy, Frank and J. Farr. 1980. "Performance Rating," *Psych. Bul.*, 87, pp. 72–107.

Larkey, Patrick and Jonathon Caulkins. 1992. "All Above Average," mimeo., Carnegie Mellon U.

Lazear, Edward. 1981. "Agency, Earnings Profiles, Productivity, and Hours Restrictions," *Amer. Econ. Rev.*, 71:4, pp. 606–20.

———. 1986. "Salaries and Piece Rates," *J. Bus.*, 59:3, pp. 405–31.

———. 1989. "Pay Equality and Industrial Politics," *J. Polit. Econ.*, 97:3, pp. 561–80.

———. 1995. *Personnel Economics*. Cambridge and London: MIT Press.

———. 1996. "Performance Pay and Productivity," NBER Working Paper 5672.

Lazear, Edward and Robert Moore. 1984. "Incentives, Productivity and Labor Contracts," *Quart. J. Econ.*, 99:2, pp. 275–96.

Lazear, Edward and Sherwin Rosen. 1981. "Rank Order Tournaments as Optimal Labor Contracts," *J. Polit. Econ.*, 89:5, pp. 841–64.

Leibowitz, Arleen and Robert Tollison. 1980. "Free Riding, Shirking and Team Production in Legal Partnerships," *Econ. Inquiry*, 18, pp. 380–94.

Leonard, Jonathon. 1987. "Carrots and Sticks: Pay, Supervision, and Turnover," *J. Labor Econ.*, 5, pp. s136–52.

Lepper, Mark; David Greene, and Richard Nisbett. 1973. "Undermining Children's Intrinsic Interest with Extrinsic Reward," *J. Personality Social. Psych.*, 28, pp. 129–37.

Leventis, Andrew. 1997a. "Cardiac Surgeons Under the Knife," mimeo., Princeton U.

———. 1997b. "The Relation Between Surgeon Age and Outcomes," mimeo., Princeton U.

Loewenstein, George and Nachum Sicherman. 1991. "Do Workers Prefer Increasing Wage Profiles?" *J. Labor Econ.*, 9, pp. 67–84.

MacLeod, Bentley. 1994. "Incentives in Organizations: An Overview of Some of the Evidence and Theory," in *Trends in Business Organization*. Horst Siebert, ed. Kiel: J.C.B. Mohr.

MacLeod, Bentley and James Malcomson. 1989. "Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment," *Econometrica*, 57, pp. 447–80.

MacLeod, Bentley and Daniel Parent. 1998. "Job Characteristics and the Form of Compensation," mimeo., USC.

Main, Brian; Charles O'Reilly, and James Wade. 1993. "Top Executive Pay: Tournament or Teamwork," *J. Labor Econ.*, 11, pp. 606–28.

Malcomson, James. Forthcoming. "New Developments in the Study of Contracts in Labor Markets," in *Handbook of Labor Economics*.

Marschke, Jerry. 1996. "Incentives in a Government Bureaucracy," mimeo., U. Chicago.

McLaughlin, Ken. 1994. "Team-Based Compensation Schemes," mimeo., Hunter College.

McMillan, John; John Whalley, and Lijing Zhu. 1989. "The Impact of China's Economic Reforms on Agricultural Productivity Growth," *J. Polit. Econ.*, 97:4, pp. 781–807.

Medoff, James and Katherine Abraham. 1980. "Experience, Performance, and Earnings," *Quart. J. Econ.*, 95:4, pp. 703–36.

———. 1981. "Are Those Paid More Really More Productive: The Case of Experience," *J. Human Res.*, 16:2, pp. 186–216.

Meyer, Margaret. 1992. "Biased Contests and Moral Hazard: Implications for Career Profiles," *Annales d'Econ. Statist.*, 25–26, pp. 165–87.

Meyer, Margaret; Paul Milgrom, and John Roberts. 1992. "Organizational Prospects, Influence Costs, and Ownership Changes," *J. Econ. Manage. Strat.*, 1:1, pp. 9–35.

Meyer, Margaret and John Vickers. 1997. "Performance Comparisons and Dynamic Incentives," *J. Polit. Econ.*, 105:3, pp. 547–81.

Milgrom, Paul. 1988. "Employment Contracts, Influence Activity and Efficient Organization," *J. Polit. Econ.*, 96, pp. 42–60.

Milgrom, Paul and John Roberts. 1988. "An Economic Approach to Influence in Organizations," *Amer. J. Soc.*, 94, pp. s154–79.

Milkovich, George and Alexandra Wigdor. 1983. "Motivation and Performance Appraisal Behavior," in *Performance Measurement and Theory*, Frank Landy, S. Zedeck and J. Cleveland, eds. Hillsdale, N.J.: Erlbaum.

Mohrman, A. and Edward Lawler. 1991. *Pay for Performance: Evaluating Performance Appraisal and Merit Pay*. Washington: National Academy Press.

Mookherjee, Dilip. 1984. "Optimal Incentive Schemes With Many Agents," *Rev. Econ. Stud.*, 51, pp. 433–46.

Murphy, Kevin J. 1986. "Incentives, Learning and Compensation: A Theoretical and Empirical Investigation of Managerial Labor Contracts," *Rand J. Econ.*, 17, pp. 59–76.

———. 1998. "Performance Standards in Incentive Contracts," mimeo., USC.

———. Forthcoming. "Pay, Performance, and Executive Compensation," in *Handbook of Labor Economics*.

Murphy, Kevin R. and Jeannette Cleveland. 1991. *Performance Appraisal: An Organizational Perspective*. Boston: Allyn and Bacon.

Nalebuff, Barry and Joseph Stiglitz. 1984. "Prizes and Incentives: Towards a General Theory of Compensation and Competition," *Bell J. Econ.*, 2, pp. 21–43.

Newhouse, Joseph. 1973. "The Economics of Group Practice," *J. Human Res.*, 8, pp. 37–56.

O'Reilly, Charles; Brian Main, and Graef Crystal. 1988. "CEO Compensation as Tournament and Social Comparison: A Tale of Two Theories," *Admin. Sci. Quart.*, 33, pp. 257–74.

Oyer, Paul. 1998. "The Effect of Sales Incentives on Business Seasonality," *Quart. J. Econ.*, 113:1, pp. 149–86.

Paarsch, Harry and Bruce Shearer. 1996. "Fixed Wages, Piece Rates, and Incentive Effects," mimeo., U. Laval.

Parent, Daniel. 1996. "Methods of Pay and Earnings: A Longitudinal Analysis," mimeo., U. Sherbrooke.

Paul, Jonathon. 1992. "On the Efficiency of Stock-Based Compensation," *Rev. Fin. Stud.*, 5, pp. 471–502.

Pencavel, John. 1977. "Work Effort, On-the-Job Screening, and Alternative Methods of Remuneration," in *Research in Labor Econonomics*, Vol. 1, R. G. Ehrenberg, ed. JAI Press, pp. 225–58.

Petersen, Trond. 1992. "Payment Systems and the Structure of Inequality: Conceptual Issues and an Analysis of Salespersons in Department Stores," *Amer. J. Soc.*, 98, pp. 67–104.

Prendergast, Canice. 1993a. "The Role of Promotion in Inducing Specific Human Capital Acquisition," *Quart. J. Econ.*, 108:2, pp. 523–34.

———. 1993b. "A Theory of Yes Men," *Amer. Econ. Rev.*, 83:4, pp. 757–70.

———. Forthcoming. "What Happens Within Firms?" in *Labor Statistics and Measurement.* John Haltiwanger, Marilyn Manser, and Robert Topel, eds. NBER; Chicago: U. Chicago Press.

Prendergast, Canice and Lars Stole. 1996. "Impetuous Youngsters and Jaded Oldtimers: Acquiring a Reputation for Learning," *J. Polit. Econ.*, 104:6, pp. 1105–34.

Prendergast, Canice and Robert Topel. 1996. "Favoritism in Organizations," *J. Polit. Econ.*, 104:5, pp. 958–78.

Putterman, Louis. 1990. "Effort, Productivity, and Incentives in a 1970s Chinese Peoples Commune," *J. Compar. Econ.*, 14, pp. 88–104.

Radner, Roy. 1985. "Repeated Principal-Agent Games With Discounting," *Econometrica*, 53, pp. 1173–98.

Raff, Daniel. 1988. "Wage Determination and the Five-Dollar Day at Ford," *J. Econ. Hist.*, 48, pp. 387–99.

Rob, Raffi and Peter Zemsky. 1997. "Cooperation, Corporate Culture and Incentive Intensity," mimeo., U. Pennsylvania.

Rosen, Sherwin. 1982. "Authority, Control, and the Distribution of Earnings," *Bell J. Econ.*, 13, pp. 311–23.

———. 1986. "Prizes and Incentives in Elimination Tournaments," *Amer. Econ. Rev.*, 76, pp. 921–39.

Rothe, H.F. 1949. "The Relation of Merit Ratings to Length of Service," *Personnel Psych.*, 2, pp. 237–42.

Rundquist, E.A. and R. H. Bittner. 1948. "Using Ratings to Validate Personnel Instruments: A Study in Method," *Personnel Psych.*, 1, pp. 163–84.

Salop, Steven and Joanne Salop. 1976. "Self-Selection and Turnover in the Labor Market," *Quart. J. Econ.*, 90:4, pp. 619–27.

Sattinger, Michael. 1993. "Assignment Models and the Distribution of Earnings," *J. Econ. Lit.*, 31, pp. 831–80.

Scharfstein, David and Jeremy Stein. 1992. "Herd Behavior and Investment," *Amer. Econ. Rev.*, 80, pp. 465–79.

Shaefer, Scott. 1994. "The Dependence of CEO Pay-Performance Sensitivity on the Value of the Firm," mimeo., Stanford U.

Shapiro, Carl and Joseph Stiglitz. 1984. "Equilibrium Unemployment as a Discipline Device," *Amer. Econ. Rev.*, 74, pp. 433–44.

Slade, Margaret. 1996. "Multitask Agency and Organizational Form: An Empirical Exploration," *Int. Econ. Rev.,* 37, pp. 465–86.

Spilerman, Seymour. 1986. "Organizational Rules and the Features of Work Careers," *Res. Social Stratification and Mobility*, 5, pp. 41–102.

Spilerman, Seymour and Hiroshi Ishida. 1994. "Stratification and Attainment in Large Japanese Firms," mimeo., Columbia U.

Staw, Barry. 1989. "Intrinsic and Extrinsic Motivation," in *Readings in Managerial Psychology*. H. Leavitt, L. Pondy, and D. Boje, eds. Chicago: U. Chicago Press.

Stein, Jeremy. 1989. "Efficient Capital Markets, Inefficient Firms: A Model of Myopic Corporate Behavior," *Quart. J. Econ.*, 104, pp. 655–69.

Stockford, Lee and H.F. Bissell. 1949. "Factors Involved in Establishing a Merit-Rating Scale," *Personnel*, pp. 94–116.

Taylor, Curtis. 1995. "The Economics of Breakdowns, Checkups, and Cures," *J. Polit. Econ.*, 103, pp. 53–74.

Tehranian, Hassan and James Waegelein. 1985. "Market Reaction to Short-Term Executive Compensation Plan Adoption," *J. Acc. Econ.*, 7, pp. 131–44.

Tirole, Jean. 1992. "Collusion and the Theory of Organizations," in *Advances in Economic Theory: Sixth World Congress*, Vol. 2. J. Laffont, ed. Cambridge: Cambridge U. Press.

Waldman, Michael. 1984. "Job Assignments, Signaling and Efficiency," *RAND J. Econ.*, 15, pp. 255–70.

Weisbach, Michael. 1988. "Outside Directors and CEO Turnover," *J. Finan. Econ.*, 20, pp. 431–60.

Weiss, Andrew. 1987. "Incentives and Worker Behavior," in *Information, Incentives and Risk Sharing*. Haig Nalbantian, ed. Towota, N.J.: Rowan and Littlefield.

Zabojnik, Jan. 1996. "Sales Maximization and Specific Human Capital," mimeo., Cornell U.

Zweibel, Jeffrey. 1995. "Corporate Conservatism and Relative Compensation," *J. Polit. Econ.*, 103, pp. 1–25.