Favoritism in Organizations
Author(s): Canice Prendergast and Robert H. Topel
Source: *The Journal of Political Economy*, Vol. 104, No. 5 (Oct., 1996), pp. 958-978
Published by: The University of Chicago Press
Stable URL: http://www.jstor.org/stable/2138948
Accessed: 18/01/2009 18:24

Favoritism in Organizations

Canice Prendergast and Robert H. Topel

University of Chicago and National Bureau of Economic Research

Objective measures of employee performance are rarely available. Instead, firms rely on subjective judgments by supervisors. Subjectivity opens the door to favoritism, where evaluators act on personal preferences toward subordinates to favor some employees over others. Firms must balance the costs of favoritism—arbitrary rewards and less productive job assignments—against supervisors' demands for authority over subordinates. We analyze the conditions under which favoritism is costly to organizations and the effects of favoritism on compensation, the optimal extent of authority, and the use of bureaucratic rules.

Economists rarely address the fact that firms are social institutions, where personal relations among coworkers, bosses, and subordinates constitute an important component of many workers' daily lives. This paper takes a step toward addressing organizations as social entities by considering how favoritism by superiors, based on personal preferences toward subordinates, affects compensation and the structure of organizations. The underlying premise of the paper is that accurate and objective measures of a worker's performance are typically unavailable. Instead performance is gauged from subjective opinions provided by superiors. This subjectivity opens the door to favoritism,

Topel's research was supported by the National Science Foundation and the Sarah Scaife Foundation through the Center for the Study of the Economy and the State at the University of Chicago Graduate School of Business. We thank Gary Becker, Sherwin Rosen, Kevin Murphy, and Lars Stole for helpful discussions; workshop participants at Chicago, Texas A&M, Stanford, Cornell, Northwestern, Wisconsin, Michigan, Tokyo, and the EMRU Labour Economics Conference, University of Lancaster, United Kingdom; and two anonymous referees. Any errors are our responsibility.

[*Journal of Political Economy,* 1996, vol. 104, no. 5]

where evaluators use their power to reward preferred subordinates beyond their true performance.

Viewing compensation from the perspective of favoritism gives rise to insights that do not arise in more standard agency frameworks. Most important for our analysis is that those in authority—whom we call supervisors—value their power to affect subordinates' welfare. Their demands for this power are traded off against the costs of favoritism caused by added arbitrariness of performance evaluations. An additional feature of our analysis is that favoritism depends on the incentives offered to the worker, since stronger incentive pay for workers reduces the accuracy of supervisors' reports. This arises because supervisors distort their evaluations more when their decisions substantially affect subordinates' welfare.[1] Distortion harms the organization because favoritism corrupts the information on which job assignments are made. Then less productive assignments occur. Arbitrariness in performance appraisals also imposes risk on workers, for which they must be compensated.

The harmful effects of favoritism have two implications for the design of rewards. First, incentive pay for workers will be deemphasized in order to constrain favoritism, both because arbitrariness adds noise to monitoring and because this noise is caused by contract choices.[2] Second, favoritism causes firms to use bureaucratic rules in pay and promotion decisions, where information about a worker's performance is not aggregated in the most efficient way ex post. We show that firms place too little weight on supervisor appraisals and other subjective opinions of performance, giving too much weight to noncorruptible measures such as seniority in compensation and promotion decisions.

These results follow when favoritism is harmful to an organization and must be constrained. A central theoretical point of the paper, however, is that favoritism generates value for those who exercise it. In our analysis, supervisors derive utility from exercising bias, which leads naturally to a demand for "power" by those in authority. Contrary to intuition, this demand can mean that favoritism benefits the organization. Then incentives will be enhanced in order to accommodate supervisors' taste for power. As such, a contribution of the paper is that it endogenously derives the benefit from being "boss" and the implications of this for compensation policies used in firms.

[1] Bjerke et al. (1987) document this effect among supervisors in the Navy. Supervisors considered the consequences of giving out high or low ratings, and they distorted evaluations to maximize the likelihood of outcomes they desired.

[2] Recent work on performance evaluation by Milgrom and Roberts (1988, 1990) and Tirole (1992) also finds that supervisor discretion can reduce reliance on incentive pay. Their results are driven by incentives different from those in our analysis, however.

Whether favoritism is harmful or beneficial depends crucially on the existence of distortions in the "market" for favoritism. If the firm can charge supervisors an optimal price for exercising their preferences and if the only cost of favoritism is the risk it imposes on workers, we show that incentives will be set *as though* favoritism did not exist. When distorted performance appraisals also harm placement decisions, the same circumstances lead to *stronger* incentive pay for workers. In this case, the firm uses a form of bureaucracy in which it commits to abiding by supervisors' decisions even when they are known to be wrong. The implication of our analysis, then, is that favoritism can yield low-powered incentives and rules constraining supervisor behavior, but only if significant distortions can be found in the market for favoritism. With this in mind, Section III considers two plausible distortions on this market: (i) supervisor risk aversion and (ii) effort choices by the supervisor. In both cases, we show how these distortions can make favoritism harmful in equilibrium, which implies low-powered incentives for workers.

We close by considering extensions of the model. First, a typical outcome of monitoring supervisors is that they compress their performance evaluations relative to their true observations.[3] We show how compression derives from the optimizing behavior of supervisors and how compression affects the costs of favoritism and the form of compensation. We then consider how residual claimancy can be used to control favoritism. Finally, we show how reward structures vary across groups that are likely to be the object of bias, such as women and minorities.

Section V provides a brief conclusion.

I. A Model of Favoritism

A. *Incentives and Technology*

The firm studied here employs a worker and a supervisor. The supervisor privately observes a nonverifiable measure of a worker's performance, given by

$$y = e + \alpha + \epsilon_s, \tag{1}$$

where e is effort exerted by the worker, α is the worker's talent for the tasks performed in the firm, and ϵ_s is measurement error that adds noise to the supervisor's observation of true performance. We assume that $\epsilon_s \sim N(0, \sigma_s^2)$. The worker's talent is also drawn from a normal distribution, $\alpha \sim N(0, \sigma_\alpha^2)$, and talent (i.e., how well the

[3] Such compression is well documented in the psychology literature (Landy and Farr 1980; Mohrman and Lawler 1983; Murphy and Cleveland 1991).

worker matches to the firm) is unknown to all parties. We assume that ϵ and α are uncorrelated.

Information provided by the supervisor is used for two purposes. The first is to provide incentives for workers to choose e appropriately. The second is to assign workers to their most productive tasks. On the basis of the supervisor's information, workers can be assigned to another task for an unmodeled future period. Note that the marginal product of ability is one in equation (1). The marginal product of talent (α) on the second task is negative one; in all other respects the jobs are similar.[4] Thus workers with $\alpha > 0$ are more productive if they are retained, so that workers with expected ability $\hat{\alpha} > 0$ will be retained in the old job. We assume that α is entirely firm-specific, so that the task change need not affect the worker's compensation. Here α is meant to reflect how well a worker matches to a particular task, for example, whether he is better in a technical or administrative position.

Preferences

The worker has exponential utility

$$v = -\exp\{-r[w - c(e)]\}, \tag{2}$$

where w is the worker's wage, $r > 0$ is the constant rate of absolute risk aversion, and the worker's cost of supplying effort is $c(e)$, with $c' > 0$, $c'' > 0$, $c'(0) = 0$, and $c'(\infty) = \infty$.

To reflect a taste for bias, the supervisor's utility depends on his own pay, w_s, and on the pay of his subordinate, w:[5]

$$v_s = w_s + \eta w. \tag{3}$$

Here η is the intensity of the supervisor's preference for the worker, so favoritism takes the form of positive or negative altruism. We assume that η is unknown to all parties except the supervisor and is learned by the supervisor only after he joins the firm and encounters the worker. The ex ante distribution of η is $N(0, \sigma_\eta^2)$, so greater values of σ_η^2 indicate greater potential bias by a supervisor; η is uncorrelated with α and ϵ.

Equation (3) embodies the agency problem faced by management in dealing with the supervisor and the worker. If the worker's pay is increasing in the supervisor's report, the supervisor can gain by overstating the performance of favored subordinates ($\eta > 0$). In the

[4] Prescott and Visscher (1980) use a similar assignment technology.

[5] None of our qualitative results is changed if v_s depends on the worker's utility rather than wage. Instead, it merely complicates the derivations without adding significant insights.

absence of some cost of biasing his reports, there is no limit on the supervisor's misinformation. So management will monitor the supervisor and penalize favoritism. It does this by obtaining its own private observation on the worker's performance, given by

$$y_f = e + \alpha + \epsilon_f, \tag{4}$$

where $\epsilon_f \sim N(0, \sigma_f^2)$.[6] The firm then compares the supervisor's report, y_s, to its own information and pays the supervisor according to[7]

$$w_s = w_0 - .5\lambda(y_s - y_f)^2. \tag{5}$$

Values of $\lambda > 0$ penalize the supervisor for positive or negative perceived bias. In what follows, λ is the "price" of exercising favoritism.

We close the model by specifying the worker's pay. Since linear compensation is optimal in the case of exponential utility and normally distributed errors,[8] the worker's pay depends linearly on the two pieces of information available to management:

$$w = \tau_0 + \tau_s y_s + \tau_f y_f. \tag{6}$$

Timing occurs as follows: (1) The firm chooses a compensation policy for the worker and the supervisor. (2) The supervisor observes η. (3) The worker exerts effort, e, and performance, y, is privately observed by the supervisor. The firm simultaneously receives its private observation, y_f. (4) The supervisor reports y_s to the firm. (5) The worker and supervisor are paid and the worker is reassigned if $\hat{\alpha} < 0$.

B. *Behavior*

The supervisor is monitored on the basis of how far his report of the worker's performance is from management's private information, so he will try to guess what management knows. If the bias in the supervisor's report is denoted by $b(\eta, y, \lambda)$, the supervisor reports

$$y_s = E(y_f | y, e^*) + b(\eta, y, \lambda), \tag{7}$$

[6] Examples of y_f include noncorruptible correlates of productivity, such as measurable dimensions of output or management's own subjective opinion of an employee's contribution. Ideally, this signal should be untainted by bias and reflect the effort of the worker. In some instances, firms may use observations for which these requirements do not strictly hold. For example, some firms obtain recommendations by many supervisors on a given worker to determine compensation, which although correlated with effort may be tainted by the bias of the other supervisors. By contrast, firms commonly use seniority to reward individuals, which, though not subject to bias, is unlikely to be correlated with effort.

[7] We do not contend that this compensation scheme dominates all others. However, as the purpose of the paper is to provide closed-form solutions to allow comparative statics, we restrict attention to compensation schemes yielding linear marginal costs.

[8] See Holmstrom and Milgrom (1987). This conclusion is valid only if the noise from favoritism is normally distributed, as will be shown below.

where e^* is the equilibrium level of effort that all parties know the worker will choose. The supervisor's best estimate of management's information is the conditional expectation of y_f given the supervisor's information:

$$E(y_f|y, e^*) = \theta y + (1 - \theta)e^*, \tag{8}$$

where $\theta = \sigma_\alpha^2/(\sigma_\alpha^2 + \sigma_s^2)$. Equation (8) illustrates the tendency of supervisors to "compress" the performance evaluations of their subordinates, understating the performance of those who do well and overstating the performance of those who do poorly (Landy and Farr 1980).[9] In Section IV, we discuss the implications of such compression.

The worker is rewarded by

$$\begin{aligned} w &= \tau_0 + \tau_s[\theta y + (1 - \theta)e^* + b] + \tau_f y_f \\ &= \tilde{\tau}_0 + \tilde{\tau}_s y + \tau_f y_f, \end{aligned} \tag{9}$$

where $\tilde{\tau}_0 = \tau_0 + \tau_s(1 - \theta)e^* + \tau_s b$, and $\tilde{\tau}_s = \tau_s\theta$. The arguments that affect b are excluded for notational simplicity. Throughout the paper, we characterize the worker's incentives by $\tilde{\tau}_s$ since it reflects the marginal return to exerting effort. The worker faces a marginal incentive of $\tilde{\tau}_s + \tau_f$ and chooses effort to satisfy

$$\tilde{\tau}_s + \tau_f = c'(e). \tag{10}$$

Given (7), the supervisor's bias in state η is chosen to

$$\max_b \eta\tau_s[\theta y + (1 - \theta)e^* + b] - \left(\frac{\lambda}{2}\right)b^2, \tag{11}$$

so

$$b = \frac{\eta\tau_s}{\lambda}. \tag{12}$$

Equation (12) is central to our analysis. It shows that bias in state η increases with the supervisor's preferences for the worker, η, and with the impact of the supervisor's report, τ_s, on pay. Thus supervisors lie more when their reports have significant implications for

[9] In the special case in which $\sigma_\alpha^2 = 0$, $E(y_f|y, e^*) = e^*$; the supervisor's information is useless. Then the supervisor knows that any heterogeneity of observed performance is generated by noise, and reporting this noise only raises his expected penalty from being monitored. He merely reports what he knows the worker will do, e^*, without revealing any of his own information. Only when $\sigma_\alpha^2 > 0$ does it pay to reveal some of what he knows, because then y provides information about management's opinion of the worker, y_f. This incentive to compress arises not only when favoritism is evident but also in any circumstance in which the supervisor's report is monitored by the observation of another.

worker welfare. This accords with empirical evidence that supervisors are more likely to distort when "money is on the line" (Landy and Farr 1980; Mohrman and Lawler 1983; Murphy and Cleveland 1991). Bias declines with λ, the price of favoritism.

Management's direct tool to offset bias is λ, and (12) might suggest that λ should be set as large as possible in order to minimize the costs of bias. This is incorrect, because the ability to control the fate of subordinates is a form of "power" that supervisors value. As the fine for perceived bias is a transfer from the supervisor to the firm, the surplus to the supervisor from exercising favoritism is $E[\eta\tau_s b] = (\tau_s^2/\lambda)\sigma_\eta^2 = \lambda\sigma_b^2$, where $\sigma_b^2 = \sigma_\eta^2\tau_s^2/\lambda^2$ is the (endogenous) variance of bias. In what follows, σ_b^2 is our technical definition of bias. Therefore, all statements regarding increased or decreased favoritism should be read in terms of this measure. Notice that expected supervisor utility increases with σ_η^2, the supervisor's ex ante taste for power, and is higher when τ_s is large or when λ is small. Supervisors value the authority that comes with being "boss," but this gain is reduced when their authority is constrained by λ.

Section III below derives the optimal choice of λ by the firm.

C. *The Firm's Objective*

The firm's objective is to maximize ex ante surplus for all parties, S, since all individual rationality constraints bind in equilibrium. If $\mu = 1/\theta$, the problem is to

$$\begin{aligned}\max_{\tau_f,\tilde{\tau}_s,\lambda} S = e - c(e) - \frac{r}{2}[\tilde{\tau}_s^2(\sigma_\alpha^2 + \sigma_s^2 + \mu^2\sigma_b^2) + \tau_f^2(\sigma_\alpha^2 + \sigma_f^2) \\ + 2\tilde{\tau}_s\tau_f\sigma_\alpha^2] + \lambda\sigma_b^2 - L(\sigma_b^2)\end{aligned} \tag{13}$$

subject to (10) and (12). This surplus has several components. First, $e - c(e)$ is expected output less the cost of effort, as is common in agency problems. The bracketed term is a risk premium. The term $\lambda\sigma_b^2$ is the surplus for the supervisor from being "boss." Finally, $L(\sigma_b^2) = 2(E\alpha|\alpha > 0)\text{prob}(\alpha > 0|\hat{\alpha} < 0)$ is the cost of assigning the worker to the wrong task, based on an inaccurate estimate of talent.[10]

[10] The term $\hat{\alpha}$ is defined by

$$\hat{\alpha} = \left(\frac{\sigma_s^2 + \mu^2\sigma_b^2}{\sigma_\alpha^2 + \sigma_f^2 + \mu^2\sigma_b^2 + \sigma_s^2}\right)(y_f - e^*) + \left(\frac{\sigma_f^2}{\sigma_\alpha^2 + \sigma_f^2 + \mu^2\sigma_b^2 + \sigma_s^2}\right)(y_s - e^*),$$

where e^* is the equilibrium level of effort. The errors caused by misallocation are then $2E(\alpha|\alpha > 0)\text{prob}(\alpha > 0|\hat{\alpha} < 0) - 2E(\alpha|\alpha < 0)\text{prob}(\alpha < 0|\hat{\alpha} > 0)\text{prob}(\hat{\alpha} > 0)$. But $\text{prob}(\hat{\alpha} > 0) = \text{prob}(\hat{\alpha} < 0) = 1/2$ and $E(\alpha|\alpha > 0) = -E(\alpha|\alpha < 0)$, so this simplifies to L above. A closed-form solution for L is provided in the Appendix.

This cost depends on τ_s because stronger incentive pay causes greater favoritism by supervisors (larger σ_b^2). This point is summarized in the following lemma, whose proof is in the Appendix.

LEMMA 1. $L'(\sigma_b^2) > 0$; greater favoritism leads to less productive job assignments. Furthermore, $L'' < 0$.

For purposes of exposition, it is useful to solve this problem in two steps. First, we derive optimal worker incentives ($\tilde{\tau}_s$ and τ_f) while holding λ fixed. Following this, we consider the optimal choice of λ. In order to clarify the various influences of favoritism on compensation, it is useful to consider the effect of σ_b^2 on total surplus:

$$\frac{dS}{d\sigma_b^2} = -\tfrac{1}{2}(r\mu^2\tilde{\tau}_s^2 - 2\lambda + 2L') = -\tfrac{1}{2}\Omega. \tag{14}$$

If $\Omega > (<)\ 0$, more favoritism reduces (increases) total surplus. Note that Ω has three components: the two marginal costs of favoritism are picked up by the risk imposed ($r\mu^2\tilde{\tau}_s^2/2$) and worker misallocation (L'), and the marginal benefit of power is λ. The implication is that when power is highly valued by supervisors, additional bias can increase the value of the firm. This point proves important in what follows.

II. Optimal Incentives under Favoritism

Straightforward calculations yield the values of $\tilde{\tau}_s$ and τ_f that maximize (13):

$$\tilde{\tau}_s^* = \frac{\sigma_f^2}{[(\sigma_f^2 + \sigma_s^2 + \mu^2\sigma_b^2)(1 + \rho\sigma_\alpha^2) + \rho\sigma_f^2(\sigma_s^2 + \mu^2\sigma_b^2)] + \left\{\frac{\mu^2\sigma_\eta^2}{r\lambda^2}[1 + \rho(\sigma_\alpha^2 + \sigma_f^2)]\Omega\right\}} \tag{15a}$$

and

$$\tau_f^* = \frac{\sigma_s^2 + \mu^2\sigma_b^2 + \dfrac{\mu^2\sigma_\eta^2}{r\lambda^2}\Omega}{[(\sigma_f^2 + \sigma_s^2 + \mu^2\sigma_b^2)(1 + \rho\sigma_\alpha^2) + \rho\sigma_f^2(\sigma_s^2 + \mu^2\sigma_b^2)] + \left\{\frac{\mu^2\sigma_\eta^2}{r\lambda^2}[1 + \rho(\sigma_\alpha^2 + \sigma_f^2)]\Omega\right\}}, \tag{15b}$$

where $\rho = rc''$. These equations illustrate two influences of favoritism on the provision of incentives. First, the terms in the brackets include the variance of bias. This effect is standard and reflects how increased risk affects incentives; $\tilde{\tau}_s^*$ falls with σ_b^2 when $r > 0$. Note that τ_f^* rises with σ_b^2, but not one-for-one with the fall in $\tilde{\tau}_s^*$, so that total incentives $\tilde{\tau}_s^* + \tau_f^*$ fall as σ_b^2 rises. Thus one effect of favoritism is to reduce incentives because of increased risk in evaluations. Second, the term in the braces reflects how the "marginal cost" of bias affects incentives. As mentioned above, Ω measures (twice) the marginal cost of increasing σ_b^2. If $\Omega > 0$, bias harms the organization on the margin, so incentives are weakened. On the other hand, if $\Omega < 0$, the benefits of authority to the supervisor exceed the costs to the worker. Then $\tilde{\tau}_s^*$ should be increased. This is a central point of our analysis.

This distinction also shows that favoritism is not formally the same as noisy monitoring, which also yields weaker incentives. If favoritism were merely the addition of exogenous noise, the terms in the braces would not appear. In our model, favoritism makes noise endogenous in that the compensation scheme itself corrupts the evaluation process.[11] These effects are captured by the terms multiplying Ω in (15a) and (15b).[12]

According to (15), complete decoupling of evaluations and rewards occurs when workers are risk neutral ($r = 0$). Then $\tilde{\tau}_s^* = 0$ and $\tau_f^* = 1$; only the firm's information is used in setting pay, though the supervisor's information is still used for placement. When $r = 0$, maximum weight is given to the noncorruptible performance measure, y_f, even though that signal may be the least accurate in equilibrium. This minimizes bias and maximizes the accuracy of the supervisor's report, which aids placement decisions. Milkovich and Wigdor (1991, p. 109) document this feature in real-world appraisal systems:

> 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 distrib-

[11] This may appear to run counter to the Revelation Principle, which holds that the equilibrium of our game can be replicated by another that involves truth telling by the supervisor. This observation is correct; however, to induce truth telling, the firm must offer the supervisor an incentive to do so. This implies that the supervisor must be allowed to affect the worker's welfare in exactly the way that occurs in our equilibrium. Therefore, the Revelation Principle yields an outcome identical to that outlined in the equilibrium above; it merely allows the supervisor to report truthfully.

[12] Another useful observation here is that the firm can require separate observations by the supervisor for rewards and for task assignment. In our model, there are no returns to task assignment, and so the supervisor will report $\hat{y}$ but continue to lie according to (12) for the "reward evaluation." While this offers an improvement over the outcome described above, it is likely to be fraught with problems. First, in reality, task assignment does involve utility changes for workers so that lying may continue to occur in reality. Second, the firm may be tempted to use the truthful report (on assignment) to reward the worker. For these reasons, we have ignored this possibility throughout the paper. However, if the assignment problem can be solved by this mechanism, all results below hold for the case in which $L' = 0$.

> ute 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.

This description of trade-offs accords exactly with our analysis.

A. *Bureaucracy*

Decision processes in many organizations are bureaucratic, dominated by procedures that systematically ignore relevant information. For instance, with the decoupling just mentioned, supervisors' reports are used for employee "feedback" and to guide placement decisions, but evaluations are given little weight in setting pay (Bretz and Milkovich 1989). Seniority rules are another example. Freeman and Medoff (1984) report that only 14 percent of nonunion firms base layoff decisions on productivity alone, ignoring seniority.[13] In many cases, pay is conditional on seniority alone, with information on productivity ignored (Spilerman 1986).

We define bureaucratic rules as situations in which the ratio of $\tilde{\tau}_s$ to τ_f is different than when information is aggregated optimally. That is, bureaucracy exists if $\tilde{\tau}_s/\tau_f \neq \sigma_f^2/(\sigma_s^2 + \mu^2\sigma_b^2)$, the unique ratio consistent with optimal information aggregation. With endogenous favoritism we have

$$\frac{\tilde{\tau}_s^*}{\tau_f^*} = \frac{\sigma_f^2}{\sigma_s^2 + \mu^2\sigma_b^2 + (\mu^2\sigma_\eta^2/r\lambda^2)\Omega}. \tag{16}$$

If favoritism is harmful on the margin ($\Omega > 0$), the firm commits to a $\tilde{\tau}_s^*$ that is "too low" and a τ_f^* that is "too high." Greater weight is given to the noncorruptible signal in a way that is ex post inefficient. It is worth emphasizing that this does not occur when bias is exogenous, for then the terms multiplying Ω in (16) disappear. Instead, bureaucracy arises solely because the supervisor's incentive to lie depends on the contract offered to the worker.

Our interpretation of bureaucracy does not always imply that su-

[13] Other examples come from the academic market. In many universities, departmental recommendations on pay and promotion must pass through layers of review committees, made up of faculty from other departments. These committees rely on letters from outside experts to monitor departmental recommendations. The economics department of a major American university grants tenure to junior faculty who have published six articles in refereed journals. One article must be in a "top five" journal, but subjective evaluations of the research by senior colleagues are given little weight. In another case, faculty members earn "points" for publishing in certain journals. The point total is the sole criterion for determining pay raises. These pay and promotion practices are alleged responses to earlier cases of favoritism.

pervisor reports are given too little weight. This is the case only if favoritism is costly on the margin ($\Omega > 0$). If $\Omega < 0$, the firm does not intervene appropriately to change a supervisor's recommendation. Instead, the supervisor's report is given "too much" weight because of his thirst for power. By our definition, this is another form of bureaucracy.[14] An additional point is that bureaucratic rules are used only when evaluations can distribute rents. In our model the assignment decision does not involve rents, so the firm aggregates information on α ex post efficiently. Therefore, bureaucracy in our view is intrinsically linked to the existence of rents, where rents are distributed at the discretion of evaluators.

III. Monitoring the Supervisor

The results described above address an indirect solution to the problem of favoritism: the *worker's* incentives are adjusted to remedy an action taken by the *supervisor.* A more direct response by the firm may be to manipulate the supervisor's rewards. The firm can use λ for this purpose, in effect creating a market for favoritism. A striking outcome is that, without the assignment problem, the existence of favoritism has no effect on worker incentives. When we further consider the costs of misallocation, favoritism *increases* worker incentives.

A. *Selling the Rights to Exert Bias*

The firm chooses λ to maximize surplus, which after some calculations yields

$$\lambda^* = \mu^2 \tilde{\tau}_s^{*2} r + 2L'. \tag{17}$$

Substituting this into the definition of $\tilde{\tau}_s^{*2}$ gives

$$\tilde{\tau}_s^* = \frac{\sigma_f^2}{(\sigma_f^2 + \sigma_s^2)(1 + \rho\sigma_\alpha^2) + \rho\sigma_f^2\sigma_s^2 - 2\dfrac{\mu^2\sigma_\eta^2}{r\lambda^2}[1 + \rho(\sigma_\alpha^2 + \sigma_f^2)]L'}. \tag{18}$$

Assume for the moment that the sorting problem does not exist, so $L' = 0$. Then σ_b^2 does not affect $\tilde{\tau}_s^*$. The firm chooses incentives as

[14] To illustrate the possibility that in some cases too much authority is offered to supervisors, the Bureau of National Affairs (1979) reports data from a survey on grievance procedures in organizations, where workers can complain about decisions made by supervisors to senior management. The bureau argues that the most striking aspect of its data, collected from 128 firms, is the reluctance of management to overturn supervisors' decisions, with almost half of the companies never overturning a decision during the period of observation.

though favoritism did not occur.[15] Therefore, favoritism is ignored when the firm is choosing the piece rate, despite adding risk to evaluations. Incentives for the worker are set *as though* $\sigma_b^2 = 0$. The only effect of favoritism is that the worker's salary is higher by the risk premium $\sigma_\eta^2/2r$. Straightforward calculations show that the surplus generated by supervisors' taste for power is also given by $\sigma_\eta^2/2r$. This demonstrates that, despite the opportunity for the abuse of power by the supervisor, the firm can harness these preferences so that welfare is generated without harming worker incentives.

In (17) and (18), greater misallocation costs (L') increase both λ and $\bar{\tau}_s^*$. The effect on λ is intuitive: higher costs of bad information raise the price that the firm charges for exerting power. But at this price, $dS/\partial\sigma_b^2 > 0$ (see [14]). Bias *raises* total surplus when power is priced optimally, so power is increased by giving greater weight to supervisors' reports in setting pay.

B. *Distorting the Market for Favoritism*

As we have just shown, favoritism need not imply low-powered worker incentives. Without distortions in the market for power, favoritism can create surplus through demands for power that are best exploited by offering the supervisor greater authority. We offer two plausible distortions to the market for power that may modify this result, leaving $\Omega > 0$ in equilibrium.

Supervisor Risk Aversion

Assume that the supervisor has exponential utility with coefficient of risk aversion $r_s > 0$. The effect of risk aversion in the standard linear compensation model (Holmstrom and Milgrom 1987) would be to limit risk by reducing λ. Further complexity arises here because the supervisor *affects* the amount of risk he faces through his choice of bias: σ_b^2 is decreasing in λ. After calculations provided in the Appendix, the optimal choice of λ solves

$$\hat{\lambda} = \frac{\sigma_b^2(r\bar{\tau}_s^{*2}\mu^2 + 2L')}{\sigma_b^2 + (\hat{\lambda}^3/2)r_s\,\mathrm{var}[E(y_f|y, e^*) - y_f]^2 - (r_s\sigma_b^2/\hat{\lambda})}. \tag{19}$$

[15] The optimal contract without favoritism (i.e., where $\sigma_\eta^2 = 0$) implies

$$\bar{\tau}_s^* = \frac{\sigma_f^2}{(\sigma_f^2 + \sigma_s^2)(1 + \rho\sigma_\alpha^2) + \rho\sigma_f^2\sigma_s^2}$$

as in (18) with $L' = 0$.

The optimal $\hat{\lambda}$ differs from λ^* in (17) by the final two terms in the denominator. The sign of these terms is ambiguous, so it is uncertain whether the introduction of risk aversion reduces $\hat{\lambda}$ from λ^*. The reason is that increasing λ has two effects. First, it increases the variance of rewards for any amount of noise. This leads to lower-powered incentives for standard reasons. But second, higher λ also reduces σ_b^2, so the risk faced by the supervisor can fall. This leads to higher λ under risk aversion than with risk neutrality, counter to normal intuition.

Whether risk aversion increases or decreases λ depends on which effect is larger. If increasing λ increases the total risk faced by the supervisor, then the firm will choose $\hat{\lambda} < \lambda^*$. This implies $\bar{\tau}_s < \bar{\tau}_s^*$. This is the case if supervisor monitoring is poor ($\text{var}[E(y_f|y, e^*) - y_f]^2$ is large). Then the firm will reduce worker incentive to combat favoritism.

Supervisor Effort

A more novel distortion arises when the supervisor exerts effort in monitoring. It is surely the case that firms monitor supervisors' activities for reasons beyond the control of favoritism. To capture this idea, we assume that the variance of the supervisor's error in observing performance declines as he applies more effort. Then the firm's choice of λ performs two functions: it affects supervisors' choice of bias, as above, but also provides incentives for supervisor effort. The choice of λ cannot yield efficient outcomes on both margins.

ASSUMPTION. The supervisor can reduce σ_s^2 by exerting effort e_s. Let $\sigma_s^2 = g(e_s)$, with $g' < 0$ and $g'' > 0$, and let the cost of exerting that effort be e_s.

We identify the "input possibility frontier" as the locus of choices of e_s and σ_s^2 that can be induced by the firm through its choice of λ. The firm uses λ to choose its preferred point on the frontier.

The supervisor's objective is to choose e_s and b to maximize $w_s - e_s$. After appropriate substitutions, this is equivalent to

$$\min_{e_s, b} \frac{\lambda}{2}\left(\frac{\sigma_s^2\sigma_\alpha^2}{\sigma_s^2 + \sigma_\alpha^2} + \sigma_f^2 + \sigma_b^2\right) + e_s \tag{20}$$

subject to $g(e_s) = \sigma_s^2$ and (12). The first-order condition gives the equilibrium relationship between e_s and σ_b^2:

$$\frac{-g'(e_s)\sigma_\alpha^4}{2(\sigma_\alpha^2 + \sigma_\epsilon^2)} = \frac{\sigma_b}{\sigma_\eta \bar{\tau}_s \mu}. \tag{21}$$

The firm can vary λ so as to induce any combination of e_s and σ_b^2 that is consistent with (21), where larger values of λ yield both higher effort and lower σ_b^2. The input possibility frontier defined by (21) slopes downward in $\{\sigma_b, e_s\}$ space as illustrated in figure 1.

The firm's objective is to maximize surplus (which is as in Sec. II minus the cost of supervisor effort) subject to the requirement that effort and bias lie on the frontier. Note from (21) that the position of the frontier depends on $\tilde{\tau}_s$. An increase in $\tilde{\tau}_s$ shifts the frontier out, as favoritism rises for any level of e_s. Immense simplification arises when the sorting problem is ignored ($L' = 0$). Then the optimal level of bias is $\sigma_b^{*2} = \sigma_\eta^2/r$ and is independent of supervisor effort.

Two cases must be considered in characterizing the firm's choice of $\tilde{\tau}_s$ and λ. In the first case, e_s is too low at the first-best level of bias, σ_b^*, given $\tilde{\tau}_s = \tilde{\tau}_s^*$ in (18). This is the case considered at B in figure 1, where the supervisor's effort is $\hat{e}_s < e_s^*$, and e_s^* is the optimal (first-best) level. By the definition of σ_b^*, the cost of reducing bias slightly from that level is second-order, but because $\hat{e}_s < e_s^*$, there is a first-order gain to increasing supervisor effort. The firm does this by increasing λ, which reduces bias along the frontier. By similar analysis, it is clear that the firm will never choose $e_s \geq e_s^*$. The result in this case is too little monitoring effort by supervisors, who are also given "too little" power: the outcome lies in the region AB.

This analysis is conditional on $\tilde{\tau}_s = \tilde{\tau}_s^*$. The firm will also increase worker incentives above $\tilde{\tau}_s^*$, shifting the frontier, since along AB there are first-order gains to increasing σ_b. To illustrate, suppose that the optimal choice of λ, conditional on $\tilde{\tau}_s^*$, generates point D in figure 2.

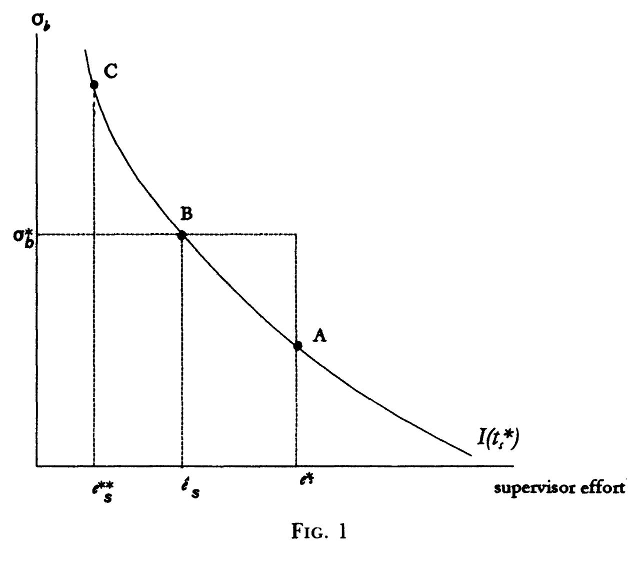

FIG. 1

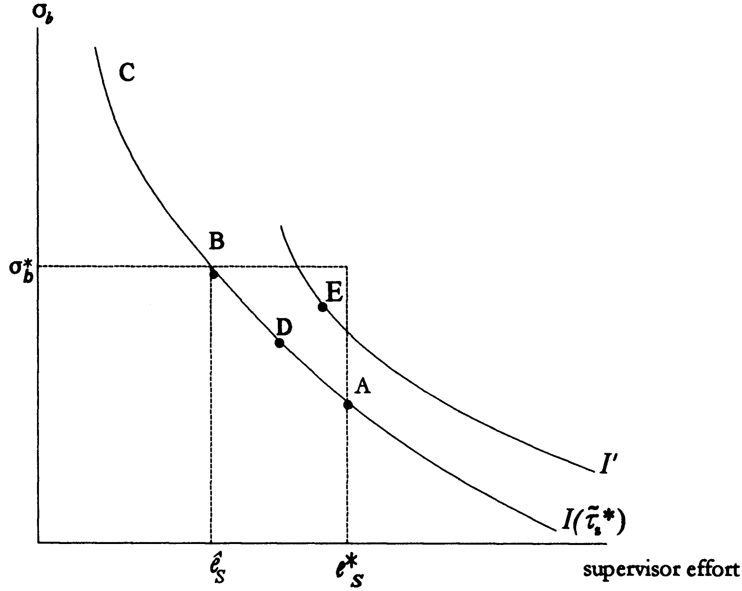

FIG. 2

This involves too little bias, so there are gains to increasing $\tilde{\tau}_s$ because doing so induces more bias. This shifts the supervisor's choice to a point such as E on frontier I'. The solution involves a higher value of $\tilde{\tau}_s$, more bias, and more supervisory effort than would occur with piece rate $\tilde{\tau}_s^*$. In this case, the supervisor's ability to control his effort yields stronger incentives for workers.

In the second case, at $\tilde{\tau}_s^*$, bias is harmful on the margin ($\Omega > 0$). This occurs when monitoring effort is too high at the first-best level of bias, σ_b^*. In figure 1, suppose that the efficient level of monitoring effort is e_s^{**}. Analogous reasoning to that discussed above shows that the firm would reduce λ, yielding a point on segment BC of the input possibility frontier, where the level of bias is inefficiently high. Then it pays to reduce $\tilde{\tau}_s$ below $\tilde{\tau}_s^*$, shifting the frontier in, because bias is harmful on the margin. In this case, greater pay equity ($\tilde{\tau}_s < \tilde{\tau}_s^*$) is profitable because it limits supervisors' incentive to exercise bias.

IV. Extensions

A. *The Effect of Compression*

In this paper, supervisors are monitored by how closely their reports correspond to information that the firm holds on the performance of the worker. This aspect of monitoring leads supervisors to compress their evaluations toward their beliefs on the firm's view of the worker. In this section we address how such compression affects incentives. Note that compression is captured by the parameter θ, where the supervisor reports $\theta y + (1 - \theta)e^* + b$ to the firm (see [8]). Here we analyze the effect of compression by considering the implication of the supervisor's reporting $y + b$, where any differences in compensation and welfare are caused by compression per se.

First, consider the case in which bias is harmful on the margin, $\Omega > 0$. Then compression harms equilibrium welfare and reduces worker incentives. The reason is that when the supervisor compresses his reports, the firm must increase the sensitivity of rewards to reported performance in order to provide incentives. In particular, for a compression parameter of θ, the firm must offer a piece rate on the supervisor's report of τ_s/θ to provide the same incentives as a piece rate of τ_s without compression. But from (12), increasing τ_s leads to more bias. In equilibrium, the firm responds to lower θ by reducing the effective piece rate $\tilde{\tau}_s$. Consequently, when favoritism is harmful on the margin, compression imposes costs on the organization.

This is not necessarily the case when the firm chooses λ optimally. Consider the case in which the supervisor is risk neutral and the sorting problem is ignored ($L' = 0$), so $\lambda^* = \mu^2\tilde{\tau}_s^{*2}r$. It follows from (18) that the optimal piece rate offered to the worker is independent of θ, the compression coefficient. Therefore, as with favoritism, understanding how compression affects incentives depends on how supervisors are monitored over perceived bias.

B. *Residual Claimancy*

An alternative means of monitoring the supervisor is to make him a residual claimant on the output of the worker. Supervisors who show favoritism will bear a direct cost, in the form of lower output from subordinates, if evaluations are arbitrary. Then the incentive to develop a reputation for fairness may resolve the problems of favoritism. Yet it does not follow that residual claimancy is the optimal solution to the problem of favoritism. For example, the ability to form a reputation for fairness need not imply that the supervisor acts honestly for the usual discounting reasons (Bull 1987). In addition, the opportunity to develop a reputation for honesty may entail inefficiencies of its own. First, with residual claimancy, the supervisor may have an incentive to renege on wages in order to increase his profits. Second, it is uncertain how reputation formation affects performance evaluation. For example, consider the case in which a supervisor provides a performance evaluation on a minority worker, where the supervisor does not wish to be labeled a racist. If the supervisor genuinely observes bad performance, what does he report? A poor evaluation may indicate a high likelihood of being racist. So it is plausible that the report would be inflated. Then the opportunity for reputation formation may result in false ratings, though for reasons other than those considered here.

C. *Observable Worker Characteristics*

Workers may vary in their propensities to be the subject of favoritism. Most obviously, minorities or women may be discriminated against because of personal characteristics. How will compensation policies adjust to reflect the possibility of this type of bias?

It is important to distinguish between differences across groups in the mean of the distribution of supervisor preferences and differences across groups in the variance of that distribution. In our analysis the mean of the supervisor's preferences has been normalized to zero. Allowing the mean of the distribution of η to vary across groups does not affect our qualitative results,[16] however, which says that expected favoritism has no effect on pay or incentives.[17] This is not the case when uncertainty about η changes. For example, assume that there is uncertainty about whether the supervisor is sexist. This is equivalent to increasing σ_η^2 for women relative to men. Intergroup differences in σ_η^2 affect compensation, but the effect depends entirely on how the market for favoritism operates.

For given $\tilde{\tau}_s$ and λ, an increase in σ_η^2 increases favoritism: $\sigma_b^2 = \tau_s^2\sigma_\eta^2/\lambda^2$. This exposes workers to greater risk and implies more allocation errors. Intuition suggests that this will lower $\tilde{\tau}_s^*$ and raise τ_f^*, with total incentives falling. This is correct when favoritism is harmful on the margin; inspection of (15a) shows that $\tilde{\tau}_s^*$ is decreasing in σ_η^2 if $\Omega > 0$. If preferences for minorities or women are more variable than for white males, less weight will be given to subjective evaluations by supervisors and more to explicit noncorruptible signals of output. This prediction is supported by evidence in Goldin (1990), who finds that women work in occupations that rely more heavily on objective performance measures. Our work also predicts that firms will seek

[16] The only complication it gives rise to is that because the supervisor values the worker's wage rather than utility in (3), the firm may increase effort merely to increase the wages of workers who are favored. We felt this effect to be uninteresting, so we ignored it by setting $E\eta = 0$ in the earlier analysis.

[17] Note the importance of continuous signals here. If signals were discrete, this would no longer be the case. For example, assume that a supervisor can report only on whether a worker is "good" or "bad." Then changes in the mean of η affect the likelihood of being reported as "good" or "bad" and make unraveling genuine performance much more difficult.

supervisors who are least likely to be sexist because sexism harms incentives and total surplus.

As above, this intuition is correct only if the costs of favoritism outweigh the benefits of authority. Consider the case in which the rights to favoritism can be sold efficiently and $L' = 0$. Then increases in σ_η^2 raise surplus, as total surplus generated by favoritism is given by $\sigma_\eta^2/2r$ and $\tilde{\tau}_s^*$ is independent of σ_η^2. In this case, piece rates are unaffected by σ_η^2. If sorting costs are introduced, $\tilde{\tau}_s^*$ is *increasing* in σ_η^2. That is, when the rights to favoritism can be sold, increases in the intensity of prejudice raise incentives, so *more* prejudice occurs. Prejudice is actually accommodated in this case, and firms may seek supervisors who particularly value authority, that is, those for whom σ_η^2 is large. Of course, workers are compensated for enduring this prejudice, so in equilibrium they suffer no ex ante harm.

V. Conclusion

The economics literature has generally ignored issues such as personal preferences toward employees and the demand for power by management. Our analysis is a first step toward understanding, first, that performance evaluation is inherently a subjective exercise and, second, that personal preferences toward employees become important determinants of the evaluations that workers receive. We have shown that compensation policies themselves affect monitoring efficiency, since supervisors lie more when money is on the line. This distorts incentives through additional risk imposed on workers, as is standard, and also harms the allocation of workers to jobs. In designing compensation and evaluation procedures, firms must balance these costs against the benefits that accrue to supervisors from the ability to exercise their preferences. Not surprisingly, how favoritism affects compensation hinges on distortions in the monitoring market for the supervisor. In the absence of any distortions, worker incentives are not reduced in order to combat favoritism. Indeed, they may increase.

The underlying premise of the paper is that sensitive compensation schemes are likely to result in behavior that affects the allocation of quasi rents, when supervisors misreport their evaluations to accord with their preferences. In this our work is related to Tirole (1992) on collusion within organizations and Milgrom (1988) and Milgrom and Roberts (1990) on rent seeking. Tirole considers a case in which subordinates bribe their superiors to obtain desirable outcomes. Inefficiencies arise in his model because the bribing mechanism is inefficient; that is, transfers between the worker and supervisor occur in inefficient ways. Milgrom (1988) and Milgrom and Roberts (1990) consider an alternative form of inefficiency in which subordinates waste valuable productive time lobbying for desirable outcomes.

In each of the papers just mentioned, an optimal response by firms will be to make compensation less sensitive to supervisor evaluations and to use bureaucratic rules placing "excess" weight on noncorruptible signals. In that sense this work is similar to ours. There are, however, a number of important differences. First, bribery and lobbying are unnecessary to illustrate the possible efficiency of these responses; all that is necessary is that supervisors have likes and dislikes toward their subordinates. Second, we posit two additional inefficiencies of subjective performance evaluation associated with their consequent arbitrariness, namely, increased risk and more inefficient task assignments. The third and most important distinction is that we illustrate the benefits of allowing supervisor discretion, since supervisors value their authority. We showed that the demand for power by supervisors fundamentally affects worker compensation. Firms may grant supervisors more power to affect the welfare of their subordinates, so favoritism may actually strengthen incentive pay. A related point is that we have illustrated the importance of distortions in the market for supervisor power in modeling the behavior and form of organizations.

This paper is also related to recent work on performance evaluation by Holmstrom and Milgrom (1991) and Baker (1992), who argue that incentive contracts can result in workers' substitution of their time and attention across activities. The typical example used is a worker who responds to an increased piece rate on output by increasing her quantity of output at the cost of lower quality. In this paper, changing worker incentives changes behavior beyond the direct effect on worker effort; here it is the *supervisor's* behavior that changes (rather than the worker's behavior) since he lies more when the worker is offered sensitive compensation. The implications of this behavior for compensation depend on whether these externalities are beneficial or harmful to organizational surplus.

Appendix

A. *Proof of Lemma 1*

The firm begins with the prior that $\alpha \sim N(0, \sigma_\alpha^2)$. It then receives a report from the supervisor that has variance $\sigma_s^2 + \sigma_b^2\mu^2 = \tilde{\sigma}_s^2$ and receives its own observation with a variance of σ_f^2. The supervisor's report and the firm's observation can be combined into a compositive signal with mean $(\sigma_s^2 y_f + \sigma_f^2 y_s)/(\sigma_s^2 + \sigma_f^2)$ with variance $\sigma_s^2\sigma_f^2/(\sigma_s^2 + \sigma_f^2)$. Then following Prescott and

Visscher (1980), we can characterize the misallocation loss as

$$L = E(\alpha|\alpha > 0) - 0.798\left[\frac{\sigma_f^2 + \tilde{\sigma}_s^2}{\sigma_f^2(\sigma_\alpha^2\sigma_f^2\tilde{\sigma}_s^2 + \sigma_f^2 + \tilde{\sigma}_s^2)}\right].$$

Then

$$\frac{\partial L}{\partial \sigma_b^2} = \frac{0.798\,\mu^2\sigma_f^6\sigma_\alpha^2}{[\sigma_f^2(\sigma_\alpha^2\sigma_f^2\tilde{\sigma}_s^2 + \sigma_f^2 + \tilde{\sigma}_s^2)]^2} > 0.$$

As bias enters only into the denominator of this expression, it should be obvious that $\partial^2 L/\partial b^2 < 0$, thus completing lemma 1.

B. *Risk Aversion*

Let $\Delta = E(y_f|y, e^*) - y_f$. The supervisor is rewarded by $s_s = w_0 - [\lambda/2(\Delta + b)^2]$. The surplus lost from risk aversion is parameterized by the variance of the supervisor's wage. As Δ and b are independent,

$$\mathrm{var}(w_s) = \left(\frac{\lambda^2}{4}\right)[\mathrm{var}(\Delta^2) + (b^2 - Eb^2)^2].$$

But

$$(b^2 - Eb^2)^2 = E\left[\frac{\mu^4\tau_s^4\sigma_\eta^4}{\lambda^4}\left(\frac{\eta^4 - 1}{\sigma_\eta^4}\right)\right]$$
$$= \frac{2\mu^4\tau_s^4\sigma_\eta^4}{\lambda^4}$$

since the coefficient of kurtosis for a normal distribution is three. Let S be the surplus generated with a risk-neutral supervisor, as in (13). Then with a risk-averse supervisor, the maximization problem is

$$\max_{\lambda,\tilde{\tau}_s,\tau_f} S - \frac{r_s\lambda^2}{8}\left[(\Delta^2 - E\Delta^2)^2 + \frac{2\mu^4\tau_s^4\sigma_\eta^4}{\lambda^4}\right].$$

Straightforward maximization yields (19). Q.E.D.

References

Baker, George P. "Incentive Contracts and Performance Measurement." *J.P.E.* 100 (June 1992): 598–614.

Bjerke, D.; Cleveland, J.; Morrison, R.; and Wilson, W. "Officer Fitness Report Evaluation Study." Navy Personnel Research and Development Center Report, no. TR 88-4. Navy Personnel Res. and Development Center, 1987.

Bretz, Robert, and Milkovich, George T. "Performance Appraisals in Large Organizations: Practice and Research Implications." Working Paper no. 89-17. Ithaca, N.Y.: Cornell Univ., Center Advanced Human Resources Studies, 1989.

Bull, Clive. "The Existence of Self-Enforcing Implicit Contracts." *Q.J.E.* 102 (February 1987): 147–59.

Bureau of National Affairs. "Policies for Unorganized Labor." Personnel Policies Firm Survey no. 125. Washington: Bur. Nat. Affairs, 1979.

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

Goldin, Claudia. *Understanding the Gender Gap: An Economic History of American Women.* New York: Oxford Univ. Press, 1990.

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

———. "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design." *J. Law, Econ. and Organization* 7 (suppl.; 1991): 24–52.

Landy, Frank J., and Farr, James L. "Performance Rating." *Psychological Bull.* 87 (January 1980): 72–107.

Milgrom, Paul R. "Employment Contracts, Influence Activities, and Efficient Organization Design." *J.P.E.* 96 (February 1988): 42–60.

Milgrom, Paul R., and Roberts, John. "The Efficiency of Equity in Organizational Decision Processes." *A.E.R. Papers and Proc.* 80 (May 1990): 154–59.

Milkovich, George T., and Wigdor, Alexandra K., eds. *Pay for Performance: Evaluating Performance Appraisal and Merit Pay.* Washington: Nat. Acad. Press, 1991.

Mohrman, A., and Lawler, E. "Motivation and Performance Appraisal Behavior." In *Performance Measurement and Theory,* edited by Frank J. Landy et al. Hillsdale, N.J.: Erlbaum, 1983.

Murphy, Kevin, and Cleveland, Jeanette N. *Performance Appraisal: An Organizational Perspective.* Boston: Allyn & Bacon, 1991.

Prescott, Edward C., and Visscher, Michael. "Organization Capital." *J.P.E.* 88 (June 1980): 446–61.

Spilerman, S. "Organizational Rules and the Features of Work Careers." In *Research in Social Stratification and Mobility,* vol. 5. Greenwich, Conn.: JAI, 1986.

Tirole, Jean. "Collusion and the Theory of Organizations." In *Advances in Economic Theory: Sixth World Congress,* vol. 2, edited by Jean-Jacques Laffont. Cambridge: Cambridge Univ. Press, 1992.