NBER WORKING PAPER SERIES

FAVORITISM IN ORGANIZATIONS

Canice Prendergast

Robert H. Topel

Working Paper No. 4427

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 1993

This research was supported by the Centel Foundation/Robert P. Reuss Faculty Research Fund at the University of Chicago Graduate School of Business (Prendergast) and by the National Science Foundation (Topel). A previous version of this paper was circulated under the title "An Agency Approach to Bias in Organizations." We are grateful for comments from Edward Lazear, Lars Stole, Raghu Rajan, and seminar participants at Cornell, Wisconsin, Michigan, Stanford, Chicago, Northwestern, Rochester, Tokyo, and Texas A&M. This paper is part of NBER's research program in Labor Studies. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

NBER Working Paper #4427
August 1993

FAVORITISM IN ORGANIZATIONS

ABSTRACT

Performance evaluations for workers are typically subjective impressions held by supervisors rather than easily quantifiable measures of output. We argue that perhaps the most important aspect this is that it gives supervisors the opportunity to exercise their personal preference towards their employees in a way that is detrimental for performance. both for incentive reasons and through misallocation of workers to jobs.

We illustrate that firms will respond to this problem in a number of ways. First, they will make compensation less sensitive to performance, even when workers are risk neutral. Furthermore, firms will typically use bureaucratic procedures for allocating rewards, even though these are known to be ex post inefficient. In addition, firms may tie wages to jobs as a means of credibly rewarding the best performers. These organizational changes are used because directly monitoring supervisors' decisions is fraught with problems, among them the creation of 'yes men,' so that the indirect mechanisms described above are likely to be optimal responses to favoritism.

Canice Prendergast
Graduate School of Business
University of Chicago
Chicago, IL 60637
and NBER

Robert Topel
Graduate School of Business
University of Chicago
Chicago, IL 60637
and NBER

Most of the economics literature on compensation and organization builds from the theory of agency. For the most part, the literature analyzes situations in which agents' performance can be controlled by tying compensation to objective performance measures such as output or sales. It ignores the fact that most compensation arrangements involve superiors' *subjective*, and hence noncontractible, judgments about employee performance. In our view, much of what is interesting about actual employment relations follows from the observation that "performance appraisal is a process by which humans judge other humans" (Milkovich and Wigdor, 1991).

Many potentially important issues arise when performance evaluations are based on noncontractible data. See Prendergast and Topel (1993) for a survey. Perhaps the most important implication of subjectivity of performance evaluation is that it provides supervisors with the opportunity to exercise favoritism or bias towards employees based on personal preferences rather than related to performance. The objective of this paper is to outline the costs that favoritism imposes on organizations and how many institutions that we observe in organizations may be a reflection of attempts to minimize the discretionary power of managers.[1] We illustrate

[1]The paper is both complementary yet has a different orientation to papers by Milgrom (1988) and Milgrom and Roberts (1988, 1990), who focus on the possibility that employees waste valuable resources

how firms may (i) introduce more equality across workers than they would otherwise want, (ii) introduce bureaucratic rules to assign resources rather than using optimal ex post assignment rules, and (iii) tie wages to specific tasks rather than separate sorting and incentive issues, to mitigate the inefficiencies of favoritism.

Before describing the theoretical model, it is useful to review the literature on discretion in performance appraisals. Perhaps the most common finding on performance appraisals is the perceived tendency toward leniency, where supervisors are reluctant to give unpleasant ratings to employees.[2] This could, of course, be merely a language problem where workers and managers "really know" that a rating of 3 out of 5 means incompetence. More useful for our purposes is data reflecting (i) that the exercise of discretion potentially varies across individuals and (ii) that this discretion also varies with the use of the performance evaluation. For example, Kraiger and Ford's (1985) survey of the effect of race on ratings reported that the race of both the rater and the ratee had an influence on performance ratings. Supervisors gave higher ratings to same-race subordinates than to subordinates of a different race. Thus supervisor discretion may operate differentially

---

trying to persuade superiors of their talents.

[2]See Bretz and Milkovich (1989) for a discussion of this literature.

across individuals. Another indication of opportunities for bias in subjective performance evaluation is the widespread assumption that the *uses* of performance appraisals affect appraisal outcomes (see Landy and Farr 1980, Mohrman and Lawler 1983, and Murphy and Cleveland 1991). In particular, ratings used to make administrative decisions such as merit pay or promotion are more inflated than are ratings used for employee feedback (see Williams et al. 1985, Reilly and Balzer 1988). Therefore, discretion appears to be exercised in situations when "money is on the line." This will be important in our analysis below.

We begin in Section 1 by assuming that a supervisor has personal preferences towards his subordinates, where these preferences reflect the extent to which he likes or dislikes the subordinate. This information is privately held by the supervisor. We assume that the role of the supervisor is to make performance appraisals on his workers. This information is potentially useful for two reasons. First, the firm would like the workers to exert effort and second, the firm also wishes to determine the talents of its employees to aid efficient task assignment. Then the impact of favoritism is that performance evaluations are contaminated by additional noise caused by the supervisor's preferences. This renders monitoring less efficient, leading to inefficiencies both

in terms of inducing workers to exert effort and assigning workers to jobs.

It should be remembered that supervisors receive utility from exercising their preferences over their subordinates, so in some circumstances favoritism may be welfare-improving.[3] The results we describe in the paper are predicated on the assumption that favoritism reduces aggregate surplus, so that the firm would like to reduce favoritism as a means of allocating resources. Of more interest than merely outlining that favoritism can lead to inefficiencies is how organizations are likely to deal with it. In Section 1, we illustrate that firms are likely to respond by introducing more equality than they would otherwise want, in the sense that the rewards offered to workers for good performance are reduced relative to a case where favoritism does not exist. It is important to bear in mind that this is not simply a *result* of the additional noise in the monitoring process; instead, the compensation scheme itself *generates* noise because more sensitive compensation schemes lead to more lying by supervisors as there is more money "on the line," consistent with the empirical evidence

[3]This differs from the work on influence costs such as Milgrom and Roberts (1988).

described above.[4] Introducing more equality does, of course, harm incentives, but it is shown both to reduce aggregate risk and to improve the allocation of workers to jobs. Section 1 therefore provides a justification for "the traditional rule of thumb among managers of performance appraisal has suggested the wisdom of decoupling the appraisal system from merit pay" (Milkovich et al., 1991, p. 109).

Section 2 extends the analysis by considering an alternative device to limit the costs of favoritism, namely, the use of bureaucratic rules. We extend the basic model to include performance measures other than the supervisor's report on the worker. For example, the other piece of information could be the worker's tenure in the company or a manager's observation. We characterize bureaucratic rules as being used when information on the worker's performance is used for compensating him in a different way from the optimal ex post way of aggregating information on the worker's performance. We show that the manager optimally commits to a rule that places "too little" weight on the supervisor's opinion in the evaluation process. Paradoxically, the manager

---

[4]"The concern has been that managers will deliberately inflate performance appraisals 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*" (Milkovich et al., 1991, p. 109, italics added).

benefits from committing to underutilize the supervisor's information to make that information better.

It is worth emphasizing again that this result is not simply that noisier performance evaluations lead to less weight being placed on the supervisor's observation. Instead, the firm places even less weight on the supervisor's opinion.

So far, we have considered indirect responses by the firm to favoritism in that we have described how changing how *a worker* is compensated can reduce malfeasance by *his supervisor*. Section 3 addresses a more direct solution, where the firm monitors the supervisor by obtaining another observation on the worker and penalizes the supervisor if he is perceived to have acted unfairly. We begin by illustrating how taking multiple observations can give rise to the first-best outcome with a risk neutral supervisor, where the firm compares its observation on the worker with the supervisor's report.

The remaining parts of Section 3 illustrate that this scenario may be unduly optimistic. First, if a supervisor is punished for exercising bias, he has an incentive not to get caught. We illustrate that supervisors are likely to make uncontroversial recommendations about workers rather than honestly report their observations. But this tendency toward the creation of "yes men," who attempt to second-guess the manager, results in costs imposed on the organization through

inefficient monitoring. Extending Prendergast (1993), we show that this incentive mitigates the desirability of punishing the supervisor for indiscretions of favoritism, so that the firm may continue to allow the supervisor to exercise discretion, despite the ability to monitor.

Section 3.2 describes an alternative cost to monitoring supervisors on performance appraisals. Supervisors typically carry out other activities than performance appraisals that could affect a worker's career. For example, in addition to directly monitoring performance, supervisors typically assign workers to tasks, provide on-the-job training, and aid workers in difficulty. Here it is argued, using results in Holmstrom and Milgrom (1992), that if distorting the worker's evaluations exhibits diminishing returns, then lying on performance appraisals and distorting training opportunities are substitutes. In particular, if the supervisor lies less on performance appraisals, he will distort more on training. The implication of this should be clear: the firm will not penalize severely on performance appraisals in order to make training more efficient. Therefore, once again, monitoring the supervisor may be fraught with problems.

Another means of aligning the supervisor's incentives with those of the organization is to make the supervisor residual claimant on output. This is addressed in Section 4.

Typically, we imagine that making the supervisor residual claimant improves the supervisor's incentives as the supervisor may feel that indiscretion may affect his reputation for fairness. The development of reputations to support honest behavior is by now well known (see Bull, 1986, for example). Rather than explore this avenue, we consider a case where reputation formation is impossible. In the absence of reputation formation, we show that the firm can reduce favoritism by tying wages to tasks, in the sense that a worker can be rewarded for good performance only by promotion to a new, better-paid job. In other words, incentives and sorting workers to jobs are linked.

The intuition for this result is as follows. Suppose that sorting and incentives are independent, in the sense that task assignment carries no wage implications. Then the supervisor will simply assign incentive money based on personal whims if he cannot credibly build up a reputation, despite being residual claimant. Realizing this, workers shirk. On the other hand, now assume that wages are attached to jobs. Then if the supervisor rewards his favorites, he is likely to suffer because his favorites may not be qualified for the tasks that carry the higher wages. This acts as a constraint on favoritism and, consequently, results in effort exertion by workers.

The paper is set out in a straightforward fashion.

Section 1 describes the basic model and the tendency toward equality in organizations. The analysis is extended in Section 2 to illustrate the likely use of bureaucratic rules, while Section 3 describes the costs of monitoring supervisors. In Section 4, we illustrate how promotion-based compensation schemes can be used to mitigate favoritism so that firms may optimally link sorting and incentive issues. We conclude with a brief discussion.

I. The Costs of Favoritism and the Tendency Toward Equality within Organizations

Favoritism can involve inefficiencies on a host of dimensions. First, employees may quit if they feel discriminated against, resulting in turnover costs and lost specific human capital. Alternatively, it may affect incentives to exert effort, as employees may feel that the additional "noise" in the monitoring process caused by favoritism may render the relationship between performance and rewards less obvious. Another problem with the existence of discretion by supervisors over their subordinates is to induce behavior by subordinates seeking to ingratiate themselves with their superiors (Milgrom and Roberts 1988, 1990, Tirole, 1992).

We consider an organization with two objectives. First, there is a desire to induce employees to exert effort, as in a

standard agency context. However, in addition, the firm must sort its employees to jobs. It is shown that an optimal response to the existence of favoritism is a tendency toward more equality than would otherwise occur. This move toward equality is not caused solely by the additional noise created by favoritism, but rather is also generated by how changing compensation affects the marginal return to exercising whatever preferences supervisors have toward their subordinates.

The Model

The firm employs two workers, who exert effort, and a supervisor, who monitors their efforts.

A. Incentives

The workers, 1 and 2, exert effort e^i, $i = 1, 2$ in producing output

$$y^i = e^i + \alpha^i + \epsilon^i, \tag{1}$$

where y^i is output privately observed by the supervisor, α^i is the ability of the worker, and ϵ^i is an error term with mean zero, which is normally distributed with variance σ_ϵ^2. The cost of effort is $c(e^i)$, where $c' > 0, c'' > 0, c'(0) = 0, c'(\infty) = \infty$.

Ability is assumed to be unknown to all agents and is perceived to be normally distributed with mean 0 and variance σ_0^2. All primitive random variables are uncorrelated across workers.

The primary objective of this section is to illustrate how favoritism changes the way workers are paid. As a result, we make use of a special case of agency problems that result in linear sharing rules, characterized by Holmstrom and Milgrom (1987), where the agents have exponential utility given by

$$v^i = -(\exp(-r(w^i - c(e^i)))). \tag{2}$$

Output is privately observed by the supervisor. However, the supervisor may distort his report on the basis of his preferences towards the subordinates. The preferences of the supervisor are assumed to be given by

$$v_s = w_s + \eta^1 w^1 + \eta^2 w^2, \tag{3}$$

where w_s is the wage of the supervisor, w^i is the wage of worker i, and η^i is the "intensity" of the supervisor's preferences towards worker i. Favoritism is therefore measured by positive or negative altruism. It is assumed that η^i is unknown to all agents except the supervisor, but is only

realized by the supervisor after he joins the firm.[5] The distribution of η^i is assumed to be normal with mean 0 and variance σ_η^2.

We assume that rather than reporting y^i about worker i, the supervisor reports $y^i + b^i$, where b^i is the extent to which the evaluation is distorted based on the supervisor's preferences. In this section of the paper, we impose a cost to the supervisor from exercising bias. More specifically, there is an assumed monitoring technology for the supervisor's reports. How such a monitoring technology should be determined is discussed at length in Section III. Here we simply assume that the cost to exercising bias is given by $D(b^i) = 1/2\, b^{i2}$. In other words, the cost of telling the truth ($b^i = 0$) is zero but the marginal cost is increasing as the supervisor increases his distortions. This cost is assumed to be a transfer from the supervisor to the firm and has no effect on aggregate welfare.

The final ingredient necessary for our results is how the workers are compensated. Holmstrom and Milgrom (1987)

---

[5]This is a potentially contentious assumption, as subordinates may have some idea about how they are perceived by the superiors. The implication of using this assumption that considered in the text is that an additional distortion that the firm must worry about is where the worker quits, as his boss does not like her. We feel that this distortion would give rise to the same responses by the firm as we describe for the model where the supervisor's preferences are unknown. We believe that what is key for the results is that the firm not know the supervisor's preferences.

have illustrated that with normally distributed errors and exponential utility functions, the optimal compensation function is of the form

$$w^i = \delta + \tau \hat{y}^i, \tag{4}$$

where $\hat{y}^i$ is the supervisor's report on worker i, τ is the piece rate on output, and δ is a salary component. (τ and δ are common across workers because workers are ex ante indistinguishable; the error terms are uncorrelated so that there is no need for relative performance evaluation.) This compensation function is only optimal if the endogenous "error term" (bias) is normally distributed. The quadratic cost function assumed above is important for this.

B. Sorting

In addition to providing workers with incentives, the firm also wishes to use the performance evaluation scheme to determine which of the two workers is more talented. In particular, we assume that the worker with the higher evaluation is assigned to another job, which we call job B, where the marginal product of talent is larger than is specified by the technology described above. In effect, that worker is promoted, while the other worker remains in her old job, which

we call job A. Note that the marginal product of ability is normalized to unity in (1). We assume that job B has a marginal product of ability given by $\gamma > 1$. In all other respects, the two jobs are the same. Then consider the implication of promoting the "wrong" worker, i.e., the worker with lower α^i, to job B. The "misallocation cost" is then given by $(\gamma - 1)(\alpha^j - \alpha^i)$ in that profits could be increased by this amount if the correct worker was assigned.[6]

Timing in the model is as follows. The firm first chooses the contract for the workers. The workers then exert effort and y^i is privately observed by the supervisor, who then distorts his report based on his preferences. Following this, the workers are compensated and the worker with the best report is assigned to the new job.[7]

The Supervisor's Incentives

As mentioned above, with normal error terms, a linear sharing rule is optimal. All that remains to show the optimality of the linear scheme is to show that bias is normally

---

[6]This setup offers no return to promotion for the worker. Consequently, she has no incentive to exert effort to appear talented and hence be promoted. This allows us to ignore the reputation effects associated with unknown ability, as developed in Holmstrom (1982).

[7]As there is some value to the supervisor's information, it is always efficient to promote the worker with the better report.

distributed under a linear sharing rule. Straightforward maximization shows that the supervisor chooses b^i to

$$\max \; w_s + \eta^1 w^1 + \eta^2 w^2 - (1/2)\,(b^i)^2 \qquad (5)$$

subject to (4), which implies that $b^i = \tau \eta^i$. Then from an ex ante perspective,

$$b^i \sim N(0, \tau^2 \sigma_\eta^2) \qquad (6)$$

A useful result is that $\partial b^i / \partial \eta = \tau$ and $\partial^2 b^i / \partial \eta \partial \tau = 1$ so that as sensitivity of pay to reported performance (τ) increases, the incentive to distort rises. In other words, when there is more money "on the line," supervisors are likely to lie more. Also note that by substituting (6) into (3) that the supervisor's expected utility gain from his ability to exercise favoritism on each worker is $E[\tau^2 \; \eta^{i2}] = \sigma_b^2$, the variance of bias. This expected gain is also larger when there is more money on the line, so that the returns to being "boss" is the right to affect others' rewards.

The Worker's Incentives

Each worker faces a piece rate of τ so she chooses effort such that $c'(e^i) = \tau$.

Objective of the Firm

As is well known, the objective of the firm is to maximize ex ante surplus as all individual rationality constraints bind in equilibrium. Therefore all individual rationality constraints are ignored and incorporated into the surplus maximization problem. This is given by

$$\max_{\delta,\tau} \sum_{i=1}^{2} [y^i - c(e^i) - \frac{r}{2}\tau^2(\sigma_s^2 + \sigma_b^2)] + 2\sigma_b^2 - (\gamma - 1)(E\alpha \mid \alpha > 0)\ prob(\alpha < 0 \mid \hat{y}^1 > \hat{y}^2) \tag{7}$$

where $\alpha = \alpha^1 - \alpha^2$ and $\sigma_s^2 = \sigma_\epsilon^2 + \sigma_0^2$.

Surplus has a number of components. First, $y^i - c(e^i)$ is output produced minus effort exerted, as is common in agency problems. The risk premium $\frac{r}{2}\tau^2[\sigma_s^2 + \sigma_b^2]$ differs from the standard linear compensation problem only in that σ_b^2 is endogenous, dependant on the firm's choice of piece rate. Next, $2\sigma_b^2$ is the expected benefit to exerting bias to the supervisor. Finally, $(\gamma - 1)\ (E\alpha \mid \alpha < 0)\ prob(\alpha < 0 \mid \hat{y}^1 > \hat{y}^2)$ is the cost of allocating the wrong worker to the new job and consists of the expected cost of misallocation $(\gamma - 1)(E\alpha \mid \alpha < 0)$ conditional on the wrong worker being promoted, times the probability that $\alpha^1 - \alpha^2 < 0$, given $\hat{y}^1 > \hat{y}^2$, i.e., the probability of misallocation.

Throughout the paper, we stress that compensation schemes themselves generate inefficiencies, as supervisors distort their observations more when their reports count for more. It is therefore useful to consider a benchmark where the variance of bias is exogenous. More specifically, suppose that σ_b^2 was exogenously given. Then the optimal piece rate would be

$$\hat{t} = \frac{1}{1 + rc''[\sigma_s^2 + \sigma_b^2]}. \tag{8}$$

This piece rate is clearly lower than if no bias existed ($\sigma_b^2 = 0$). However, as we shall see, in the case where favoritism reduces welfare, the firm will choose a piece rate lower than this again.

Given the structure of the model, it is not clear whether exercising bias is a good or a bad. Clearly it involves costs to the workers, but at the same time it makes the supervisor better off. The objective of this paper is to consider the harmful effects of favoritism: if these are welfare-improving trades, there is little reason for firms to develop schemes to monitor their existence. In order to restrict attention to the case where favoritism is harmful, we assume that the cost to the workers exceeds the benefits that accrue to the supervisor

by Assumption 1.[8]

Assumption 1: $r\hat{\tau} > 2$.

From (7), Assumption 1 is a sufficient condition for surplus to be decreasing in σ_b^2 at a piece rate of $\hat{\tau}$, as the risk imposed on the workers reduces welfare more than the benefits to the supervisor.

Lemma 1 illustrates the first reason for imposing equality in organizations.

Lemma 1: Let $L(\tau) = (\gamma - 1)\,(E\alpha \,|\, \alpha > 0)\,\text{prob}(\alpha < 0 \,|\, \hat{y}^1 > \hat{y}^2)$ be the misallocation cost for a contract with piece rate τ. Then $L'(\tau) > 0$, so that a higher piece rate leads to worse allocation of workers to jobs.

Proof: See Appendix.

Lemma 1 illustrates one of the principal points of the paper, namely, that as the piece rate rises, there is more

[8]Note that this assumption does not imply that no bias offers highest utility, even conditional on an exogenous effort choice by the worker. This is because for small τ, the marginal benefit to the supervisor from exercising favoritism is greater than the cost to the worker.

money "on the line," which gives supervisors a greater incentive to distort their reports on workers. In turn, this leads to more costly mistakes in job placement.

The Optimal Choice of τ

The objective of the firm is to maximize (7) subject to $c'(e^i) = \tau$ and $b^i = \tau\eta^i$. This yields the first-order condition

$$1 - c'(e^i) = rc'c''[\sigma_s^2 + \sigma_b^2] - 2c'c''\sigma_\eta^2[r\tau^2 - 2] + L'(\tau). \quad (9)$$

Simplifying and using the fact that $\tau = c'(e^i)$ yields

$$\tau^* = \frac{1 - L'(\tau)}{1 + rc''[\sigma_s^2 + \sigma_b^2] + 2(r\tau^2 - 2)c''\sigma_\eta^2} < \hat{\tau} \quad (10)$$

if Assumption 1 holds. Equation (10) illustrates the major points of Section 1. Increasing the sensitivity of pay to reported performance leads to more lying by the supervisor, which is costly to the firm through misallocation costs. Consequently, the firm reduces τ in order to limit the effects of supervisor discretion. Similarly, increasing the piece rate increases the risk faced by the workers (as the supervisor lies more) which also reduces the optimal sensitivity of pay to

performance. In addition, note that there is a marginal effect of increasing τ on the noise in the monitoring process, given by the last term in the denominator of (10), so that $\tau^* < \hat{\tau}$. In this sense, it is not that bias simply adds noise to monitoring, which would cause the the firm to reduce τ. Instead, the reward structure is determined by the realization that the compensation scheme itself generates noise in the evaluation process.[9]

II. Bureaucracy

Section I illustrates how favoritism can be mitigated by making compensation less sensitive to performance evaluations than it would be in the absence of supervisor bias. We now illustrate how an organization may use bureaucratic rules for the same purpose.

Bureaucratic rules pervade many facets of organizations. We typically think of bureaucracy in terms of "red tape," but more generally it refers to the use of rigid rules for allocating resources. In the context of human resource management, for example, many organizations commit

[9]Note that if the misallocation problem is small and workers are risk neutral, then the optimal τ exceeds unity. More generally, this point illustrates that firms should design jobs in such as way as to benefit from the utility that individuals receive from exercising authority over others.

to rules that give priority to seniority in promotion decisions, and often require minimum experience requirements in a particular job before an employee can be promoted, independent of the worker's talents.

The bureaucratic nature of these practices is not that seniority should be correlated with promotion prospects or that, on average, some experience is necessary before promotion, but rather that firms commit themselves to rules for allocating resources rather than deciding on appropriate assignments on a case-by-case basis. Consequently, we characterize decisions as "bureaucratic" when the firm commits itself to a compensation rule that does not efficiently aggregate information ex post. Instead, the firm could attach "too much" weight to a less informative signal and "too little" to others when compensating employees.

The job allocation problem described above plays no particularly interesting role in the use of bureaucracy in organization so it is ignored in this section of the paper. Here, the firm's only objective is to induce effort exertion at minimum cost. Similarly, ability plays no important role here so we assume that $\sigma_0^2 = 0$. Given this assumption, we can simply concentrate on the incentives of a single worker, so that all worker superscripts are ignored throughout this section.

In order to model this process, it is necessary to include another piece of information on the worker's productivity. We assume that the firm receives an independent observation on productivity given by

$$y_f = e + \epsilon_f + \alpha, \tag{11}$$

where $\epsilon_f \sim N(0, \sigma_f^2)$. This information could either be some objective information on the worker's performance, such as quantity produced, or it could simply be a manager's impression of the worker's performance.

For future reference, it is useful to note that the firm's piece of information can potentially be used for two purposes. First, as emphasized in this section, it can be used as an additional piece of information to monitor the performance of the worker. However, it can also be used to monitor the supervisor's favoritism, as the firm's observation can be compared to the supervisor's to get an observation on the bias exerted by the supervisor. This issue is addressed at length in the next section. In this section, we assume that the firm's observation is used only as an additional instrument to compensate the worker.

As a benchmark, it is worthwhile to consider the case where the variances of the monitoring errors for the firm and

supervisor are exogenously assumed to be σ_f^2 and σ_s^2, respectively. Then it is straightforward to show that the optimal compensation function

$$w = \alpha + \tau_s \hat{y}_s + \tau_f y_f \tag{12}$$

is given by

$$\tau_s = \frac{\sigma_f^2}{\sigma_f^2 + \sigma_s^2 + c'' r \sigma_f^2 \sigma_s^2} \tag{13}$$

and

$$\tau_f = \frac{\sigma_s^2}{\sigma_f^2 + \sigma_s^2 + c'' r \sigma_f^2 \sigma_s^2}. \tag{14}$$

Of particular importance is that $\frac{\tau_s}{\tau_f} = \frac{\sigma_f^2}{\sigma_s^2}$ so that the weights attached to the supervisor's information relative to the firm's are exactly in proportion to their respective variances. This is precisely the optimal means of aggregating information on the worker's performance ex post, where the firm's ex post estimate of the worker's effort is

$$\hat{e} = \frac{\sigma_f^2 \hat{y}_s + \sigma_s^2 y_f}{\sigma_f^2 + \sigma_s^2}, \tag{15}$$

so that the relative value of $\hat{y}_s$ in estimating effort is given by

σ_f^2/σ_s^2. Thus greater uncertainty about the firm's information increases reliance on the supervisor's report.

The purpose of what follows is to illustrate that when bias is endogenously chosen by the supervisor, the firm will commit to a rule that attaches less weight to the supervisor's information than the optimal means of aggregating information ex post.

The firm's objective is now to

$$\max \; y - c(e) - \frac{r}{2}[\tau_s^2\sigma_s^2 + \sigma_f^2\tau_f^2] + \sigma_b^2, \tag{16}$$

where σ_b^2 is as defined in Section 1, subject to $\tau_s + \tau_f = c'(e)$ and $b = \tau_s\eta$. The first-order conditions to this problem are

$$1 - c'(e) = c''r[\tau_s\sigma_s^2 + \tau_s^3\sigma_\eta^2] - 2\tau_s^3\sigma_\eta^2 \tag{17}$$

and

$$1 - c'(e) = c''r[\tau_f\sigma_f^2]. \tag{18}$$

These imply that

$$\frac{\tau_s}{\tau_f} = \frac{\sigma_f^s}{\sigma_s^2 + (r\tau_s^2 - 2)\sigma_\eta^2} < \frac{\sigma_f^2}{\sigma_s^2}, \tag{19}$$

if favoritism is welfare reducing, i.e., if $r\tau_s^2 < 2$. Therefore, the firm commits itself to a rule that does not aggregate informa-

tion in an efficient fashion, in order to reduce the incentive for the supervisor to exercise bias.

This point is worth emphasizing. This section does not simply imply that the additional noise induced by favoritism leads to less weight placed on the supervisor's report. Rather, the firm will commit itself to a rule that places *even less* weight than is used when information is aggregated efficiently, despite the lower value of the supervisor's information. It is in this sense that bureaucratic rules are used, placing "too much" weight on statistics about performance that cannot be corrupted.

It is worth noting here that this solution relies on the firm having credible commitment power to use this solution. It is clear that ex post the firm would like to attach more weight to the supervisor's report than in the bureaucratic process described above. With this in mind, it is not surprising that firms often use objective criteria, such as seniority or quantity of output produced, when making these decisions rather than other more implicit measures which are harder to measure, which the firm might be tempted to use.

III. Monitoring the Supervisor

The remedies considered in Sections I and II are indirect solutions to the problem of favoritism as *the worker's*

compensation is being varied to remedy transgressions by *the supervisor.* A more direct response is to consider monitoring *of the supervisor* by the firm, where the supervisor's compensation is adversely affected if the firm feels that the supervisor is lying about performance appraisals.

The typical solution for monitoring supervisors is to obtain multiple observations on a worker's performance. This is useful for two reasons. First, more observations on the worker's performance improve the monitoring mechanism, as in any agency setting. However, multiple monitoring also serves to monitor the behavior of the monitors, because the report of any one monitor can be compared to others to estimate likely transgression.[10]

3.1. The Use of Multiple Observations

In the previous sections, the inefficiencies of favoritism can be parameterized by σ_b^2. This is treated as a reduced form here, where σ_b^2 is conditional on some exogenous compensation function with piece rate τ for the worker, as described in Sections I and II. The supervisor observes y, as

[10]In many organizations, punishment for potential transgressions is likely to occur in a vague fashion, such as a stain on the supervisor's reputation that adversely affects his chances of promotion. This issue is circumvented here by simply assuming a compensation function where the supervisor is rewarded and punished directly on the basis of a test statistic that is correlated with favoritism.

above but reports $\hat{y}$, while the firm observes y_f. The test statistic used to compensate the supervisor is

$$\hat{y} - y_f = B. \tag{20}$$

The supervisor is then assumed to be compensated by the function

$$w_s = \mu - \lambda B^2, \tag{21}$$

where the supervisor is penalized on the basis of how far the test statistic is from zero. Note the similarity between this compensation scheme and the reduced-form cost function used in Sections I and II.[11]

The parameter λ measures the intensity with which the firm penalizes the supervisor for perceived transgressions. The salary component μ is chosen merely to satisfy the supervisor's individual rationality constraint and plays no further role in the analysis. The supervisor chooses b to

$$\max_b \; E[w_s + \eta w], \tag{22}$$

---

[11]There is actually a continuum of compensation schemes that can induce the first-best as the supervisor is risk neutral. The quadratic structure is chosen only for simplicity. It should also be noted that we have introduced convexity in the cost function as a substitute for concavity in the supervisor's preferences.

where w is the wage of the worker. This can be simplified in a straightforward fashion to

$$\max_b \quad E[\ \mu - \lambda(\hat{y} - y_f)^2 + \eta w], \tag{23}$$

The objective of the supervisor is to predict the firm's observation to maximize his monetary rewards. Therefore the supervisor will report $Ey_f + b$ where Ey_f is the supervisor's best estimate of the firm's observation. This simplifies to

$$\max_b \quad E[\ \mu - \lambda(Ey_f - y_f)^2 - \lambda\sigma_b^2 + \eta w], \tag{24}$$

with the first-order condition for b characterized by $b = [\tau/\lambda]\eta$, so that the variance of bias, conditional on λ, is

$$\sigma_b^2(\lambda) = \left(\frac{\tau}{\lambda}\right)^2 \sigma_\eta^2. \tag{25}$$

(Note that in Sections I and II, λ was normalized to unity.)

Given the supervisor's incentive, it would appear that the optimal policy for the firm to follow is to increase λ from zero because $b \to 0$ as $\lambda \to \infty$. Therefore a potentially simple solution outlined to this agency problem is to penalize the supervisor for perceived transgressions in terms of showing

favoritism.[12]

There is one potentially important caveat, however, in that this strategy relies on the supervisor's report being invertible to determine his report. Assume that the contract is chosen so that b is negligible. Then the supervisor reports his best guess of the firm's observation which is given by

$$\tilde{y} = \frac{\sigma_0^2 y + \sigma_s^2 e*}{\sigma_0^2 + \sigma_s^2}. \tag{26}$$

where e* is the equilibrium effort level of the worker. Then as σ_i^2 and e* are commonly known, the firm can invert $\tilde{y}$ to induce y, the supervisor's true observation. If this is the case, then the potential exercise of bias results in no distortions and the first-best monitoring scheme can be implemented.

However, this critically relies on the assumption that $\sigma_0^2 > 0$, i.e., that the supervisor's observation carries some information for predicting the firm's observation, conditional on knowing the worker's effort in equilibrium. But if $\sigma_0^2 = 0$, the supervisor knows that the worker's effort is e^*, and his best guess of the firm's observation is also e^*, independent of his own observation. In other words, once the supervisor is moni-

[12]More generally, the firm can induce any "level" of bias by changing λ.

tored, he never uses his own information; he simply states what he knows the worker's effort to have been.

This problem arises only because the error terms in monitoring the worker are uncorrelated across the supervisor and the firm, in that there is no common information in the reports of the supervisor and the firm for this problem to arise, aside from deterministic effort. But in most scenarios, this seems unlikely, as whatever bad luck pervades the performance of an employee is likely to be at least imperfectly observed by both monitors. This is the case where $\sigma_0^2 > 0$, so that the assumption of unknown ability plays an important role in the analysis. Therefore, subject to this caveat, the firm can use multiple observations to induce the first-best monitoring scheme. The remainder of this section is devoted to describing realistic extensions of the model that imply this is an unduly optimistic impression.

3.2. Problems with Monitoring Supervisors

A. Risk Aversion

The solution outlined above clearly involves considerable risk for the supervisor. In effect, the limiting result described above rewards the supervisor well if the supervisor's report is within some tight bounds of the firm's observation. Otherwise, the supervisor is poorly paid. This can obviously

involve considerable deadweight losses if the supervisor is risk averse.

This issue is well known in the agency literature and there appears little purpose in formally modeling these costs. Instead, the main focus of this section will be to illustrate alternative constraints on organizations monitoring supervisors who have private information.

B. Yes Men

The way we have described monitoring supervisors is that when firms decide to use multiple observations to reward the supervisor, it gives supervisors more incentive to tell the truth. However, as alluded to above, this is not quite precise, as it really gives supervisors an incentive not to be *caught* lying, which is not the same as telling the truth. This section extends Prendergast (1993) to illustrate how offering supervisors incentives of the type described above can result in the creation of yes men, who, rather than tell the truth, act merely to mimic the opinion of their superior.

This incentive arose above where the supervisor reported his estimate of the manager's observation. However, a key feature of the analysis above is that the supervisor's report can be inverted to *uniquely* determine y. In this section, we illustrate that if the supervisor can invest resources to

determine a noisy measure of the manager's opinion, this result is no longer true.

In particular, assume that the supervisor, in addition to observing y, also observes a noisy statistic on the manager's observation via

$$y_\omega = y_f + \omega, \tag{27}$$

where $\omega \sim N(0, \sigma_\omega^2)$. There are many realistic interpretations of y_ω. For example, a senior manager may have certain priorities, such as emphasizing "quantity" over "quality" or the manager may himself have preferences towards certain workers himself.

We assume that the supervisor needs to expend some resources in determining the manager's views. We also assume that $\sigma_\omega^2 = g(e_s)$, where e_s is effort exerted by the supervisor at cost $d(e_s)$, where $d' > 0$, $d'' > 0$, $d'(0) = 0$, and $d'(\infty) = \infty$. We assume that $g' < 0$, $g'' > 0$, and $g(0) = \infty$. Therefore the supervisor can expend effort to determine what his manager would like to hear, but he has no idea of his manager's preferences unless he exerts some effort ($g(0) = \infty$).[13]

---

[13]It should also be clear that the supervisor's observation on the worker could depend on effort exerted. This issue is addressed at length in Prendergast (1993) and is ignored here for simplicity.

Consider the supervisor's estimate of the manager's opinion, conditional on his effort decision. Given y, y_ω, and knowledge of the worker's incentives (which yield e*), the supervisor's estimate of y_f is given by a normal distribution with mean

$$\hat{y}_f = \mu_0 y + \mu_1 e_* + \mu_2 y_\omega, \tag{28}$$

where

$$\mu_0 = \frac{\sigma_0^2 \sigma_\omega^2}{\sigma_0^2\sigma_s^2 + \sigma_0^2\sigma_f^2 + \sigma_f^2\sigma_s^2 + \sigma_\omega^2(\sigma_0^2 + \sigma_s^2)}, \tag{29}$$

$$\mu_1 = \frac{\sigma_s^2 \sigma_\omega^2}{\sigma_0^2\sigma_s^2 + \sigma_0^2\sigma_f^2 + \sigma_f^2\sigma_s^2 + \sigma_\omega^2(\sigma_0^2 + \sigma_s^2)}, \tag{30}$$

and

$$\mu_2 = 1 - \mu_0 - \mu_1. \tag{31}$$

Thus when the firm monitors the supervisor's performance, the supervisor reports a convex combination of his three pieces of information. This in itself is not necessarily a problem. Consider the case where $\sigma_\omega^2 = \infty$, (i.e., where $e_s = 0$), so that the supervisor gets no indication on the manager's opinion. Then as $\mu_2 = 0$, the supervisor reports $\tilde{y}$ as in equation (26) so that y can be perfectly inferred.

The interesting problem arises when $\sigma_\omega^2 < \infty$ or $e_s > 0$. Then the worker's report depends on two variables that are unknown to the manager, y and y_ω, but only y is relevant for

determining the true value of performance. To see this, note that the manager knows μ_i, for all i, in equilibrium so that the uncertainty facing the manager over the true value of y is given by the random variable

$$z = \mu_0 y + \mu_2 y_\omega. \tag{32}$$

The conditional distribution of y given z is normal with mean y and variance $[\mu_2/\mu_0]^2\sigma_\omega^2 > 0$, so the existence of "yes men" introduces greater uncertainty than when the supervisor reports honestly. If the supervisor reports honestly, the variance of $\hat{y}$ is given by

$$V* = \frac{\sigma_s^2\sigma_f^2\sigma_0^2}{\sigma_f^2\sigma_0^2 + \sigma_f^2\sigma_s^2 + \sigma_0^2\sigma_s^2}, \tag{33}$$

But as the manager can only invert the supervisor's report to determine z, the estimate of the variance of performance is given by

$$\frac{\left[\sigma_s^2 + \left(\frac{\mu_2}{\mu_0}\right)^2\sigma_\omega^2\right]\sigma_f^2\sigma_0^2}{\left[\sigma_s^2 + \left(\frac{\mu_2}{\mu_0}\right)^2\sigma_\omega^2\right][\sigma_f^2 + \sigma_0^2] + \sigma_f^2\sigma_0^2} > V* \tag{34}$$

so that monitoring by the firm causes the supervisor to monitor less effectively. Note that the costs of monitoring are strictly decreasing in σ_ω^2 as $[\mu_2/\mu_0]^2\ \sigma_\omega^2$ is decreasing in σ_ω^2.

Our purpose is not simply to show how this incentive for conformity can result in efficiency losses. Instead we study the implication of this incentive for how the firm should compensate the supervisor. As above, we assume that the supervisor's compensation function is given by (22), which equals

$$w_s = \mu - \lambda(\mathrm{E}y_f - y_f + b)^2. \tag{35}$$

Routine calculations show this equals

$$w_s = \mu - \lambda(\mu_0^2\sigma_s^2 + (\mu_2-1)^2\sigma_f^2 + \mu_2^2\sigma_\omega^2 + b^2). \tag{36}$$

The supervisor then chooses e_s to

$$\max_{e_s,b}\ \mu - \lambda(\mu_0{}^2\sigma_s{}^2 + (\mu_2-1)^2\sigma_f{}^2 + \mu_2{}^2\sigma_\omega{}^2 + b^2) + \eta w - d(e_s) \tag{37}$$

subject to $\partial w/\partial b = \tau$. It is straightforward to show that $\partial w_s/\partial e_s > 0$ because more effort implies that the supervisor can better predict his manager's opinion. However, of more importance for our analysis is that $\partial e_s/\partial\lambda > 0$ so that as the manager imposes greater costs on the supervisor for perceived

infractions, the supervisor responds by spending more time on unproductive activities designed to determine the manager's opinion. In this sense, the tendency towards conformity increases as the supervisor's compensation depends more upon it. On the other hand, it is straightforward to show that $\partial\sigma_b^2/\partial\lambda < 0$. The choice of λ is therefore not as simple as in the previous section, where its only implication was to reduce favoritism. Here, there is a countervailing incentive where the supervisor responds by distorting his report more and more, and expending more resources to determine his superior's opinion. More formally, the firm's profits fall as σ_ω^2 falls, both because the monitoring mechanism becomes less precise, as described above, and because the effort exerted by the supervisor must ultimately be paid for by the firm through the supervisor's IR constraint. For these reasons, it is easy to see that it is no longer optimal to choose $\lambda = \infty$ but, instead, the firm trades off the costs of favoritism with the costs of conformity.

C. Multi-tasking

The model described above assumes that the only activity carried out by the supervisor is to provide performance evaluations. This is clearly a restrictive assumption; supervisors typically carry out a host of activities that potentially

affect the welfare of their subordinates. For example, supervisors provide on-the-job training for employees, assign them to tasks with greater or smaller returns to talent, decide on layoffs, and determine work schedules. The purpose of this section is to describe how the firm will typically restrict its punishment of the supervisor for perceived infractions over favoritism in order to restrict substitution of harmful activities on other dimensions.

To take a specific example, assume that the supervisor provides on-the-job training to the worker, in addition to providing performance evaluations. Assume further that the firm cannot monitor training; all that it can do is penalize the supervisor for perceived favoritism in performance evaluations. (In other words, the supervisor can be penalized only if it is felt that he has lied.) Then consider the behavior of a supervisor who dislikes his subordinate ($\eta < 0$). He has two instruments by which he can harm the disliked worker: either lie on performance evaluations or make sure that the worker's performance is poor by denying him training opportunities. The key point here is that as the firm increases the punishment on perceived lies (as λ increases), the supervisor is likely to substitute into activities that makes his report truthful. He can sabotage his subordinate's performance. This constitutes a cost for the firm from punishing the supervisor and results

in more lenient penalties on the supervisor.

Formally illustrating this result requires a considerable extension. First, we need to consider the impact of training on productivity. Second, the logic described above is predicated on the assumption that as the supervisor lies less on performance evaluations, his marginal incentive to sabotage increases. This requires an assumption of diminishing returns to harming the worker, which is not provided in the model above as the compensation function is linear and the supervisor's preferences are linear. Therefore, to formally model this phenomenon requires imposing strict concavity in the supervisor's preferences. If these extensions are made to the model, then it can be shown (Prendergast (1993)) that the supervisor's ability to sabotage worker's true output mitigates the firm's incentive to penalize infractions on performance appraisals. This conclusion accords with similar analysis by Holmstrom and Milgrom (1992) on multi-tasking.

IV. Tying Wages to Jobs

In large organizations, wages and salaries are commonly "tied" to a particular job or task, perhaps with some small discretionary range within tasks.[14] Wages are

[14]See, e.g., Groshen (1991) on data illustrating how wages are typically tied to job assignments.

associated with a particular task, independently of the identity of the worker performing that task. On the surface, this design seems to remove much of the firm's discretion in rewarding heterogeneous workers, which may harm incentives. We argue that tying wages to jobs may actually improve performance by removing supervisors' opportunities to exercise bias in setting pay. The point is related to our earlier discussion of bureaucratic rules in organizations.

The idea is that tying wages to jobs raises the supervisor's cost of inefficiently rewarding favorites. We assume that the supervisor's reward is based on the total output of his subordinates. When the supervisor has full discretion over both job assignments and compensation, he can dole out monetary rewards to favored workers while assigning the most talented workers optimally. This harms incentives. But when wages are tied to tasks, the output gains from optimal assignments are also tied to paying workers appropriately. Then favoritism in compensation carries a cost of inefficient job assignments. Therefore, tying wages to jobs can act as a mitigating factor in the exercise of supervisor discretion.

We model this by assuming that the supervisor is the full residual claimant of his subordinates' output. Then the real issue is commitment by the supervisor. Ex ante, the

supervisor would like to commit to efficient compensation, as this guides incentives and increases total output. Then strong reputations will eliminate the type of favoritism we model here. For this reason, we assume that reputation formation is prohibitively costly.

We return to the basic model of Section I with a single employee, but where the employment relation is for two periods, with no discounting. Here the employee can either be retained in his current job, job A, after period one, where the marginal return to ability is unity, or promoted to job B, where the marginal product of ability is $\gamma > 1$. For simplicity, it is assumed that in all other respects the jobs are identical. It is further assumed that the supervisor is residual claimant on output for the two periods. The supervisor is not monitored over his performance evaluations, for the reasons described in Section III. Note that on efficiency grounds the worker should be promoted to job B if her perceived ability exceeds 0 after period one.

4.1. Separating Sorting and Incentive Issues

Assume first that there is no requirement that wages change upon changing tasks. As there are no benefits to the worker (and consequently, the supervisor) from being promoted, the supervisor will always assign workers efficiently, as

he is residual claimant. Therefore, if the perceived ability of the worker exceeds 0, the worker will be promoted to job B. However, in the absence of a monitoring scheme, the supervisor cannot induce any effort from his subordinates. This follows immediately from backward induction. At the end of period 2, the supervisor will base his recommendations solely on the basis of his preference, so there is no incentive to work in period 2. But as there is no return or opportunity to the supervisor from developing a reputation, the supervisor will similarly act on his whims in period 1, despite the fact that he is residual claimant. Therefore, by backward induction, the worker never exerts effort, which clearly involves efficiency losses, and the firm should offer a wage independent of performance evaluations.[15]

4.2. Tying Incentives and Sorting

Assume that the firm develops an alternative organizational structure where the supervisor has no discretion about wages within a job, but can promote workers to a job with a higher wage. More specifically, assume that the organization designates a wage of L to job A and W to job B, where W > L. The issue at hand here is whether the firm can use such a

[15]This is clearly caused by the inability of the supervisor to develop a credible reputation for fairness. Similar qualitative results are likely to hold when reputation acquisition is costly.

procedure to induce more honesty from the supervisor.

A key aspect is that the supervisor now incurs a cost by offering a high wage to an undeserving worker because productivity falls. To see this, note that the supervisor's utility from offering job B to the worker rather than job A is $\eta W + \gamma\alpha_1 - [\eta L + \alpha_1]$, where α_1 is the perceived ability of the worker after period 1. Therefore, a job B assignment will be offered if

$$\eta > \frac{(1-\gamma)\alpha_1}{W-L}. \tag{38}$$

Suppose that the worker is sufficiently capable to be promoted on efficiency grounds. Then (38) implies that there is some probability that the worker will be promoted even if he is disliked by his supervisor.[16]

To show how a policy that ties wages to jobs can be beneficial to the firm, consider the worker's incentives. The worker is given the promotion if

$$\frac{(\gamma-1)\alpha_1}{W-L} + \eta > 0. \tag{39}$$

Note that

---

[16]Note that as in Section 1, as the return to promotion tends to zero, the first best allocation of workers to jobs arises.

$$\alpha_1 = \frac{\sigma_0^2(y - Ee)}{\sigma_0^2 + \sigma_\epsilon^2} \tag{40}$$

where Ee is expected effort, equal to e^* in equilibrium. This implies that in equilibrium the distribution of the LHS of (39) is a normal distribution with mean 0 and variance

$$\left(\frac{(\gamma - 1)^2}{(W - L)^2} \frac{\sigma_0^4}{\sigma_0^2 + \sigma_\epsilon^2} + \sigma_\eta^2 \right).$$

Let this distribution function be characterized by Φ, with density ϕ. The worker chooses e to

$$\begin{aligned} \max \quad & \left[1 - \Phi\left(\frac{\gamma - 1}{W - L}(\alpha + e - Ee) + \eta\right)\right] u(W) \\ & + \Phi\left(\frac{\gamma - 1}{W - L}(\alpha + e - Ee) + \eta\right) u(L) - c(e) \end{aligned}, \tag{41}$$

where $u(x)$ is the worker's utility from wage x. Her effort decision is characterized by

$$\left[\frac{(\gamma - 1)}{W - L} \frac{\sigma_0^2}{\sigma_0^2 + \sigma_\epsilon^2} \right] \phi(0)[u(W) - u(L)] = c'(e) \tag{42}$$

in equilibrium. As the LHS of (42) is positive (as $\gamma > 1$ and $\sigma_0^2 > 0$), the worker exerts positive effort. In other words, the

prospect of appearing talented results in the worker exerting effort despite the preferences of the supervisor. This result occurs *only because wages are tied to jobs*, which gives the supervisor a means of committing to acting honestly. It is in this sense that tying wages to jobs can enhance efficiency.

Note, however, the tradeoffs involved. By tying wages to jobs, workers do indeed exert effort but, from (38), at the cost of sometimes misallocating workers. This arises if the supervisor's preferences are so strong that he is willing to put up with the costs of misallocation. The alternative is not to provide incentives at all, which, though obviously costly, has the benefit of not involving misallocation. The optimal policy depends on parameter values, measuring the benefits of incentive provision relative to the benefits of correctly allocating workers to jobs.

Conclusion

This paper is an initial attempt to understand how organizations respond to favoritism. To deal with favoritism and bias, we argue that firms may rely on bureaucratic rules for compensation and promotion, they may use more interpersonal equality than would otherwise be efficient, and may tie wages to jobs rather than individuals. Here, these management strategies are meant offset the inefficiencies that

can be caused by the personal likes and dislikes of supervisors.

The main inefficiency elaborated on in the paper is that favoritism renders monitoring less efficient. The response of the firm is not simply a *reaction* to this additional noise; instead, the instruments described are designed on the realization that the efficiency of the monitoring system is *caused* by compensation schemes. It is in this sense that favoritism is a theoretically different construct from an inefficient monitoring scheme.

Before concluding, it is worthwhile to consider the importance of some maintained assumptions. We have assumed that workers do not know or act on supervisors' preferences. We think it unlikely that this significantly affects our results. Similarly, we have assumed that ability is unknown to all parties. But the key point for many of our findings is that the firm does not know workers' talents, so it must rely on the supervisor's reports. Overall, we feel that our results are robust to alternative specifications and assumptions.

The impact of favoritism goes well beyond the narrow confines that we have described. For example, in some cases personal relationships can enhance productivity. This aspect of productivity enhancement has been ignored here but presumably can make the monitoring problem described in Section III far more problematic; a manager showing favoritism

towards a worker may be doing so because it increases productivity. We have also ignored any role for reputations in reducing the impact of favoritism. Yet firms are often concerned with being "fair," or even appearing so, particularly in the context of the possibility of discrimination against women or minorities. How this affects behavior is unclear. For example, is this incentive for reputation building enough to eliminate favoritism? Or, can it lead to positive discrimination towards some otherwise unfavored group?

Our limited objective was to demonstrate that many issues of favoritism can be described in a typical agency framework, rather than relying on notions of fairness and reputation. In this way, our analysis provides a foundation for studying how organizations may deal with problems of bias.

Appendix

Proof of Lemma 1

Note that since $E\alpha|\alpha < 0$ is a constant, we need only consider prob($\alpha < 0|\ \hat{y}^1 > \hat{y}^2$) when considering the effect of compensation on sorting. $\hat{y}^1 > \hat{y}^2$ implies that $\alpha > -\mu$ where $\mu = \epsilon + b^1 - b^2$ and $\epsilon = \epsilon^1 - \epsilon^2$. Conditional on μ, this occurs with probability $1 - F_\alpha(-\mu)$, where F_α is the distribution of α. The probability that $\alpha < 0$, conditional on $\alpha > -\mu$, is $\min(0, F_\alpha(0) - F_\alpha(-\mu))$. However μ is unknown ex ante so that we must integrate over μ. So prob($\eta < 0|\hat{y}^1 > \hat{y}^2$) is

$$\min\left[0, \int_{-\infty}^{\infty} [F_\eta(0) - F_\eta(-\mu)]\,\partial F_\mu(\mu)\right].$$

If $\mu < 0$, errors are never made as worker 1 is harmed by the supervisor's report. Therefore,

$$\min\left[0, \int_{-\infty}^{\infty} [F_\alpha(0) - F_\alpha(-\mu)]\,\partial F_\mu(\mu)\right]$$
$$= \frac{1}{2} - \int_{0}^{\infty} F_\alpha(-\mu)\,\partial F_\mu(\mu).$$

Let $b = b^1 - b^2$. Then since $E(\epsilon + b) = 0$, increasing τ is equivalent to a mean preserving spread of the distribution. A

mean preserving spread of μ decreases

$$\int_0^\infty F_\alpha(-\mu)\partial F_\mu(\mu)$$

as more weight is placed in the tails of the μ distribution. Hence errors are made more frequently as more weight is placed in the tails of the distribution. □

References

Bretz, R., and Milkovich, G. "Performance Appraisals in Large Organizations: Practice and Research Implications," Working Paper No. 89-17, Center for Advanced Human Resources Studies, Cornell University, Ithaca, NY, 1989.

Bull, C., (1987), "The Existence of Self Enforcing Contracts," *Quarterly Journal of Economics.*

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

Holmstrom, B. and P. Milgrom (1987), "Aggregation and Linearity in Agency Contracts." *Econometrica.*

Holmstrom, Bengt, and Milgrom, Paul. "Multi-task Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design," *Journal of Law, Economics, and Organization* 7 (1991): 24-52.

Holmstrom, Bengt, and Milgrom, Paul. "Measurement Costs and Organization Theory: The Firm as a Cluster of Attributes," mimeographed, Yale University, 1992.

K. Kraiger and J. Ford (1985), "A Meta-Analysis of Ratee Race Effects in Performance Ratings," *Journal of Applied Psychology*, 70(1), 56-65.

F. Landy and J. Farr (1980), "Performance Rating," *Psychological Bulletin* 87(1), 72-107.

P. Milgrom and J. Roberts (1988), "An Economic Approach to Influence Activities and Organizational Responses," *American Journal of Sociology*, 94, S154-79.

P. Milgrom and J. Roberts (1990), "The Efficiency of Equity in Organizational Decision Processes," *American Economic Review*, 80(2), 154-59.

G. Milkovitch, et al. *Pay for Performance*. Washington, D.C.: National Academy Press, 1991.

A. Mohrman and E. Lawler (1983), "Motivation and Performance Appraisal Behavior," in F. Landy et al., eds., *Performance Measurement and Theory*, Hillsdale, NJ: Enbaum.

K. Murphy and J. Cleveland (1991), *Performance Appraisal: An Organizational Perspective*, Boston: Allyn & Bacon.

Prendergast, Canice. "A Theory of 'Yes Men'," *American Economic Review*, forthcoming, September 1993.

Prendergast, Canice (1993), "An Agency Approach to Bias in Organizations," mimeo, University of Chicago.

Prendergast, Canice, and Topel, Robert H. "Discretion and Bias in Performance Appraisal," forthcoming, *European Economic Review*, 1993.

C. Reilly and W. Balzer (1988), "Effects of Purpose on Observation and Evaluation of Teaching Performance," mimeographed, Bowling Green University.

Tirole, J. (1992), "Collusion and the Theory of Organizations," in J. Laffont (ed.), *Advances in Economic Theory*, Vol. II, Cambridge University Press.

K. Williams, P. Wickert, and R. Peters (1985), "Appraisal Salience: Effects of Instructions to Subjectively Organize Information," in *Proceedings of the Southern Management Association Meetings*, Orlando, FL.