Review of Economic Studies (2001) 68, 45–66 0034-6527/01/00030045$02.00

Performance, Promotion, and the Peter Principle

JAMES A. FAIRBURN
University of Sussex

and

JAMES M. MALCOMSON
University of Oxford

First version received May 1997; *final version accepted June* 2000 (*Eds.*)

This paper considers why organizations use promotions, rather than just monetary bonuses, to motivate employees even though this may conflict with efficient assignment of employees to jobs. When performance is unverifiable, use of promotion reduces the incentive for managers to be affected by influence activities that would blunt the effectiveness of monetary bonuses. When employees are risk neutral, use of promotion for incentives need not distort assignments. When they are risk averse, it may—sufficient conditions for this are given. The distortion may be either to promote more employees than is efficient (the Peter Principle effect) or fewer.

> ''*Promotions serve two roles in an organization. First, they help assign people to the roles where they can best contribute to the organization's performance. Second, promotions serve as incentives and rewards*.'' (Milgrom and Roberts (1992, p. 364))

> ''*Promotions are used as the primary incentive device in most organizations, including corporations, partnerships, and universities ... This ... is puzzling to us because promotion-based incentive schemes have many disadvantages and few advantages relative to bonus-based incentive schemes*.'' (Baker, Jensen and Murphy (1988, p. 600))

1. INTRODUCTION

The dual role for promotions, to assign people to the jobs that best suit their abilities and to provide incentives, is part of a well-established tradition in labour economics. Empirical studies of individual firms by Rosenbaum (1984), Lazear (1992) and Baker, Gibbs and Holmström (1994*a*, 1994*b*), provide evidence consistent with that tradition. But, as the quote above from Baker, Jensen and Murphy indicates, promotions have many disadvantages as incentive schemes, disadvantages discussed both in that paper and in Milgrom and Roberts (1992). Perhaps the most obvious is the potential conflict between the two roles for promotions. If a firm provides incentives by promoting those who have performed well in a job, it may simply transfer them to a job to which they are not well suited, a mild version of the Peter Principle which in its original version (Peter and Hull (1969)) took the form ''people are promoted to their levels of incompetence''. Why not then use promotions to assign employees to jobs and monetary bonuses to provide incentives? And, if promotions have to be used to motivate employees, need the Peter Principle or other distortions necessarily result? These are the questions that motivate this paper.

The essence of our answer to the first question is as follows. Organizations, particularly large ones, must typically delegate assessments of individual performance to managers. The incentives of those managers may not be perfectly aligned with the goals of the organization, so they may misuse monetary bonuses. Bonuses are a reward to employees for work already done—to the employer, giving bonuses to those who perform best is important only to maintain the long term reputation of the organization for rewarding good performance. A manager who may move on may not care much about the long term loss of the organization's reputation that results from misallocation of bonuses, especially if it takes time for employees to assess how well they have performed, realize what is going on, and adjust their behaviour accordingly. Employees, however, care a lot about whether they get a bonus. In reporting performance, the manager may thus be more influenced by other considerations, whether employees offer favours, indulge in other types of influence activities discussed in Milgrom (1988) and Milgrom and Roberts (1988), or are simply favourites as in Prendergast and Topel (1996). Monetary bonuses are particularly susceptible to this. In the extreme, if employees can simply bribe the manager to give them a bonus, the incentives the bonus scheme is intended to provide are completely destroyed.

Use of promotions to provide incentives can mitigate this effect. As long as managerial rewards are tied in some way to the short term performance of the section for which the manager is responsible, the manager has an interest in ensuring that the appropriate employees are promoted. That reduces the incentive to succumb to influence activities when allocating employees to jobs with different pay. Indeed, we show that, if the manager's reward is made sufficiently sensitive to the performance of the section next period, the organization can eliminate such influence activities completely. Moreover, if employees are risk neutral, it can do that without compromising either efficient effort by employees or efficient assignment to jobs. If employees are risk averse, there will typically be some distortion in assignments but that may be a small price to pay to improve incentives for performance in the face of influence activities. We investigate when this distortion takes the form of promoting employees who would not be promoted for assignment reasons alone, the Peter Principle effect.

The standard principal–agent approach to providing incentives for performance is to use a contract with performance related pay. See Hart and Holmström (1987) for a survey. That approach, however, relies on employee performance being a suitable basis for a legally enforceable contract. In many real world applications, this is simply not the case. Performance is often not *verifiable* in the manner required: for reasons explored in Holmström and Milgrom (1991), it is too complex or multidimensional to be written unambiguously into a contract that could be enforced by the courts.

Non-verifiability of performance is not in itself sufficient to account for the use of promotions. Even without a contract, employers may pay bonuses for good performance because of their concern for reputation and so induce employees to exert effort beyond what is minimally enforceable. Since promotions are used widely for incentives even in large organizations that develop reputations in other respects, why not in regard to personnel policy? Moreover, as in the repeated principal–agent models of Radner (1981, 1985), such arrangements can be structured to provide risk averse employees with insurance. An employer's concern for reputation can also support a promotion structure that both provides incentives for performance and sorts employees by ability as shown in MacLeod and Malcomson (1988), though in a model with all employees doing the same tasks and promotion merely sorting them into ranks, so there is no potential for conflict between incentives and assignment to tasks.

In practice, however, reputation can at best enforce only what can be controlled by those with a concern for it. Owners, or senior managers, can presumably check reasonably easily what pay scales are, how many employees are in each job category and grade, and how many have been promoted in the last year, by calling the personnel division if necessary. Their concern for the organization's reputation can then be relied upon to ensure that it provides a promised number of bonuses or promotions, and that salaries or bonus payments are at agreed levels. What is much harder to check is how well any individual employee has performed. In practice, organizations have to rely for this information on the reports of more junior managers and supervisors, who may have less concern with the organization's long term reputation.[1] If the role of junior managers could be fully duplicated by senior management, there would be no reason to use junior managers in the first place. In practice there are tiers of management, and the behaviour of managers must be taken into account.[2]

Tournaments are an alternative to reputation as a means of ensuring that firms reward good performance when performance is unverifiable. The literature on tournaments stems from the work of Lazear and Rosen (1981), Holmström (1982), Green and Stokey (1983) and Nalebuff and Stiglitz (1983). With a standard principal–agent contract but non-verifiable performance, the employer faces an incentive to renege on the contract *ex post* by claiming that employee performance is at a level meriting only the lowest wage permitted by the contract. When there is more than a single employee and the employer uses a tournament, that option is removed. As observed by Malcomson (1984) and Bhattacharya (1983), and in a rather different context by Carmichael (1983), tournaments have the attractive property of fixing the wage bill. Some employee must have achieved the best performance level, some other the second best, and so on and, since the employer must pay each wage to some employee, there is no reason why the prizes should not be awarded according to the true performance ranking.

Promotions are one way to implement tournaments in practice. Lazear and Rosen (1981), Malcomson (1984), Rosen (1986) and Baker, Jensen and Murphy (1988) all explore the use of promotions in this way for incentive purposes. Yet, in principle, tournaments are more flexible than this. Prizes in a tournament can take the form of monetary bonuses that do not even potentially conflict with efficient assignment to jobs. Moreover, Malcomson (1986) showed that, for large numbers of employees, tournaments with monetary prizes can under appropriate conditions replicate the effects of an optimal performance related pay contract of the standard principal–agent type. Thus the existing literature does not provide a reason why firms should use promotions, with their potential conflict with efficient assignment, rather than monetary bonuses to implement a tournament.

Promotions emerge in the present paper because, with non-verifiable performance, the implementation of performance related incentive structures is delegated to managers and, as indicated above, a promotion scheme provides a way of disciplining managers that a bonus scheme does not. Characteristics of promotions important for this are (1) that managers have performance related rewards which ensure they care whether those promoted are the most suitable for the higher level jobs, and (2) that the firm attaches pay increases to promotions. Use of promotions to provide incentives has been studied by Kahn and Huberman (1988), Waldman (1990), Prendergast (1993), Fairburn and Malcomson (1994) and Mori (1998) but with no manager susceptible to influence activities and hence no role for promotions in disciplining managers. Prendergast and Topel (1996)

1. Although both senior and junior managers are finitely lived, the reputation of the former can be transferred on, for example by the means discussed by Kreps (1990), much more plausibly than is the case with the latter.

2. Incentive contracting in a multi-tiered organization without promotions is studied by Tirole (1986). Subsequent contributions on collusion in organizations are reviewed by Tirole (1992) and Laffont and Martimort (1997).

include in their analysis a manager who, for entirely exogenous reasons unrelated to productivity, favours particular employees and thus distorts performance evaluations and assignment decisions. The present analysis recognizes that a manager may distort assignment decisions because of endogenously created influence activities that do not depend on inherent favouritism. A distinct literature on the wage increases following from promotion stems from Waldman (1984) and has been developed by Bernhardt and Scoones (1993) and Bernhardt (1995). In those papers, assignment to a higher level job reveals favourable information about an employee's productivity to other potential employers, resulting in a large wage increase that induces the current employer to distort assignments. Gibbs (1995) discusses the incentive effects such promotions provide. There wage incentives are a by-product of job assignment that reveals information, whereas here job assignment is used intentionally as a way of providing wage incentives. For further discussion of these issues, see the surveys by Gibbons and Waldman (1999) and Prendergast (1999).

In the next section, we set out the model. In Section 3, we show that performance related pay is an ineffective means of providing incentives in an organization in which employees can bribe managers with impunity. The same holds true for a tournament with monetary prizes. In Section 4 we shift attention to promotion schemes for employees who differ in their suitability for different jobs and for whom initial performance provides a signal of that suitability. We show that use of promotions can dramatically improve the provision of incentives. Subsequent sections consider the consequences of using promotions to provide incentives. In Section 5 we examine the properties of promotion schemes when employees are risk neutral. In Section 6 we extend the analysis to risk averse employees and examine the conditions under which the Peter Principle applies. Section 7 concludes.

2. THE MODEL

A firm hires N employees from a population comprising high (H) and low (L) ability types in proportions q and $1-q$ respectively. Employees work for 2 periods, $t = 1, 2$. The utility of an employee in period t from receiving wage w_t and exerting effort $a_t \in [0, \bar{a}]$ is $u(w_t) - v(a_t)$. $u(w_t)$ is continuously differentiable and satisfies $u'(w_t) > 0$ and $u''(w_t) \leqq 0$. The disutility of effort $v(a_t)$ is continuously differentiable and satisfies $v'(a_t) > 0$, $v''(a_t) > 0$ and

$$v(0) = 0; \qquad \lim_{a_t \to 0} v'(a_t) = 0; \qquad \lim_{a_t \to \bar{a}} v'(a_t) = \infty. \tag{1}$$

The assumptions in (1) rule out uninteresting corner solutions.

The firm is risk neutral and has two types of job, h and l. The performance, x_t, of an employee of type $i \in \{L, H\}$ in job $j \in \{l, h\}$ in period t is

$$x_t = a_t + z_t, \qquad \text{for } t = 1, 2, \tag{2}$$

where z_t is a random term that is independently distributed for each employee and has mean s_j^i. This formulation implies a unique efficient effort level $\tilde{a}$ defined by $v'(\tilde{a}) = 1$. High ability employees are best suited to job h and low ability to job l, that is, $s_l^H < s_h^H$ and $s_l^L > s_h^L$. An employee's performance in each period is observed by the firm's manager but (unlike in Holmström (1999), Gibbons (1986), Gibbons and Murphy (1992) and Fairburn and Malcomson (1994) who use related models) not by other employees, the firm or any outside party. We consider both the case in which it is observed by the employee and the case in which it is not. At the time of hiring, the ability of an employee is unknown to firm, manager and employee alike. The expected performance of an unscreened worker is higher in job l than in job h, so unscreened employees are optimally assigned to job l. That is, given the proportion q of the H type in the population,

$$qs_l^H + (1-q)s_l^L > qs_h^H + (1-q)s_h^L. \tag{3}$$

First period performance in job l provides information about any given employee's type and thus his or her suitability for reassignment in the second period. To simplify notation, let z without a subscript denote z_1 and $f(z|s_j^i)$ its density function. We assume $f(z|s_l^i)$ is continuously differentiable in z and has the monotone likelihood ratio property (MLRP) $\partial[f(z|s_l^H)/f(z|s_l^L)]/\partial z > 0$ or, equivalently,

$$\frac{f'(z|s_l^H)}{f(z|s_l^H)} > \frac{f'(z|s_l^L)}{f(z|s_l^L)}, \tag{4}$$

so better performance in job l indicates an employee better suited to job h. Given effort a_1 by unscreened employees in period 1, the density function for performance x_1 is

$$g(x_1 - a_1) \equiv q\,f(x_1 - a_1|s_l^H) + (1-q)f(x_1 - a_1|s_l^L). \tag{5}$$

Let $G(\cdot)$ and $F(\cdot|\cdot)$ denote the distribution functions corresponding to $g(\cdot)$ and $f(\cdot|\cdot)$ respectively. Knowing $z = x_1 - a_1$ (that is, actual performance less the effort level), the manager updates the probability of each particular employee being type H from q to $q(z)$ according to Bayes' Rule:[3]

$$q(z) = \frac{q\,f(z|s_l^H)}{q\,f(z|s_l^H) + (1-q)f(z|s_l^L)}. \tag{6}$$

If the supports of $f(z|s_l^H)$ and $f(z|s_l^L)$ are disjoint, then the performance signals are fully informative about type and $q(z)$ equals either 0 or 1 for all z. If the supports of $f(z|s_l^H)$ and $f(z|s_l^L)$ are the same, then $0 < q(z) < 1$ for all z. Supports may also be partially overlapping: in this case $q(z) = 0$ or 1 for some performance signals and $0 < q(z) < 1$ for others. We assume that the second case applies, denote the common support by $[\underline{z}, \bar{z}]$ (it may be that $\underline{z} = -\infty$ and/or that $\bar{z} = \infty$) and define $\underline{q} \equiv q(\underline{z})$ and $\bar{q} \equiv q(\bar{z})$. The monotone likelihood ratio property (4) ensures that $q'(z) > 0$ for all $z \in [\underline{z}, \bar{z}]$.

Given z in period 1, an employee's expected performance in period 2 in jobs h and l is given by

$$h(z) = q(z)s_h^H + [1-q(z)]s_h^L, \tag{7}$$

$$l(z) = q(z)s_l^H + [1-q(z)]s_l^L, \tag{8}$$

respectively. Clearly $s_l^L \leqq l(z) \leqq s_l^H$ and $s_h^L \leqq h(z) \leqq s_h^H$ for all $z \in (\underline{z}, \bar{z})$. Let $R(z)$ be the expected productivity gain from reassigning an employee with first period performance z from job l to job h

$$R(z) = h(z) - l(z) = (s_h^L - s_l^L) + [(s_h^H - s_l^H) - (s_h^L - s_l^L)]q(z). \tag{9}$$

Given MLRP, $R'(z) > 0$: the higher an employee's first period performance, the greater the gain from reassigning that employee from job l to job h. The performance level that places an employee on the margin between being efficiently assigned to job l and to job h

3. Although unable to observe effort levels, the manager knows the employees' utility function and the contract they face, and can thus calculate the chosen effort.

is $\tilde{z}$ defined by

$$R(\tilde{z}) = h(\tilde{z}) - l(\tilde{z}) = 0. \tag{10}$$

To ensure that it is efficient to assign some employees to each job, we assume $\underline{z} < \tilde{z} < \bar{z}$.

The firm's manager also works for two periods, is risk neutral and has pay related to the firm's performance. In other contexts, performance-related pay might be used to induce the manager to put effort into monitoring employees. Its purpose here is to reduce the manager's temptation to be bribed or otherwise influenced into accepting activities that reduce profits. Given risk neutrality, there is no loss from making managerial compensation in each period linear in money. It is useful for what follows to allow it to depend separately on the revenue and wage components of profits and on the number M_t of employees the manager assigns to job h in period t, thus giving managerial utility in period t

$$V_t = \gamma_t + \alpha(\text{total revenue}_t - kM_t) - \beta(\text{wage bill}_t) + \text{total bribes}_t, \qquad \text{for } t = 1, 2, \tag{11}$$

for some constants α, β, k and γ_t for $t = 1, 2$. If $\beta = \alpha$ and $k = 0$, the performance related part of the manager's reward is simply a share of profits. The results that follow are unaffected by imposing the same α, β and k for both periods.

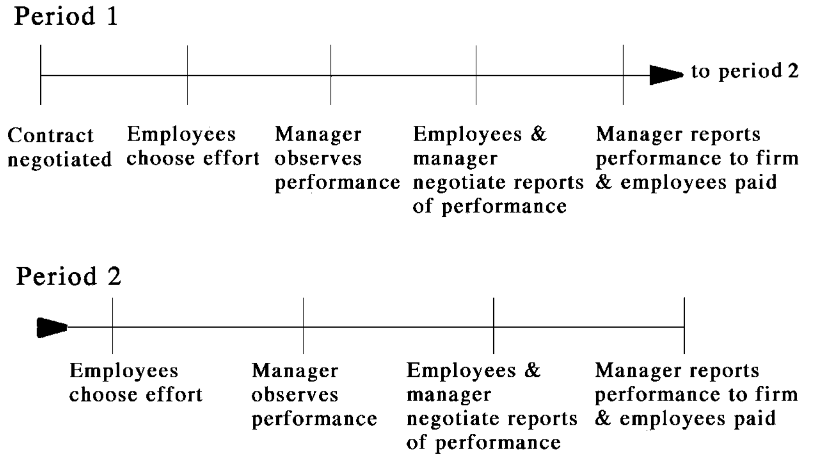

FIGURE 1

Timing of events

The timing of the model is illustrated in Figure 1. Initially a "contract" is agreed, stipulating payment to employees according to reports on performance made by the manager to the firm. (The quotation marks indicate that the contract need not be legally enforceable but, if not, is instead, as typically the case in practice, enforced by the reputation of the firm. We do not, however, model that reputation explicitly.) Precise forms of contract are discussed below. Employees then choose their first period effort levels, after which first period performance is observed by the manager. Following that, manager and employees negotiate over the performance reports the manager makes to the firm. Negotiation takes the form of employees simultaneously offering bribes for specified performance reports, which the manager can either accept or reject before making reports to the centre. Bribes are not detected by the firm. Employees are then paid according to the contract, given what the manager has reported. These steps, with the exception of the initial contract agreement, are repeated for period 2.

3. INCENTIVES NOT BASED ON PROMOTION

Suppose the firm sets wages in period 2 that are independent of the job to which an employee is assigned, so there is no role for promotion in providing incentives for performance in period 1. To provide incentives, it might use individual performance related pay based on the manager's reports of performance. Pay is then a function of the manager's report of performance in the current period and, in period 2, possibly also the previous period. Suppose the firm sets $k = 0$. (This is not important for the conclusions.) Consider first period 2 for a worker with performance y_1 already reported for period 1 and let $w_2(y_1, y_2)$ denote the pay the firm makes to the employee in period 2 when the manager reports that employee's period 2 performance as y_2. At the time the manager makes that report, employee effort for the period has already been determined, so what the manager reports makes no difference to output in that period. Moreover, the manager no longer works after period 2 and has no concern about the future implications of the reported performance. Hence, in the absence of bribes and any constraints on the set of permitted reports, the manager always reports a performance level y_2 that minimizes $\beta w_2(y_1, y_2)$. Let $\underline{y}_2(y_1)$ denote such a report. In order to receive a more favourable report, $\tilde{y}_2(y_1)$ say, where $w_2(y_1, \tilde{y}_2(y_1)) > w_2(y_1, \underline{y}_2(y_1))$, the employee must offer a bribe $b(y_1, \tilde{y}_2(y_1))$ which at least compensates the manager for not reporting $\underline{y}_2$. The wage gain to the employee from having $\tilde{y}_2$ reported is $w_2(y_1, \tilde{y}_2(y_1)) - w_2(y_1, \underline{y}_2(y_1))$. The loss to the manager in profit related pay from reporting $\tilde{y}_2$ is $\beta[w_2(y_1, \tilde{y}_2(y_1)) - w_2(y_1, \underline{y}_2(y_1))]$. Thus, there is scope for a mutually acceptable bribe if and only if

$$(1-\beta)[w_2(y_1, \tilde{y}_2(y_1)) - w_2(y_1, \underline{y}_2(y_1))] \geqq 0. \tag{12}$$

If $\beta > 1$, there is no mutually acceptable bribe: what the employee is prepared to pay is insufficient to compensate the manager for the loss in profit related pay. In this case, the manager simply reports the level $\underline{y}_2(y_1)$ for all employees, no matter what their actual performance levels. If by contrast $\beta < 1$, there is a mutually acceptable bribe. The joint gains are then largest if the manager reports a performance that maximizes $w_2(y_1, y_2)$. Whatever the bribe that divides those gains, it is in the interests of both for the manager to make such a report so, again, reported performance is independent of actual performance. Thus, in both these cases, what the manager reports to the firm is independent of the employee's actual performance and this removes any incentive effect from the performance related pay system. Finally, if $\beta = 1$, the only mutually acceptable bribe is one that exactly equals the difference in the wage. Thus the net payoff (wage less bribe) to the employee will be $w_2(y_1, \underline{y}_2)$ whatever the employee's actual performance. In that case too, therefore, the incentive effects of performance related pay are completely removed. A similar argument applied to period 1 establishes that reported performance in period 1 is independent of actual performance in period 1 and, hence, there are no incentives for performance in period 1 either. What drives this argument is simply the finite lifetime of the manager resulting in reputation effects unravelling in the standard way as a consequence of backward induction. The same result holds for any number of periods in the manager's working lifetime as long as it is finite.

Individual performance related pay thus provides no incentives for employees to exert effort when there are no limitations on bribes. By varying β, the parameter linking the manager's pay to the wage bill, the firm can alter the amount of bribes changing hands, but no level of β induces the manager to report performance truthfully. The firm cannot do better than set $\alpha > 0$ and $\beta = 0$, in which case the manager assigns employees efficiently to h and l jobs in period 2, but with no incentives for effort.

If individual performance related pay fails to provide incentives, would a tournament fare any better? With a tournament, the manager is required to rank the performance of employees in each period and the contract specifies a set of N wages, $w_1 \geqq w_2 \geqq \cdots \geqq w_N$, which are paid to employees in each period according to the performance evaluations submitted by the manager for that period. The advantage traditionally associated with such a scheme is that the wage bill in each period $\sum_{i=1}^{N} w_i$ is fixed, so profits cannot be increased by misrepresenting performance, see Malcomson (1984). With a manager susceptible to bribes, however, a tournament scheme of this type is no more effective at providing incentives for effort than individual performance related pay. The manager simply allocates more favourable reports to employees offering higher bribes. In a way similar to an auction with known common values, competition between employees in offering bribes then ensures that all employees' wages net of bribes are equal to the lowest wage on offer and thus independent of the actual performance achieved. Thus, as with individual performance related pay, there are no incentives for effort.[4]

Even if the firm were able to detect and punish bribery sufficiently to prevent money changing hands, other forms of influence activity can have a similar effect. Suppose employees offer time-consuming favours valued by the manager to influence the manager's ranking of performance and the manager ranks employees according to the favours provided. Unlike bribes, employees necessarily incur the cost of these favours whatever ranking they ultimately receive. The influence process with a tournament then becomes an *all pay auction* with multiple prizes. Such auctions are discussed by Clark and Riis (1998). Just as with monetary bribes, an employee's payoff net of the cost of influence activities does not depend on job performance, so there are no incentives for effort.

Alternatively, suppose the firm can detect whether the manager responds to bribes or influence activities with some positive probability less than one. If the maximum penalty that can be imposed for being caught is limited (for example, to dismissal), the fear of being caught may deter the manager from responding to influence activities when the gains are small but may not do so when the gains are large. Formally, if the present discounted value of the manager's utility is V if not bribed, V^b (excluding the bribe) if bribed and not caught, and $\bar{V}$ (excluding the bribe) if bribed and caught, the latter occurring with probability p, the manager will not accept bribes of less than $\underline{b} = V - V^b + p(V^b - \bar{V})$. In that case, the firm can use small monetary bonuses without the incentive effect being removed by bribes but may not be able to offer all the incentives it would like if that would require large bonuses. There may then be reasons to use the additional incentives that we show can be provided by promotion.

4. PROMOTION BASED INCENTIVES

We showed in the previous section how individual performance related pay and tournaments fail to provide incentives for employees if managers can accept bribes without risk of being caught. In this section we switch attention to incentives provided by promotion. Promotion schemes take two forms. In the first, promotion results in higher pay (and possibly higher status) but no change in job responsibilities. In the second, promotion also involves doing a different job. With the first form, it makes no difference to the

4. Details are in a working paper version available from the authors on request.

manager's future profit related pay which employees are promoted, so promotion is just a different form of monetary bonus (or prize in a tournament, if the number of promotions is specified in advance) whose incentives for effort can be undone by bribes or other influence activities as in the previous section. In contrast, the second form of promotion can, even with unrestricted bribery, induce employees to exert effort *provided employees differ in their suitability for different jobs*.

To see the incentives that promotion to a different job can provide, consider as an alternative to ineffective monetary bonuses the following contract between the firm and its employees. All employees are hired into job l for period 1 and paid wage w_1. On the basis of the manager's reports of performance, employees are then either kept in job l for period 2 and paid w_l, or promoted to job h and paid w_h, with $w_h > w_l$. With no prospect of subsequent promotion, second period effort will be at the minimum level, 0. We know from the previous section that performance related bonuses in period 2 will not change effort. Given that, we can without ambiguity use a without a subscript to denote first period effort.

The firm's freedom to set wages may be constrained by what other potential employers observe. If, as in Waldman (1984), Gibbons (1986), MacLeod and Malcomson (1988), Ricart i Costa (1988) and Bernhardt and Scoones (1993), other potential employers can observe the level to which an employee is assigned, there may be a lower limit on the wage w_h the firm can pay if it is to retain any promoted employees. This lower limit will depend on how many employees it promotes. For the present, we assume for simplicity that other employers cannot observe the contract offered by the firm or assignments within it, so the only constraint on second period wages is $w_l \geqq \underline{w}$, which ensures that the lowest paid employees cannot be compelled to continue working for the firm for less than the exogenously given wage $\underline{w}$ they can obtain outside. We return to this issue later.

Once first period performance levels have been realized, employees can bribe the manager in return for promotion. Given the reward function (11), the manager will recommend for promotion an employee with first period performance z provided that employee offers a bribe b that satisfies

$$b + \alpha[R(z) - k] - \beta(w_h - w_l) \geqq 0, \tag{13}$$

where $R(z)$, defined in (9), is the expected revenue gain from promoting an employee with performance z, β is the share of the additional wage cost deducted from the manager's pay, and k is the penalty the firm imposes on the manager for each additional promotion. Let $b(z)$ denote the lowest bribe that will get the employee promoted,

$$b(z) = \alpha[k - R(z)] + \beta(w_h - w_l). \tag{14}$$

Since $R(z)$ is monotonic, this can be inverted to give the lowest performance $z(b)$ in the first period for which a bribe of b is sufficient to ensure promotion

$$z(b) = R^{-1}\left[k - \frac{b}{\alpha} + \frac{\beta}{\alpha}(w_h - w_l)\right], \quad \text{for } b \geqq 0. \tag{15}$$

There are two cases to consider. An employee may know the manager's assessment of performance z (*informed employees*) and thus exactly how much to bribe the manager to obtain promotion, namely $b(z)$. Alternatively, an employee may not know the manager's assessment of performance (*uninformed employees*) and thus not know what bribe is required for promotion. This latter case may be particularly relevant when the manager's assessment is of a highly subjective nature. We start with this case.

4.1. *Uninformed employees*

Let $P(b)$ denote the employees' subjective probability of promotion if bribe b is offered given by[5]

$$P(b) = \text{Prob}\,\{z \geqq z(b)\} = 1 - G[z(b)] = 1 - G\left\{R^{-1}\left[k - \frac{b}{\alpha} + \frac{\beta}{\alpha}(w_h - w_l)\right]\right\}. \tag{16}$$

The derivative of this is given by

$$P'(b) = g\left\{R^{-1}\left[k - \frac{b}{\alpha} + \frac{\beta}{\alpha}(w_h - w_l)\right]\right\}R^{-1\prime}\left[k - \frac{b}{\alpha} + \frac{\beta}{\alpha}(w_h - w_l)\right]\Big/\alpha. \tag{17}$$

The expected utility in period 2 from offering bribe b is

$$u(w_l) + P(b)[u(w_h - b) - u(w_l)], \tag{18}$$

and the first-order condition that must be satisfied for an optimal bribe b^* greater than zero is

$$P'(b^*)[u(w_h - b^*) - u(w_l)] - P(b^*)u'(w_h - b^*) = 0. \tag{19}$$

Our concern here is with two questions. First, is the optimal bribe less than the wage gain from promotion, so that the possibility of promotion actually provides an incentive for effort in period 1? Second, is it possible for the firm to ensure that employees do not try to influence the manager's decisions by offering bribes? On the first of these, note that $u'(w) > 0$ and that both $P(b)$ and $P'(b)$ are nonnegative. Thus, unless $[u(w_h - b^*) - u(w_l)] > 0$, the left-hand side of (19) will be negative for any bribe b^* that gives positive probability of promotion ($P(b^*) > 0$). Hence, it can never be optimal for employees to offer a bribe $b = w_h - w_l$ if that results in a positive probability of promotion. (Since employees actually pay the bribe only if promoted, it makes no difference what bribe they offer if there is no probability of promotion.) The intuition is as follows. A bribe for promotion equal to the wage gain would result in no utility gain from being successful. A slightly lower bribe would still leave some probability of promotion but in this case there would be a utility gain from being successful.

The firm can also ensure employees do not offer bribes. From (16) and (17) it is clear that, as α gets large, $P(b)$ converges to $1 - G[R^{-1}(k)]$ and $P'(b)$ to zero for finite β and w_h. Then it is certainly the case that $b^* = 0$ for α sufficiently large provided $R^{-1}(k) < \bar{z}$. The intuition here is that, as α gets larger, the cost to the manager of promoting those less productive in job h than in job l increases, so the effect of increasing the bribe on the probability of promotion gets smaller. But, with $R^{-1}(k) < \bar{z}$, there is some probability of promotion even without a bribe for α sufficiently large and the smaller the bribe the bigger the gain from promotion. Thus at some point the gain in the probability of promotion from offering a bribe is more than outweighed by the loss from the reduced gain from being promoted. If the firm sets $\beta = 0$, it follows from (15) that the manager will promote all employees with first period performance of $\hat{z}$ or better, where $\hat{z}$ is defined by $R(\hat{z}) = k$. We thus have the following result.

5. In this expression, we use the convention that, for y such that $R^{-1}(y) > \bar{z}$, then $G[R^{-1}(y)] = 1$ and for y such that $R^{-1}(y) < \underline{z}$, then $G[R^{-1}(y)] = 0$.

Proposition 1. *With $\alpha > 0$, any optimal bribe that results in a positive probability of promotion is less than $w_h - w_l$. For $k < R(\bar{z})$ and α sufficiently large, the unique optimal bribe is zero and, when $\beta = 0$ also, the manager promotes all employees with first period performance of $\hat{z}$ or better, where $\hat{z}$ is defined by $R(\hat{z}) = k$.*

Since with a risk neutral manager there is no efficiency loss from making α large and setting $\beta = 0$, the firm can always costlessly induce the manager to promote all those with performance of $\hat{z}$ or better for any $\hat{z}$ it wants. Just *how* large α needs to be depends on the probability distributions and the wages the firm sets but, in practice, α may not need to be that large—all that is necessary is for it to be large enough that $P'(b)$ is sufficiently small that (19) cannot be satisfied for any positive b.

There are other forms of the manager's contract that are equally good from the firm's point of view. For example, the firm does not lose profits if employees pay some bribes, and it therefore has to increase the wage for those promoted to provide the same incentives, as long as it can reduce payments to the manager by the expected value of the wage increase. Thus, it may well be possible for the firm to obtain the critical performance level it wants for promotion at no additional cost with a lower value of α and employees engaging in some influence activities. Moreover, the same effect can be achieved with the manager's reward based on profits (rather than on revenues and costs separately) by setting $\beta = \alpha$ and choosing k appropriately or with the manager's reward independent of the number of employees promoted by setting $k = 0$ and choosing the ratio β/α appropriately. Since our concern here is with the effectiveness of a promotion system and not with the precise form of managerial compensation, we do not pursue this issue further.

4.2. *Informed employees*

We turn now to informed employees who know the manager's assessment of their performance. In that case, an employee with first period performance z offers the manager the minimum bribe acceptable for promotion, namely $b(z)$ defined in (14), as long as that is nonnegative and no more than $(w_h - w_l)$. For $b(z) < 0$, the employee does not offer a bribe. That is the case for $z > z^h$ where z^h is defined by

$$\alpha[k - R(z^h)] + \beta(w_h - w_l) = 0. \tag{20}$$

For $b(z) > w_h - w_l$, there is no bribe the employee would be prepared to offer for which the manager would give promotion. This will be the case for $z < z^l$ where z^l is defined by

$$\alpha[k - R(z^l)] + (\beta - 1)(w_h - w_l) = 0. \tag{21}$$

In this case, it makes no difference what bribe less than $w_h - w_l$ the employee offers since it will never be accepted. We then have the following result.

Proposition 2. *With informed employees*:

1. *employees with first period performance $z > z^h$ offer no bribe and are promoted*;
2. *employees with first period performance $z \in [z^l, z^h]$ offer a bribe of $b(z)$ defined in* (14);
3. *employees with first period performance $z < z^l$ are not promoted.*

When $\beta = 1$, then $R(z^l) = k$.

With informed employees, those on the margin of promotion are always prepared to engage in some form of activity to influence the manager's promotion decisions because z^l is always less than z^h. However, the resulting bribes impose no efficiency loss on the firm because their expected value can always be deducted from the lump sum component γ_t of the manager's pay. Moreover, by appropriate choice of contract, the firm can induce the manager to set any level of performance it likes as the critical level for promotion z^l. This is most obvious for the case given in the proposition—$\beta = 1$ implies $R(z^l) = k$, and thus $z^l = \hat{z}$ as defined in Proposition 1. Then the only difference between the two cases is that informed employees with performance between $\hat{z}$ and z^h give the manager bribes in order to obtain promotion. As with the case of uninformed employees, there are other forms of managerial contract that achieve the same effect for the firm but we do not pursue this issue here.

5. OPTIMAL PROMOTION SCHEMES

Propositions 1 and 2 establish the period 2 payoffs and assignments of employees for a given contract with the firm. In this section, we discuss the firm's optimal choice of contract.

We know from those propositions that there is no efficiency loss from the firm setting α, β and k such that employees with performance $z \geqq \hat{z}$ are promoted, for any $\hat{z}$ the firm chooses. Because $z = x_1 - a$, the first period performance level critical for promotion is $\hat{x} = \hat{z} + a$. The firm's choice of $\hat{z}$ and wage levels w_h and w_l determines the first period effort chosen by employees. As standard in principal–agent models, we can think of the firm itself choosing this effort as long as the promotion scheme is constrained to make the employees' effort choice incentive compatible. In what follows, we deal in detail only with the case of uninformed employees and α sufficiently large that no bribes are paid. Then those promoted receive a net income of w_h. If α is low enough for bribes b to be paid, then w_h below must be interpreted as the net income $w_h - b$. In the case of informed employees, there are always some employees who pay bribes which complicates the details of the analysis but does not affect the essential points.

With no bribes, an uninformed employee's expected utility at the start of employment is

$$U_0 = u(w_1) - v(a) + \rho\{[1 - G(\hat{x} - a)]u(w_h) + G(\hat{x} - a)u(w_l)\}, \tag{22}$$

where ρ is the discount factor employees apply to period 2. The first-order condition for the employee's optimal choice of first period effort, a, is

$$-v'(a) + \rho g(\hat{x} - a)[u(w_h) - u(w_l)] = 0. \tag{23}$$

For what follows, we assume for simplicity that the functions $v(\cdot)$ and $g(\cdot)$ are such that U_0 is strictly concave in a so that there is a unique a that satisfies (23) for any given contract. A sufficient condition for this is that $G(\cdot)$ is convex, the Convexity of the Distribution Function Condition (CDFC), but this is considerably stronger than we require. The first-order condition (23) determines the employee's first period effort for any given contract and thus acts as a constraint on what the firm can achieve. In recognizing this constraint, the firm can substitute $\hat{z} = \hat{x} - a$ into (23) to replace that condition by

$$-v'(a) + \rho g(\hat{z})[u(w_h) - u(w_l)] = 0. \tag{24}$$

The mean of z for an employee of type i in job l is s_l^i, $i \in \{H, L\}$. At the start of employment the firm's expected profits are thus

$$\Pi_0 = a + qs_l^H + (1-q)s_l^L - w_1 + \delta\left\{\int_{\underline{z}}^{\hat{z}} l(z)g(z)dz + \int_{\hat{z}}^{\bar{z}} h(z)g(z)dz - G(\hat{z})w_l - [1-G(\hat{z})]w_h\right\}, \tag{25}$$

where δ is the discount factor the firm applies to period 2, and $h(z)$ and $l(z)$, defined in (7) and (8), are the expected performance levels in the two jobs of an employee with first period performance z.

An optimal contract is a choice of $(a, w_1, w_l, w_h, \hat{z})$ that maximizes Π_0 subject to the employees' first-order condition, the participation constraint $U_0 \geqq \bar{U}$, where $\bar{U}$ is the two period expected utility employees can obtain on the market, and the constraint $w_l \geqq \underline{w}$, which ensures that employees cannot be compelled to work for the firm in the second period at a wage lower than they could receive elsewhere. An optimal contract is thus the solution to

$$\max_{a, w_1, w_l, w_h, \hat{z}} \Pi_0 \text{ subject to } U_0 \geqq \bar{U},\ w_l \geqq \underline{w} \text{ and (24).} \tag{26}$$

The optimal choice of $\hat{z}$ is always internal to the interval $[\underline{z}, \bar{z}]$. Otherwise, either all employees are promoted or none are promoted and, in either case, there is no incentive for effort in period 1. But then the firm would certainly increase profits by assigning employees efficiently to jobs, that is, setting $\hat{z} = \tilde{z}$ which is internal to the interval $[\underline{z}, \bar{z}]$. Thus the first-order conditions for a maximum $(a^*, w_1^*, w_l^*, w_h^*, \hat{z}^*)$ with multipliers μ, η and λ attached to the constraints are as follows

$$a:\quad 1 - \mu v'(a^*) - \lambda v''(a^*) = 0, \tag{27}$$

$$w_1:\quad -1 + \mu u'(w_1^*) = 0, \tag{28}$$

$$w_h:\quad -\delta[1-G(\hat{z}^*)] + \rho u'(w_h^*)\{\mu[1-G(\hat{z}^*)] + \lambda g(\hat{z}^*)\} = 0, \tag{29}$$

$$w_l:\quad \left.\begin{aligned} -\delta G(\hat{z}^*) + \rho u'(w_l^*)\{\mu G(\hat{z}^*) - \lambda g(\hat{z}^*)\} \leqq 0, \\ w_l^* \geqq \underline{w}, \end{aligned}\right\} \tag{30}$$

$$\hat{z}:\quad \begin{aligned} &\delta[l(\hat{z}^*) - h(\hat{z}^*)]g(\hat{z}^*) + \delta g(\hat{z}^*)(w_h^* - w_l^*) \\ &\quad - \rho[u(w_h^*) - u(w_l^*)][\mu g(\hat{z}^*) - \lambda g'(\hat{z}^*)] = 0, \end{aligned} \tag{31}$$

plus the relevant conditions on the multipliers and the constraints. The brace in (30) indicates a pair of complementary inequalities.

The first result concerns risk neutral employees (and also, to avoid infinite loans, $\rho = \delta$).

Proposition 3. *If employees are risk neutral and $\rho = \delta$, a profit-maximizing contract with incentives provided by promotion ensures efficient effort $a^* = \tilde{a}$ in the first period and efficient assignment $(\hat{z}^* = \tilde{z})$ in the second period.*

Proof. Given risk neutrality, we can normalize $u(w) = w$. Then equation (28) implies $\mu = 1$ and equations (29) and (30) imply $\lambda = 0$. With these, condition (27) implies $a^* = \tilde{a}$ and (31) implies $\hat{z}^* = \tilde{z}$, as defined by (10). ||

The result shows that with risk neutral employees the profit-maximizing promotion contract yields both efficient effort for the first period ($a=\tilde{a}$) and efficient assignment for the second. Thus, in this case, there is no Peter Principle effect. The essential intuition is that, for any given expected second period wage, there is no efficiency cost when employees are risk neutral to choosing the gap between w_l and w_h in order to induce efficient effort. This is true for *any* given $\hat{z}$ and hence $g'(\hat{z})$, so the firm can choose the efficient level of $\hat{z}$ without any adverse consequences for efficient effort. Moreover, risk neutrality ensures that, for any w_h, w_l and $g'(\hat{z})$, there is no efficiency loss from adjusting the first period wage w_1 to satisfy the employee's participation constraint with equality. Of course, second period effort is always zero when incentives are provided only by promotion because there is no possibility of future promotion to motivate employees. But in this respect, providing incentives by promotion is no worse than using the monetary bonuses discussed in Section 3.

The model is consistent with the cohort effects described in Baker *et al*. (1994*a*, *b*) because there is no need for w_l and w_h to be related to market conditions in period 2 to achieve efficiency. They can be set at the time employees are hired. Indeed, if employees are risk averse, that is precisely what it is efficient to do. Wage increases over time within cohorts then occur because of internal commitments, not because wages are pushed up by the outside market.

Under the assumption used here that other potential employers do not observe anything that happens within the firm, the wage they are prepared to offer employees in period 2 is independent of the contract of their first period employer. Allowing that wage to depend on that contract creates no problem for Proposition 3 because there is no efficiency loss from raising both w_l and w_h by the same amount to ensure that w_l matches what is available elsewhere and reducing w_1 so that the participation constraint $U_0 \geqq \bar{U}$ continues to hold with equality. Even if, as in Waldman (1984), Gibbons (1986), Ricart i Costa (1988) and Bernhardt and Scoones (1993), other potential employers observe wages or job assignments for period 2 and so can tailor their wage offers to employees accordingly, Proposition 3 continues to hold. The reason is that the firm can commit to having both w_h and w_l higher than any other potential employer is prepared to pay and still reduce w_1 so that the participation constraint $U_0 \geqq \bar{U}$ holds with equality. Commitment to the promotion scheme can thus overcome the inefficiency that arises in those models because the firm has an incentive not to promote employees in order to hide information about their ability from competing employers. (Gibbons (1986) reaches a similar conclusion.)

Proposition 3 can be generalized to many ability and job types. With additional time periods, ongoing effort and assignment efficiency is possible as long as the sets of outputs of the two types do not become disjoint, at least in the case of uninformed employees.[6] New information merely causes the firm to reassign employees between jobs, a case examined in greater detail in Fairburn (1994). If they become disjoint, then type is known precisely, and with it the efficient assignment for the future. As in Gibbons (1986), there is then no point in employees exerting effort to influence the assessment of their type, and incentives can continue only if the firm commits to inefficient promotion standards.

6. With informed employees, the probability of further promotion and demotion for given effort will depend on past performance and it may not be possible to induce efficient effort from employees with different past performances unless promotion decisions are biased. Meyer (1991) analyses another reason for biasing promotions.

6. ASSIGNMENT DISTORTIONS AND THE PETER PRINCIPLE

We know from standard principal–agent theory that, when employees are risk averse, there is a trade off between incentive provision and the provision of insurance. In the model analysed here, there is a third consideration: the efficiency of period 2 assignments. To understand why risk aversion may result in distortion of period 2 assignments from the efficient level, consider the employees' first-order condition for effort (23). There is an obvious trade off between insurance and effort incentives because full insurance would require $w_h = w_l$ but then there would be no incentive for effort. Suppose the firm were to consider a small change in the criterion for promotion $\hat{z} = \hat{x} - a$ from the efficient level $\tilde{z}$. Because it is starting from the efficient level, the efficiency loss from the resulting inefficient assignment is of second order. If, however, the change increases $g(\hat{z})$, it enables the firm to maintain the same effort while reducing the gap between w_h and w_l. This improves the trade off between incentives and insurance in period 2. This improvement is achieved by a small change in $\hat{z}$ from $\tilde{z}$ in whichever direction increases $g(\hat{z})$. If $g'(\tilde{z}) < 0$, this trade off is improved by reducing $\hat{z}$ below the efficient level (promoting too many employees as in the Peter Principle).

This is not, however, a sufficient argument for inefficient assignment because a change in $\tilde{z}$ changes the probability of promotion and thus the expected utility of employees and profits of the firm. To be both feasible and worthwhile, the change in the probability of promotion must be accompanied by a change in at least one of the wages that ensures neither loses. Suppose the change is in w_h. To keep expected utility of employees constant, the changes in $\hat{z}$ and w_h must, when starting from $\tilde{z}$, satisfy

$$dU_0 \equiv \rho[1 - G(\tilde{z})]u'(w_h)dw_h - \rho g(\tilde{z})[u(w_h) - u(w_l)]d\hat{z} = 0, \tag{32}$$

so

$$dw_h = \frac{g(\tilde{z})}{[1 - G(\tilde{z})]} \frac{[u(w_h) - u(w_l)]}{u'(w_h)} d\hat{z}. \tag{33}$$

(The indirect effect on expected utility through a consequential change in a is of second order by the envelope theorem.) The effect of the changes in $\hat{z}$ and w_h on expected profits derived from (25) is

$$d\Pi_0 = da - \delta[1 - G(\tilde{z})]dw_h + \delta g(\tilde{z})(w_h - w_l)d\hat{z}, \tag{34}$$

or, with substitution for dw_h from (33),

$$d\Pi_0 = da + \delta g(\tilde{z})\left\{-\frac{[u(w_h) - u(w_l)]}{u'(w_h)} + w_h - w_l\right\}d\hat{z}. \tag{35}$$

It follows from $w_h > w_l$ and strict concavity of $u(w)$ that $u'(w_h)(w_h - w_l) < u(w_h) - u(w_l)$ and hence that the term in braces in (35) is negative. Thus, ignoring the effects on effort for the moment, reducing $\hat{z}$ below $\tilde{z}$ and reducing w_h to leave expected utility unchanged increases expected profits. From (24), the effect on effort is given by

$$-v''(a)da + \rho g(\tilde{z})u'(w_h)dw_h + \rho g'(\tilde{z})[u(w_h) - u(w_l)]d\hat{z} = 0, \tag{36}$$

or, substituting for dw_h from (33) and solving for da,

$$da = \rho\left\{\frac{g(\tilde{z})^2}{[1 - G(\tilde{z})]} + g'(\tilde{z})\right\}[u(w_h) - u(w_l)]d\hat{z}/v''(a). \tag{37}$$

Thus if $g'(\tilde{z})$ is sufficiently negative to make the term in braces in (37) non-positive, profits are certainly increased by reducing $\hat{z}$ below $\tilde{z}$ and thus by distorting assignments. What this requires is that the positive direct effect on effort of reducing $\hat{z}$ below $\tilde{z}$ when $g'(\tilde{z}) < 0$ is sufficiently strong to outweigh the negative indirect effect of the compensating reduction in w_h.

We can repeat this exercise holding constant w_h and instead changing w_l to maintain the level of expected utility as $\hat{z}$ is raised above $\tilde{z}$ to yield the following result.

Proposition 4. *With risk averse employees* ($u''(w) < 0$), *either of the following conditions is sufficient for* $\hat{z}^* \neq \tilde{z}$, *so period* 2 *assignments are inefficient*:

$$-\frac{g'(\tilde{z})}{g(\tilde{z})^2} \geqq \frac{1}{1-G(\tilde{z})}, \tag{38}$$

$$\frac{g'(\tilde{z})}{g(\tilde{z})^2} \geqq \frac{1}{G(\tilde{z})}. \tag{39}$$

The strength of this result is that it is independent of the precise form of the employees' utility function, provided only that it is strictly concave. This follows from the fact that $\tilde{z}$, defined by (10), depends only on the distributions of the performance measures and the proportions of high and low ability types in the population. Thus, for any given distributions of performance measures and proportions of types in the population, one can calculate $\tilde{z}$ and check whether (38) or (39) is satisfied without needing to make any assumption about the precise form of the employee's utility function. It is straightforward to construct numerical examples for which (38) holds and examples for which (39) holds.

The weakness of the proposition is that it does not tell us in which direction the distortion goes, in the Peter Principle direction of setting $\hat{z}^* < \tilde{z}$, and so promoting more employees than is efficient, or in the opposite direction. The reason is that the argument used to derive (38) and (39) is local, not global. Lemma 1 in the Appendix shows, however, that if (38) holds at an optimum the assignment distortion is in the direction of the Peter Principle, whereas if (39) holds the assignment distortion is in the opposite direction to the Peter Principle. But, if $f(z|s_l^i)$ is exponential with parameter $1/s_l^i$ for $i = L, H$, then (38) is satisfied for *every* z. We then have the following result which is proved in the Appendix.

Proposition 5. *Suppose the distribution of* z *is exponential with parameter* $1/s_j^i$, *that is*, $F(z|s_l^i) = 1 - e^{-z/s_l^i}$. *Then, with risk averse employees*, $\hat{z}^* < \tilde{z}$ *so period* 2 *assignments are distorted in the direction of the Peter Principle.*

Obviously, the exponential distribution does not satisfy CDFC so this result applies only when the disutility of effort function is sufficiently convex to ensure that expected utility is concave in effort.

The level of $\underline{w}$, the outside wage in period 2, can also have an impact on assignment distortions if it puts a binding constraint on how low w_l can be. If it is not binding and the discount rates of the firm and employees are the same, an optimal contract always has $w_l^* < w_1^* < w_h^*$ but a high level of $\underline{w}$ can result in the first of these inequalities being reversed. The following result explores the effect.

Proposition 6. *Suppose* $\rho \leqq \delta$ *and employees are risk averse. If* $w_1^* \leqq w_l^*$ *and* $g'(\hat{z}^*) \geqq 0$, *then* $\hat{z}^* > \tilde{z}$ *so period* 2 *assignments are distorted in the opposite direction to the Peter Principle.*

Proof. Note first that, with $w_l^* < w_h^*$, risk aversion implies $u'(w_l^*) > u'(w_h^*)$. It then follows from (29) and (30) that $\lambda > 0$. It follows from (31), (10) and $R'(z) > 0$ that $\hat{z}^* > \tilde{z}$ if and only if

$$\delta g(\hat{z}^*)(w_h^* - w_l^*) - \rho[u(w_h^*) - u(w_l^*)][\mu g(\hat{z}^*) - \lambda g'(\hat{z}^*)] > 0, \tag{40}$$

or, substituting $\mu = 1/u'(w_1^*)$ from (28) and rearranging,

$$\frac{1}{u'(w_1^*)} - \frac{\delta}{\rho} \frac{w_h^* - w_l^*}{u(w_h^*) - u(w_l^*)} < \lambda \frac{g'(\hat{z}^*)}{g(\hat{z}^*)}. \tag{41}$$

Risk aversion and $w_h^* > w_l^*$ imply that

$$\frac{1}{u'(w_1^*)} < \frac{w_h^* - w_l^*}{u(w_h^*) - u(w_l^*)}, \quad \text{if } w_1^* \leqq w_l^*. \tag{42}$$

Thus, with $\lambda > 0$, $\rho \leqq \delta$, and $w_1^* \leqq w_l^*$, (41) certainly holds if $g'(\hat{z}^*) \geqq 0$ and hence $\hat{z}^* > \tilde{z}$. ||

This proposition gives conditions under which fewer employees are promoted than would be efficient, so the assignment distortion is in the opposite direction to the Peter Principle. A high floor on the wage for those not promoted increases the expected wage in period 2 and, to keep overall expected utility unchanged, the firm reduces the first period wage. But that distorts inter-temporal income smoothing (which, with $\rho \leqq \delta$, would have the expected marginal utility of income no less in period 2 than in period 1) by increasing the marginal utility of income in period 1. Reducing the probability of promotion counteracts this by increasing the expected marginal utility of income in period 2 because it increases the probability of receiving the lower wage w_l. When $g'(\hat{z}^*) \geqq 0$, that effect is reinforced by the incentive effect on effort of changing the probability of promotion that we discussed at the beginning of this section. Having $w_1^* \leqq w_l^*$ corresponds to the firm setting an upward sloping wage profile even for those employees not promoted because of the outside offers available in period 2.

We conclude this section by sketching a scenario in which the firm will indeed set such a wage profile. The constraint $w_l \geqq \underline{w}$ is consistent with the alternative jobs available to employees in period 2 having output independent of ability (so there is no adverse selection resulting from period 2 employees having been screened), paying the same wage $\underline{w}$ to all, and (like second period jobs within the firm) involving no more than minimal effort ($a = 0$). Suppose that in the initial hiring market there are more potential employees than there are jobs in which ability affects output, and plenty of jobs in which it does not, so that the alternative jobs determine the market utility available to employees. The contracts that firms offer must, therefore, provide utility $\bar{U} = (1+\rho)u(\underline{w})$. From (22) the two period expected utility of an employee under an optimal contract can be written

$$U_0^* = u(w_1^*) - v(a^*) + \rho u(w_l^*) + \rho[u(w_h^*) - u(w_l^*)][1 - G(\hat{x}^* - a^*)]. \tag{43}$$

Moreover, that a^* is chosen optimally implies

$$-v(a^*) + \rho[u(w_h^*) - u(w_l^*)][1 - G(\hat{x}^* - a^*)] \geqq 0, \tag{44}$$

since the employee could certainly have chosen $a^* = 0$ and, since $v(0) = 0$, $a^* = 0$ would have ensured that (44) is satisfied. Since (44) contains all the terms in a^* in (43), optimal choice of a can do no worse than this. Hence it follows that

$$U_0^* \geqq u(w_1^*) + \rho u(w_l^*). \tag{45}$$

Finally note that (28) implies $\mu > 0$ so that the participation constraint $U_0 \geqq \bar{U}$ binds and

$$U_0^* = (1+\rho)u(\underline{w}). \tag{46}$$

Combining (45) and (46), we see that $u(w_1^*) + \rho u(w_l^*) \leqq u(\underline{w}) + \rho u(\underline{w})$. But since $w_l^* \geqq \underline{w}$, it follows that $w_1^* \leqq \underline{w}$ and hence $w_1^* \leqq w_l^*$.

7. CONCLUSION

In this paper we set out to address two questions. First, why do firms use promotions to motivate employees when there appear to be other incentive schemes that are better suited to the task? Second, what are the consequences of using promotions to provide incentives?

In regard to the first question, we have shown that there are plausible circumstances in which other incentive schemes have limited effect. Although firms can develop reputations for honest behaviour, they must in practice delegate important aspects of their personnel policy to managers who may not be likewise motivated. A firm's reputation can ensure that it is in its interest to pay the salaries and bonuses it has said it would pay, even in the absence of strict legal compulsion to do so. But this will not provide effective incentives if the firm is responding honestly to information passed to it by managers who are themselves succumbing to short run expediency. If, after a contract has been proposed and performance realized, managers stand to gain by responding to the interests of employees, then employees themselves will recognize that what is important is not so much to generate a good performance in the first place as to effectively influence the manager when the time comes for performance evaluations to be made to the centre. In these circumstances, performance related pay for employees loses its effectiveness. Furthermore, tournament schemes based on relative performance, which have often been assumed resistant to manipulation by employers, likewise lose their effectiveness if they rely on monetary bonuses or on promoting employees in rank while continuing to do the same job. However, promotion schemes that involve promoting high performance employees to different jobs can still be effective.

The important thing about promotion schemes of this sort is not that they resemble tournaments but that employees differ in their suitability for different jobs, and that good performance in one job is a reasonable indicator of suitability for some other job. Given these, a manager who receives performance related pay can be given an incentive to allocate employees to the job to which they are most suited, an incentive that offsets that to succumb to influence activities. These insights, we believe, extend beyond the particular model formulation we have adopted here. Of course, the promotion mechanism will work only if the same manager retains responsibility for promoted employees so that the overall performance on which the manager's compensation is based is improved by the manager making the correct promotion decisions. Thus to make it work, firms must ensure that managers bear the full output consequences of their promotion decisions. Only promotion schemes for which this is the case can be the result of the effects analysed here.

In practice, of course, there is always a chance that bribery or other influence activities may be detected and punished by the firm, so managers may not accept if the gains are small. This may be enough to ensure that managers allocate smaller bonuses or pay rises in the way the firm would like, so the firm can provide some effort incentives in these ways. That there are limits to this, however, is consistent with the evidence in Gibbs (1995) that the firm he studied did not adjust other incentives for those employees whose promotion prospects declined, with the result that their performance declined too, and with the big increases in (present discounted value of future) pay coming with promotion, as found by Baker, Gibbs and Holmström (1994*b*).

A promotion scheme can certainly be used to provide *some* incentives but there is the risk that providing incentives will be at the expense of distorting the assignment of employees to the jobs to which they are most suited. In the model used here, assignment distortions do not occur if employees are risk neutral. Moreover, effort levels of those working for promotion are efficient, though performance slackens off in the period before retirement when no further promotions are possible. However, when employees are risk averse, it is not just incentives and insurance that are traded off against each other, as in the standard principal–agent problem, but typically assignment efficiency too.

Our analysis indicates that the distortions to assignments underlying the Peter Principle may in fact have a sound basis in optimal incentive theory. When promotions are used to provide incentives and employees are risk averse, standards for promotion may be laxer than would be efficient for assignment reasons alone. The firm consciously has some employees promoted to jobs beyond their capability, in the sense that they would be more productive in lower ranking jobs. But it is also quite possible for the distortion to go in the opposite direction, with criteria for promotion being tougher than would be efficient. These assignment distortions occur even when ability is one-dimensional. If ability is multi-dimensional, employee performance in a lower level job may not be a good indicator of suitability for a higher level job. In that case, promoting those who perform best may result in inefficient assignments, and the Peter Principle result may apply, for an alternative reason. Exploration of that alternative awaits future research.

APPENDIX

A.1 *Optimal promotion schemes with risk averse employees*

Lemma 1. *With uninformed employees who are risk averse, a profit maximizing contract with incentives provided by promotion has*:

$$\text{(i)}\quad \hat{z}^* < \tilde{z} \qquad \textit{if } -\frac{g'(\hat{z}^*)}{g(\hat{z}^*)} \geqq \frac{g(\hat{z}^*)}{1-G(\hat{z}^*)}; \tag{47}$$

$$\text{(ii)}\quad \hat{z}^* > \tilde{z} \qquad \textit{if } \frac{g'(\hat{z}^*)}{g(\hat{z}^*)} \geqq \frac{g(\hat{z}^*)}{G(\hat{z}^*)}. \tag{48}$$

Proof. With $w_h^* > w_l^*$, which is necessary for incentives from the employees' first-order condition (24), risk aversion implies $u'(w_h^*) < u'(w_l^*)$. With conditions (29) and (30), this implies $\lambda > 0$.

Comparison of (31) and (10) reveals that $\hat{z}^* < \tilde{z}$ if and only if

$$(w_h^* - w_l^*) - [u(w_h^*) - u(w_l^*)]\frac{\rho}{\delta}\left[\mu - \lambda\frac{g'(\hat{z}^*)}{g(\hat{z}^*)}\right] < 0. \tag{49}$$

Note that, if the following condition is satisfied

$$\frac{\rho}{\delta}\left[\mu - \lambda\frac{g'(\hat{z}^*)}{g(\hat{z}^*)}\right] \geqq \frac{1}{u'(w_h^*)}, \tag{50}$$

then concavity of $u(\cdot)$ alone is sufficient to ensure that $\hat{z}^* < \tilde{z}$ because

$$(w_h^* - w_l^*) - [u(w_h^*) - u(w_l^*)]\frac{1}{u'(w_h^*)} = \frac{1}{u'(w_h^*)}[u(w_l^*) + u'(w_h^*)(w_h^* - w_l^*) - u(w_h^*)], \tag{51}$$

which is negative by concavity of $u(\cdot)$, given $w_h^* > w_l^*$. But also note that (29) can be written

$$\frac{\rho}{\delta}\left[\mu + \lambda \frac{g(\hat{z}^*)}{1 - G(\hat{z}^*)}\right] = \frac{1}{u'(w_h^*)}, \tag{52}$$

and so a sufficient condition for (50) is that

$$-\frac{g'(\hat{z}^*)}{g(\hat{z}^*)} \geqq \frac{g(\hat{z}^*)}{1 - G(\hat{z}^*)}. \tag{53}$$

Thus, if (53) is satisfied, then certainly $\hat{z}^* < \tilde{z}$. This proves part (i).

Conversely, if the following is satisfied

$$\frac{\rho}{\delta}\left[\mu - \lambda \frac{g'(\hat{z}^*)}{g(\hat{z}^*)}\right] \leqq \frac{1}{u'(w_l^*)}, \tag{54}$$

concavity of $u(\cdot)$ alone is sufficient to ensure that $\hat{z}^* > \tilde{z}$ by a similar type of argument. But note also that (30) can be written

$$\frac{\rho}{\delta}\left[\mu - \lambda \frac{g(\hat{z}^*)}{G(\hat{z}^*)}\right] \leqq \frac{1}{u'(w_l^*)}, \tag{55}$$

and so a sufficient condition for (54) is

$$\frac{g'(\hat{z}^*)}{g(\hat{z}^*)} \geqq \frac{g(\hat{z}^*)}{G(\hat{z}^*)}. \tag{56}$$

Thus if condition (56) is satisfied, certainly $\hat{z}^* > \tilde{z}$. This proves part (ii). ||

A.2. *Proof of Proposition* 5

For the exponential distribution

$$f(z|s_j^i) = \frac{1}{s_j^i} e^{-z/s_j^i}, \qquad s_j^i > 0, z \in [0, \infty). \tag{57}$$

This has mean s_j^i as assumed. Then

$$f'(z|s_j^i) = -\frac{1}{s_j^{i2}} e^{-z/s_j^i} \quad \text{and} \quad F(z|s_j^i) = 1 - e^{-z/s_j^i}. \tag{58}$$

Note that since

$$g(z) = q\,f(z|s_l^H) + (1-q)\,f(z|s_l^L), \tag{59}$$

we have, whatever the distribution $f(z|s_j^i)$,

$$g'(z) = q\,f'(z|s_l^H) + (1-q)\,f'(z|s_l^L), \tag{60}$$

$$G(z) = qF(z|s_l^H) + (1-q)F(z|s_l^L). \tag{61}$$

Thus, for the exponential case, we have

$$g(z) = \frac{q}{s_l^H} e^{-z/s_l^H} + \frac{1-q}{s_l^L} e^{-z/s_l^L}, \tag{62}$$

$$g'(z) = -\frac{q}{s_l^{H2}} e^{-z/s_l^H} - \frac{1-q}{s_l^{L2}} e^{-z/s_l^L} = \frac{1}{s_l^H} g(z) + \left(\frac{1}{s_l^H} - \frac{1}{s_l^L}\right)\frac{1-q}{s_l^L} e^{-z/s_l^L}, \tag{63}$$

$$G(z) = q(1 - e^{-z/s_l^H}) + (1-q)(1 - e^{-z/s_l^L}) = 1 - q\,e^{-z/s_l^H} - (1-q)e^{-z/s_l^L}, \tag{64}$$

$$1 - G(z) = qe^{-z/s_l^H} + (1-q)e^{-z/s_l^L}. \tag{65}$$

Now, provided $g(\hat{z}^*) > 0$ and $1 - G(\hat{z}^*) > 0$, (47) is equivalent to

$$-g(\hat{z}^*)[1 - G(\hat{z}^*)] \geqq g(\hat{z}^*)^2. \tag{66}$$

Substitution of the expressions for the exponential distribution into this (and writing z for $\hat{z}^*$ for notational convenience) gives the equivalent condition

$$\left[\frac{q}{s_l^{H2}}e^{-z/s_l^H}+\frac{1-q}{s_l^{L2}}e^{-z/s_l^L}\right][qe^{-z/s_l^H}+(1-q)e^{-z/s_l^L}]\geqq[q\,f(z|s_l^H)+(1-q)\,f(z|s_l^L)^2], \tag{67}$$

or

$$\left(\frac{q}{s_l^H}e^{-z/s_l^H}\right)^2+q(1-q)e^{-z/s_l^H}e^{-z/s_l^L}\left(\frac{1}{s_l^{H2}}+\frac{1}{s_l^{L2}}\right)+\left(\frac{1-q}{s_l^L}e^{-z/s_l^L}\right)^2$$
$$\geqq\left(\frac{q}{s_l^H}e^{-z/s_l^H}\right)^2+2q(1-q)e^{-z/s_l^H}e^{-z/s_l^L}\left(\frac{1}{s_l^H s_l^L}\right)+\left(\frac{1-q}{s_l^L}e^{-z/s_l^L}\right)^2, \tag{68}$$

or

$$\frac{1}{s_l^{H2}}+\frac{1}{s_l^{L2}}\geqq\frac{2}{s_l^H s_l^L}, \tag{69}$$

or, since s_l^H and $s_l^L>0$,

$$\frac{s_l^{H2}+s_l^{L2}}{(s_l^H s_l^L)^2}\geqq\frac{2s_l^H s_l^L}{(s_l^H s_l^L)^2}, \tag{70}$$

or

$$s_l^{H2}+s_l^{L2}-2s_l^H s_l^L\geqq 0, \tag{71}$$

or

$$(s_l^H-s_l^L)^2\geqq 0, \tag{72}$$

which is certainly true. Thus (47) is certainly satisfied in this case and, by Lemma 1, $\hat{z}^*<\tilde{z}$. ||

Acknowledgements. A very preliminary version of this paper was presented at the European Summer Symposium in Economic Theory at the Studienzentrum Gerzensee in Switzerland, July 1994, and subsequent versions were presented at the Bellaterra seminar in Barcelona and at Harvard and Toulouse. We are grateful to participants in these seminars, in particular to George Baker, Paul Milgrom, Elu von Thadden and Jean Tirole, for helpful comments. The paper was written while the first author was visiting the Institut d'Anàlisi Econòmica. The hospitality of the Institut and financial support from the EU Human Capital and Mobility programme (contract no. ERBCHBGCT940577) are gratefully acknowledged.

REFERENCES

BAKER, G., GIBBS, M. and HOLMSTRÖM, B. (1994*a*), "The Internal Economics of the Firm: Evidence from Personnel Data", *Quarterly Journal of Economics*, 109, 881–919.

BAKER, G., GIBBS, M. and HOLMSTRÖM, B. (1994*b*), "The Wage Policy of a Firm", *Quarterly Journal of Economics*, 109, 921–955.

BAKER, G. P., JENSEN, M. C. and MURPHY, K. J. (1988), "Compensation and Incentives: Practice vs. Theory", *Journal of Finance*, 43, 593–616.

BERNHARDT, D. (1995), "Strategic Promotion and Compensation", *Review of Economic Studies*, 62, 315–339.

BERNHARDT, D. and SCOONES, D. (1993), "Promotion, Turnover, and Preemptive Wage Offers", *American Economic Review*, 83, 771–791.

BHATTACHARYA, S. (1983), "Tournaments and Incentives: Heterogeneity and Essentiality" (Graduate School of Business Working Paper No. 695, Stanford University).

CARMICHAEL, L. (1983), "Firm-Specific Human Capital and Promotion Ladders", *Bell Journal of Economics*, 14, 251–258.

CLARK, D. J. and RIIS, C. (1998), "Competition over More Than One Prize", *American Economic Review*, 88, 276–289.

FAIRBURN, J. A. (1994), "Career Paths and Incentives in Organizations", in *Promotions, Incentives and the Market for Corporate Control* (PhD thesis, University of Southampton).

FAIRBURN, J. A. and MALCOMSON, J. M. (1994), "Rewarding Performance by Promotion to a *Different* Job", *European Economic Review*, 38, 683–690.

GIBBONS, R. (1986), "Optimal Incentive Schemes in the Presence of Career Concerns" (Mimeo, MIT).

GIBBONS, R. and MURPHY, K. J. (1992), "Optimal Incentive Schemes in the Presence of Career Concerns: Theory and Evidence", *Journal of Political Economy*, 100, 468–505.

GIBBONS, R. and WALDMAN, M. (1999), ''Careers in Organizations: Theory and Evidence'', in O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*, Volume 3B (Amsterdam: Elsevier Science).

GIBBS, M. (1995), ''Incentive Compensation in a Corporate Hierarchy'', *Journal of Accounting and Economics*, 19, 247–277.

GREEN, J. R. and STOKEY, N. L. (1983), ''A Comparison of Tournaments and Contracts'', *Journal of Political Economy*, 91, 349–364.

HART, O. and HOLMSTRÖM, B. (1987), ''The Theory of Contracts'', in T. F. Bewley (ed.), *Advances in Economic Theory: Fifth World Congress* (Cambridge: Cambridge University Press).

HOLMSTRÖM, B. (1982), ''Moral Hazard in Teams'', *Bell Journal of Economics*, 13, 324–340.

HOLMSTRÖM, B. (1999), ''Managerial Incentive Problems: A Dynamic Perspective'', *Review of Economic Studies*, 66, 169–182.

HOLMSTRÖM, B. and MILGROM, P. (1991), ''Multitask Principal–Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design'', *Journal of Law, Economics, and Organization*, 7, 24–52.

KAHN, C. and HUBERMAN, G. (1988), ''Two-Sided Uncertainty and 'Up-or-Out' Contracts'', *Journal of Labor Economics*, 6, 423–444.

KREPS, D. M. (1990), ''Corporate Culture and Economic Theory'', in J. E. Alt and K. A. Shepsle (eds.), *Perspectives on Positive Political Economy* (Cambridge: Cambridge University Press).

LAFFONT, J.-J. and MARTIMORT, D. (1997), ''The Firm as a Multicontract Organization'', *Journal of Economics and Management Strategy*, 6, 201–234.

LAZEAR, E. P. (1992), ''The Job as a Concept'', in W. J. Bruns, Jr. (ed.), *Performance Measurement, Evaluation and Incentives* (Boston: Harvard Business School Press).

LAZEAR, E. P. and ROSEN, S. (1981), ''Rank–Order Tournaments as Optimal Labor Contracts'', *Journal of Political Economy*, 89, 841–864.

MacLEOD, W. B. and MALCOMSON, J. M. (1988), ''Reputation and Hierarchy in Dynamic Models of Employment'', *Journal of Political Economy*, 96, 832–854.

MALCOMSON, J. M. (1984), ''Work Incentives, Hierarchy, and Internal Labor Markets'', *Journal of Political Economy*, 92, 486–507.

MALCOMSON, J. M. (1986), ''Rank–Order Contracts for a Principal With Many Agents'', *Review of Economic Studies*, 53, 807–817.

MEYER, M. A. (1991), ''Learning from Coarse Information: Biased Contests and Career Profiles'', *Review of Economic Studies*, 58, 15–41.

MILGROM, P. R. (1988), ''Employment Contracts, Influence Activities, and Efficient Organization Design'', *Journal of Political Economy*, 96, 42–60.

MILGROM, P. and ROBERTS, J. (1988), ''An Economic Approach to Influence Activities in Organizations'', *American Journal of Sociology*, 94, Supplement, S154–S179.

MILGROM, P. and ROBERTS, J. (1992) *Economics, Organization and Management* (Englewood Cliffs, NJ: Prentice-Hall International).

MORI, P. A. (1998), ''Promotion Careers with Multiple Punitive Assignments'', *Economics Letters*, 60, 369–373.

NALEBUFF, B. J. and STIGLITZ, J. E. (1983), ''Prizes and Incentives: Towards a General Theory of Compensation and Competition'', *Bell Journal of Economics*, 14, 21–43.

PETER, L. J. and HULL, R. (1969) *The Peter Principle: Why Things Always Go Wrong* (New York: Morrow).

PRENDERGAST, C. (1993), ''The Role of Promotion in Inducing Specific Human Capital Acquisition'', *Quarterly Journal of Economics*, 108, 523–534.

PRENDERGAST, C. (1999), ''The Provision of Incentives in Firms'', *Journal of Economic Literature*, 37, 7–63.

PRENDERGAST, C. and TOPEL, R. (1996), ''Favoritism in Organizations'', *Journal of Political Economy*, 104, 958–978.

RADNER, R. (1981), ''Monitoring Cooperative Agreements in a Repeated Principal–Agent Relationship'', *Econometrica*, 49, 1127–1148.

RADNER, R. (1985), ''Repeated Principal–Agent Games with Discounting'', *Econometrica*, 53, 1173–1198.

RICART I COSTA, J. E. (1988), ''Managerial Task Assignment and Promotions'', *Econometrica*, 56, 449–466.

ROSEN, S. (1986), ''Prizes and Incentives in Elimination Tournaments'', *American Economic Review*, 76, 701–751.

ROSENBAUM, J. E. (1984) *Career Mobility in a Corporate Hierarchy* (San Diego: Academic Press).

TIROLE, J. (1986), ''Hierarchies and Bureaucracies: On the Role of Collusion in Organizations'', *Journal of Law, Economics and Organization*, 2, 181–214.

TIROLE, J. (1992), ''Collusion and the Theory of Organizations'', in J.-J. Laffont (ed.), *Advances in Economic Theory: Sixth World Congress*, Volume 2 (Cambridge: Cambridge University Press).

WALDMAN, M. (1984), ''Job Assignments, Signalling, and Efficiency'', *Rand Journal of Economics*, 15, 255–267.

WALDMAN, M. (1990), ''Up-or-Out Contracts: A Signaling Perspective'', *Journal of Labor Economics*, 8, 230–250.