TI 2026-028/VII
Tinbergen Institute Discussion Paper

A Bright Side of Peter-Principle Promotions

Robert Dur[1]
Kimiyuki Morita[2]
Takeharu Sogo[3]

1 Erasmus University Rotterdam and Tinbergen Institute

2 Senshu University

3 SKEMA Business School

A Bright Side of Peter-Principle Promotions[*]

Robert Dur[†], Kimiyuki Morita[‡], and Takeharu Sogo[§]

June 1, 2026

Abstract

Promotion to managerial roles is often strongly based on workers' performance in their current roles rather than on their managerial aptitude—the Peter principle. This paper provides a novel rationale for this apparent mismatch when promotions serve no incentive role. We model workers who care about meeting managers' expectations, and managers with interpersonal projection bias: when forming expectations about workers' performance, they give excessive weight to their own past performance. We show that organizations benefit from promoting high-performing workers to managers even when they have low managerial aptitude, because such managers hold high expectations, inducing greater effort.

"*At Paypal we never promoted anybody based upon their management skills. We promoted everybody based on their crowd. So if you wanted to be running the design team, you had to be the best designer. If you wanted to run the engineering team, you had to be the best engineer.* [...] *If you promoted that person, you're not going to demoralize your team. You have enthusiasm, energy from the rank and file* [...] *it is the antithesis of hiring a general manager, someone who's expertise is really at managing.*"

Keith Rabois, former executive at Paypal, LinkedIn, and Square in an interview in 2025 on YouTube: link (2:21-4:57)

1 Introduction

Excellent workers are not always excellent managers (Weidmann et al., 2026). Yet organizations often promote high-performing workers over those with the highest managerial qualities. This is known as the 'Peter principle' (Peter and Hull, 1969), which describes "organizations' tendency to promote workers who excel at their current jobs while downplaying or ignoring their aptitude for management" (Benson et al., 2019, p.2091).

This paper offers a novel reason for why this apparent mismatch may benefit organizations. We argue that a manager who performed very well as a worker holds higher expectations of workers' performance than a manager who struggled as a worker. The reason is interpersonal projection bias: people's tendency to assume that others share their own states or preferences. The manager's higher expectations in turn foster workers' productivity if workers care about meeting their manager's expectations (e.g., to avoid feelings of guilt or shame, or to feel pride or virtue). Thus, appointing former high-performing workers may boost

[*]We are grateful to Dan Bernhardt, Rosaria Distefano, Martin Dufwenberg, seminar participants at Erasmus University Rotterdam, and participants of the Contract Theory Workshop in Osaka for useful comments and suggestions.

[†]Erasmus University Rotterdam, Tinbergen Institute, CESifo, and IZA. E-mail: dur@ese.eur.nl

[‡]Senshu University. E-mail: kimiyuki1229@gmail.com

[§]SKEMA Business School—Université Côte d'Azur. E-mail: takeharu.sogo@skema.edu

organizational productivity and profitability. Indeed, we show that when choosing managers, organizations optimally trade off a person's managerial quality and her past performance as a worker.

Since the reason for promoting former high-performing workers in our model is to boost worker effort, one might conjecture that introducing incentive pay for workers *reduces* the benefit of Peter-principle promotions—consistent with the empirical finding of Benson et al. (2019) that Peter-principle promotions are less prevalent when workers' incentive pay is stronger. However, we identify a new channel through which incentive pay and Peter-principle promotions can instead complement each other. The reason is that high-ability workers respond more strongly to performance pay. Hence, appointing a manager who has higher expectations about workers' ability can become more attractive for the organization: such a manager expects even higher effort from workers in the presence of pay-for-performance, strengthening the motivational effect. Thus, rather than substituting, incentive pay and Peter-principle promotions can reinforce each other.

Our theory rests on two empirically supported assumptions. First, people exhibit interpersonal projection bias (also called social projection or the false-consensus effect). A rich literature in psychology and a growing number of studies in economics find strong support for this bias in various contexts, e.g. Ross et al. (1977), Marks and Miller (1987), Krueger and Clement (1994), Van Boven and Loewenstein (2003), Van Boven et al. (2003), Epley et al. (2004), Engelmann and Strobel (2012), Buchanan (2020), and Gagnon-Bartsch and Rosato (2024). Most closely related to our context, Bushong and Gagnon-Bartsch (2024) find evidence of interpersonal projection bias in a real-effort experiment in the laboratory and Cordes and Müller (2026) in field survey data on managers.

Second, people care about meeting others' expectations. Laboratory experiments in economics have provided extensive evidence that people tend to feel guilt when not living up to others' expectations and—anticipating this—adjust their behavior accordingly (for overviews, see Cartwright (2019) and Battigalli and Dufwenberg (2022)). Our model follows Khalmetski et al. (2015), who extended Battigalli and Dufwenberg (2007)'s classic model of guilt aversion to include positive feelings when exceeding others' expectations, finding strong support in their experimental data. Outside of the laboratory, Eden (1992) finds field-experimental evidence and Kierein and Gold (2000) provide meta-analytic evidence consistent with the idea that "leaders who expect more get more" in work organization settings (an effect often called 'Pygmalion effect').

The existing literature rationalizes the Peter principle by pointing to promotions' dual role: assigning the right people to management roles and incentivizing workers (Baker et al., 1988; Milgrom and Roberts, 1992). Whereas the assignment role calls for promoting workers with the highest managerial aptitude, organizations may instead promote the best-performing workers to incentivize workers' effort or risk-taking. Our rationale differs and applies even when organizations have sufficient alternative means to incentivize workers, so promotions are used solely to select the best person for a management job.[1] Yet, Peter-principle promotions emerge as optimal under two conditions: workers care about meeting their manager's expectations and managers exhibit interpersonal projection bias.

Another well-known rationalization of the Peter principle is described in Lazear (2004) and applies to performance of promoted workers more generally, not just those in a management role. Lazear (2004) argues that if the chance of promotion increases with stochastic performance, regression to the mean implies that promoted workers' post-promotion performance is on average lower than their pre-promotion performance. In our model, there is no uncertainty about abilities and characteristics at the time of promotion, so this

[1] In Fairburn and Malcomson (2001), the organization has multiple means to incentivize workers, but still may find it optimal to use promotions for incentives, because the assignment of promotions may be less susceptible to influence activities of workers than the assignment of monetary rewards.

argument does not apply.[2]

Our theory can help explain mixed empirical findings on whether high-performing workers become effective managers. In an influential paper, Benson et al. (2019) find that managers who were top salespeople tend to perform worse than managers who were less successful salespeople. Weidmann et al. (2026) find the same in a laboratory setting. A systematic review by Schleu and Hüffmeier (2021) of the empirical literature in economics, management, and psychology paints a rich picture, with some studies finding a negative relationship, others a positive relationship, and some no relationship between performance as an employee and as a manager. This holds even when restricting attention to high-powered studies. They conclude that the "empirical results are inconsistent [...] without an apparent systematic" (p.2). This ambiguous picture calls for theories—such as ours—that predict when to expect a negative versus positive relation.[3]

We extend our basic model in three ways. First, we allow for external hiring of a professional manager with no hands-on experience as a worker. Such a manager may excel in managerial quality, but her expectations of workers' performance are arguably based on what can be rationally expected, which may fall short of internal candidates' biased expectations. This reduces worker motivation, because workers work harder to meet their manager's expectations. Hence, a similar trade-off arises as in the basic model: the professional manager is appointed over the best internal candidate only if her managerial quality is sufficiently higher than that of the best internal candidate—that is, her advantage in managerial quality must outweigh the motivational advantage of the best internal candidate.

Second, as discussed above, we show that performance pay and Peter-principle promotions can be complements rather than substitutes.

Third, we study the optimality of Peter-principle promotions when workers can quit the organization after learning their manager's characteristics. Peter-principle promotions may increase turnover: low-ability workers may quit when assigned to a manager who performed very well in her role of worker, because the hard work that is needed to avoid disappointing such a manager is particularly costly for low-ability workers. Whether this increased turnover benefits or harms the organization depends on parameters. If wages are high, low-ability workers contribute negatively to profits, making increased turnover—and hence Peter-principle promotions—more profitable. If wages are low, increased turnover may reduce profits, deterring the organization from promoting high-ability workers.

2 Model

We consider an organization that lasts for two periods. In each period, the organization hires a group of new workers of exogenous size. Workers differ in three ways: productive ability θ; sensitivity to pride, shame, and guilt η; and management quality m. Their productive ability θ and management quality m are independently distributed according to the distributions $F(\theta)$ and $G(m)$ over $(\underline{\theta}, \overline{\theta})$ and $(\underline{m}, \overline{m})$, respectively, with $\underline{\theta} > 0$

[2] Dickinson and Villeval (2012) provide lab-experimental evidence for Lazear (2004)'s model. Two recent economic theory papers on the Peter principle are Fang and Noe (2022) and Fahn and Klein (2025). Fang and Noe (2022) study a model where workers choose between strategies that differ in risk and argue that 'overpromotion' may help in getting on average a better selection of people for the job. Fahn and Klein (2025) argue that reassignment of workers to new tasks can be profitable even if it results in a mismatch between employees' talents and tasks, because it helps organizations to exploit workers' overconfidence. Effects of promotions and demotions on worker confidence are also studied in Ishida (2006). Kamphorst and Swank (2013) study the effect of manager's perception of worker's ability on worker's behavior and performance in a model where workers have imperfect information about their own ability (see also Bénabou and Tirole (2003) and Swank and Visser (2007)).

[3] Schleu and Hüffmeier (2021) describe two other complementary theories: Avery et al. (2003) emphasize the importance of experience as a worker for management effectiveness, while Goodall and Bäker (2014) stress the importance of expertise obtained in the worker role.

and $\underline{m} > 0$. Sensitivity is independently and discretely distributed: a worker is either sensitive ($\eta > 0$) or not ($\eta = 0$). The proportion of sensitive workers in the population is $\rho \in (0, 1)$.

Upon hiring, neither the worker nor the organization knows the worker's ability, sensitivity, and management quality. New workers are randomly assigned to teams of given size, each headed by a manager. The characteristics of period-1 managers are exogenously given. They retire at the end of period 1 and are replaced by new managers recruited from the pool of period-1 workers. Our focus is on the organization's decision of which workers to promote to management at the end of period 1. Promoted workers lead period-2 teams and receive a wage that exactly satisfies their participation constraint in period 2. We assume that the period-2 reservation utility of all period-1 workers equals $\overline{U}$, independent of their characteristics. Non-promoted workers quit (or move to another position in the organization where their characteristics are payoff-irrelevant for the organization)[4] and also enjoy $\overline{U}$. Thus, all period-1 workers obtain $\overline{U}$ in period 2, regardless of promotion. Consequently, promotions play no incentive role.

Let us briefly elaborate on why we make this assumption. In many organizational settings, promotions come with higher compensation or status, which motivates workers to work harder in period 1. This incentive role of promotions, together with their matching role (assigning the right workers to management positions), constitutes a dual role central to the traditional explanation for Peter-principle promotions (Baker et al., 1988; Milgrom and Roberts, 1992). By shutting down the incentive role—i.e., assuming all workers receive the same period-2 utility regardless of promotion—we deliberately focus on the matching role only. This allows us to examine whether Peter-principle promotions can arise in such a setting, without the confounding effects of incentive-based promotion motives emphasized in the prior literature.[5]

Each worker chooses effort e at cost $(1/\theta)e$ after learning his type and his manager's identity. We assume that the wage of workers, w, is so high that no worker quits (we relax this assumption in section 4.3). Effort translates into performance following the production function $q = e$.

An insensitive worker's utility function reads:

$$U = w - \frac{1}{\theta}e. \tag{1}$$

Insensitive workers have no incentive to exert effort—there are no current or future gains from doing so—so they choose $e = 0$.

A sensitive worker's utility function is:

$$U = w - \frac{1}{\theta}e + \eta \ln\left(1 + q - \overline{q}\right) \tag{2}$$

where q is the worker's performance and $\overline{q}$ is his manager's expectation of worker's performance. In the spirit of Battigalli and Dufwenberg (2007) and Khalmetski et al. (2015), sensitive workers care about how their performance compares to their manager's expectation, feeling shame/guilt when $q < \overline{q}$ and pride/virtue when $q > \overline{q}$. Consequently, they suffer a utility loss when falling short of their manager's expectation and enjoy exceeding it. The functional form—natural logarithm with the "1" within the brackets—ensures that utility

[4] This implies that the organization's opportunity cost of promotion is identical across workers. In contrast, in Légeret, Zehnder and Tur (2025)'s model, an organization may hesitate to promote a high-performing worker because the person would no longer be active as worker.

[5] In line with our assumption, organizations in practice may not always want to use promotions to managerial roles as an incentive device, particularly when they have other means to incentivize workers. These include pay-for-performance (which we study in 4.2) or promotions to non-managerial positions (for instance, from a 'junior' to a 'senior' position, with an accompanying rise in pay and possibly status, but no managerial responsibilities). In the latter case, promotions to managerial and non-managerial roles coexist: the former serves a matching purpose while the latter handles incentives.

is concave and that workers neither gain nor lose from these feelings when exactly meeting the expectation. We adopt:

Assumption 1. $\eta\underline{\theta} \geq 1$.

This guarantees weakly positive effort for all sensitive workers, including the least able ($\underline{\theta}$).

Managers are interpersonally projection-biased regarding their workers' ability θ: they believe workers' abilities are more similar to their own than they actually are. The degree of projection bias is captured by $\alpha \in [0, 1]$. When $\alpha = 0$, the manager is unbiased and thus holds rational expectations about workers' ability and output. When $\alpha = 1$, the manager is fully biased, believing workers' ability equals her own θ. We describe manager's beliefs when $\alpha \in (0, 1)$ in section 3.2.

Worker i's effort e_i yields worker performance q_i, which the manager transforms into final team output:

$$Q = (1 - \lambda)\sum_i (q_i) + \lambda m$$

where λ represents the relative importance of management quality. For simplicity we assume that workers' performance and manager's management quality are additive.

Observing workers' performance q_i (and thus θ_i) and managerial quality m_i, the organization selects managers from the pool of workers at the end of period 1.[6] Its objective is to maximize expected period-2 output, which is equivalent to profit maximization because all non-promoted workers quit (or move to another role where their characteristics are irrelevant) and all promoted managers earn the same wage, regardless of their characteristics.

3 Analysis

3.1 Effort choice

As discussed, the insensitive worker chooses $e = 0$. For the sensitive worker, given the manager's expectation $\overline{q}$, the first-order condition (FOC) for optimal effort is

$$\frac{\eta}{1 + e - \overline{q}} = \frac{1}{\theta}.$$

Solving yields $e = \overline{q} + \eta\theta - 1$. Assumption 1 guarantees an interior solution for any $\overline{q} > 0$. Effort rises in the manager's expectation $\overline{q}$, the worker's ability θ, and his sensitivity η.

3.2 Manager's expectation

We first consider an unbiased manager. Using the solution for sensitive workers' effort, and noting that proportion $1 - \rho$ of workers is insensitive and exerts no effort, the unbiased manager's expectation of the worker's output satisfies:

$$\overline{q}_{UB} = \mathbb{E}[q_i(\theta_i)] = \rho\{\overline{q}_{UB} + \eta\mathbb{E}[\theta] - 1\}.$$

[6] Managerial qualities such as communication, coordination, and leadership skills are typically observable through ordinary workplace interactions before formal promotion, or can be professionally assessed in an assessment center. If managerial qualities are not perfectly known, m_i should be interpreted as the organization's best estimate based on available information. Replacing m_i with its conditional expectation leaves the analysis unchanged.

Solving yields:

$$\bar{q}_{UB} = \frac{\rho}{1-\rho}\{\eta\mathbb{E}[\theta] - 1\} > 0, \tag{3}$$

which increases with the expected ability $\mathbb{E}[\theta]$, the sensitivity η, and the fraction of sensitive workers ρ. Note that we need $\rho < 1$ for an interior solution, i.e., some workers must be insensitive.[7]

We next consider a biased manager. We assume that managers exhibit *interpersonal projection bias* concerning workers' ability: the manager believes her workers' ability θ is, to some extent, similar to her own ability, denoted by θ_M. Following Gagnon-Bartsch et al. (2021), the manager with ability θ_M perceives the worker's ability as a random variable $\hat{\theta}(\theta_M) = (1-\alpha)\theta + \alpha\theta_M$, where $\alpha \in [0,1]$ represents the degree of bias.[8] We use "hat" notation for subjective beliefs, distinguishing them from unbiased beliefs. If $\alpha = 1$, the manager is fully biased. In this case, she believes that her worker's ability is identical to her own, θ_M, and therefore holds the expectation as follows:

$$\bar{q}_{FB}(\theta_M) = \frac{\rho}{1-\rho}\{\eta\theta_M - 1\},$$

which increases with θ_M. If $\alpha = 0$, we revert to the unbiased case with expectation $\bar{q}_{UB}$ described by (3).

We now consider the case of $\alpha \in (0,1)$. For given θ_M, let $\hat{F}(\cdot \mid \theta_M)$ be the distribution of $\hat{\theta}(\theta_M)$. For any $x \in \left[(1-\alpha)\underline{\theta} + \alpha\theta_M, (1-\alpha)\overline{\theta} + \alpha\theta_M\right]$, we have

$$\hat{\theta}(\theta_M) \leq x \iff (1-\alpha)\theta \leq x - \alpha\theta_M \iff \theta \leq \frac{x - \alpha\theta_M}{1-\alpha}.$$

Thus, $\hat{F}(x \mid \theta_M) = F\left(\frac{x-\alpha\theta_M}{1-\alpha}\right)$. We have

$$\mathbb{E}[\hat{\theta} \mid \theta_M] = (1-\alpha)\mathbb{E}[\theta] + \alpha\theta_M.$$

Thus, the perceived worker ability $\hat{\theta}$ increases with the manager's ability θ_M, and satisfies $\mathbb{E}[\hat{\theta} \mid \theta_M] = \mathbb{E}[\theta]$ when $\theta_M = \mathbb{E}[\theta]$. Therefore, a manager with $\theta_M > \mathbb{E}[\theta]$ overestimates workers' ability, while a manager with $\theta_M < \mathbb{E}[\theta]$ underestimates it, and the magnitude of these over- and under-estimations increases with α.

The expectation of the biased manager with θ_M is given by

$$\bar{q}_B(\theta_M) = \mathbb{E}[q_i(\hat{\theta}_i) \mid \theta_M] = \rho\{\bar{q}_B(\theta_M) + \eta\mathbb{E}[\hat{\theta} \mid \theta_M] - 1\}.$$

Solving yields

$$\bar{q}_B(\theta_M) = \frac{\rho}{1-\rho}\{\eta\mathbb{E}[\hat{\theta} \mid \theta_M] - 1\}. \tag{4}$$

Since $\mathbb{E}[\hat{\theta} \mid \theta_M]$ is increasing in θ_M, the biased manager's expectation $\bar{q}_B(\theta_M)$ is higher when her own ability as a worker, θ_M, is higher. Moreover, the worker's optimal effort is increasing in $\bar{q}_B(\theta_M)$, and hence increases with θ_M. We call it *the incentive effect of promoting a candidate with higher* θ_M.

Relative to an unbiased manager, a biased manager with $\theta_M > \mathbb{E}[\theta]$ holds higher expectations and induces higher effort, whereas a biased manager with $\theta_M < \mathbb{E}[\theta]$ holds lower expectations and induces lower effort. These differences become more pronounced as α increases.

[7] Alternatively, quadratic effort costs could ensure an interior solution, albeit with added complexity (see Online Appendix).

[8] Morita et al. (2025) follow a similar approach to study how interpersonal projection among agents affects organization's incentive to use joint and relative performance evaluation.

3.3 Promotion

At the end of period 1, the organization decides which workers to promote to a management position. The expected period-2 output of a team led by a manager with management quality m is

$$Q = (1-\lambda)\sum_i \mathbb{E}[q_i] + \lambda m.$$

First, consider the case of $\alpha = 0$. In this case, all potential managers hold exactly the same expectation $\overline{q}_{UB}$. Consequently, a worker's effort and productivity are independent of the manager's characteristics. Since team output increases with the manager's managerial quality, the organization optimally promotes the workers with the highest possible m, making their earlier performance as a worker inconsequential for promotion. Hence, when managers do not suffer from projection bias, our model predicts that the Peter principle does not apply.

We now consider the case of $\alpha > 0$. The expected period-2 output of a team led by a biased manager with characteristics (θ_M, m) is given by

$$\begin{aligned} Q_B(\theta_M, m) &= (1-\lambda)n\rho\mathbb{E}[\overline{q}_B(\theta_M) + \eta\theta - 1] + \lambda m \\ &= (1-\lambda)n\rho\left\{\frac{\rho}{1-\rho}\left(\eta\mathbb{E}[\hat{\theta} \mid \theta_M] - 1\right) + \eta\mathbb{E}[\theta] - 1\right\} + \lambda m \end{aligned} \tag{5}$$

where n denotes the exogenous number of workers in the team.

The organization promotes the worker with (θ_M, m) who generates the highest period-2 output $Q_B(\theta_M, m)$, where $Q_B(\theta_M, m)$ is increasing in both arguments.[9] The marginal rate of technical substitution (MRTS) of ability as a worker (θ_M) for managerial ability (m) is given by

$$\text{MRTS} \equiv \frac{dm}{d\theta_M} = -\frac{\dfrac{\partial Q_B}{\partial \theta_M}}{\dfrac{\partial Q_B}{\partial m}} = -\frac{1-\lambda}{\lambda} n \frac{\rho^2}{1-\rho}\eta\alpha < 0. \tag{6}$$

Hence, the Peter principle applies: the organization optimally trades off managerial quality and work ability.

Figure 1 illustrates the isoquant curves; their slope is given by the MRTS in (6). From (6), it is clear that as the relative importance of workers' effort over managerial ability increases (i.e., λ decreases), $|\text{MRTS}|$ increases (since MRTS < 0) and the isoquant curves become steeper—that is, more units of managerial quality can be substituted for a one-unit increase in ability as a worker to keep output constant. The same holds for increases in team size (n), the share of sensitive workers (ρ), their sensitivity (η), and the degree of interpersonal projection bias (α). When $\alpha = 0$, θ_M has no impact on period-2 output; thus, MRTS $= 0$ and the isoquant curves are horizontal. These observations imply the following:

Proposition 1. *If the candidates for promotion are unbiased, the organization promotes the worker with the highest managerial quality m. If the candidates exhibit interpersonal projection bias, it promotes the worker with (θ_M, m) who generates the highest period-2 output $Q_B(\theta_M, m)$. Moreover, the MRTS is negative and its magnitude $|MRTS|$—which measures the relative importance of ability as a worker θ_M compared to*

[9] Because insensitive workers exert zero effort, their output provides no new information about their ability. We assume the organization evaluates these workers using the ex-ante expected ability $\mathbb{E}[\theta]$. If an insensitive worker is promoted, the organization anticipates that they will hold managerial expectations based on $\theta_M = \mathbb{E}[\theta]$. Thus, their promotion case is evaluated on the pair $(\mathbb{E}[\theta], m)$.

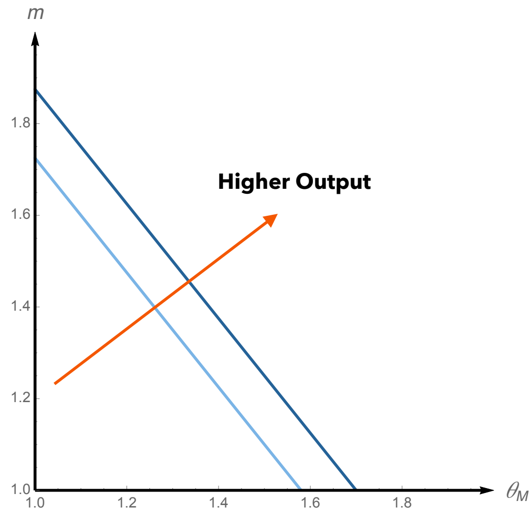

Figure 1: Isoquant curves

In the figure, $\theta \sim U[1,2]$ so $\mathbb{E}[\theta] = 1.5$. Parameter values are $\lambda = 2/3$, $n = 10$, $\rho = 1/2$, $\eta = 1$, $\alpha = 1/2$. The lighter blue line depicts the isoquant curve for $Q_B = 2.4$; the darker, for $Q_B = 2.5$.

managerial quality m in the promotion criterion—increases as n, ρ, η, or α increases and as λ decreases.

Thus, when promotion candidates are biased, the Peter principle can arise. The logic is as follows. A projection-biased worker with high ability as a worker, when promoted, forms higher expectations of her subordinates. These higher expectations induce greater effort from sensitive workers, as they seek to avoid falling short of those expectations or to exceed them. For this reason, the organization does not always promote the workers with the highest managerial ability, but trades off the candidate's managerial ability and his ability as a worker.[10]

Our explanation of the Peter principle differs starkly from the dominant explanation in the literature. In Baker et al. (1988), Milgrom and Roberts (1992), and Fairburn and Malcomson (2001), promoting high-performing workers provides effort incentives *ex ante*, but the organization suffers *ex post* because workers with high ability as workers do not necessarily possess high managerial ability. Consequently, such promotion rules may not persist when firms can demote managers with low managerial quality. In contrast, in our model, promoting high-performing workers provides effort incentives *ex post* to the subordinates of the newly promoted manager. Thus, promotion rules consistent with the Peter principle can persist even when demotion is possible.[11]

Proposition 1 yields several empirical implications. The model predicts that Peter-principle promotions are more likely to be observed when (i) the degree of interpersonal projection bias is higher (larger α); (ii) the proportion of sensitive workers is larger (higher ρ); (iii) the degree of sensitivity η is higher;[12] (iv) team size n is larger; and (v) managerial quality is relatively less important for output (lower λ). All of

[10] One might ask whether the organization could instruct any manager to announce high expectations, replicating the motivational benefit without sacrificing managerial quality. In our model, $\overline{q}$ is determined by the manager's genuine belief shaped by her own θ_M through projection bias, and workers correctly infer $\overline{q}$ from observable characteristics. An announced expectation detached from these would not be credible, and thus would lack motivational power—workers' motivation stems from their manager's genuine expectations rather than from announced targets, which are cheap talk.

[11] A recent empirical study by Mujcic and Oswald (2025) finds that demotions are quite commonplace, challenging the conventional wisdom that demotions are rare (Baker et al. (1994)). See also Howlett (2004) and Belzil and Bognanno (2008).

[12] Bohns and Flynn (2013) and Daniels and Robinson (2019) discuss workplace features affecting the strength of workers' guilt and shame in response to their performance, including worker autonomy, specificity of performance feedback, and outcome interdependence.

these strengthen the incentive effect of promoting a candidate with higher θ_M. Section 5 discusses how the core mechanism of our model can be empirically tested. Crucially, our theory predicts that Peter-principle promotions persist even in organizations that regularly demote underperforming managers—in contrast to the dominant explanation in the literature mentioned above, under which demotion possibilities may reduce or eliminate such promotions.

4 Extensions

4.1 External managers

In the basic model, we analyzed the case of promoting a current worker to a manager. In practice, however, firms may instead hire an external professional manager with superior managerial quality (e.g., leadership or coordination skills). We now compare these two options.

Suppose that the organization can hire a *professional manager* who has no experience or expertise in the tasks of workers and forms the unbiased expectation $\overline{q}_{UB}$ based on $\mathbb{E}[\theta]$, as given by (3). Let m^{OUT} denote the professional manager's managerial quality. The alternative is to promote a current worker—*specialist manager*—who forms the biased expectation $\overline{q}_B(\theta_M)$ based on their own ability θ_M, as given by (4). Let m^{IN} denote the specialist manager's managerial quality.

For simplicity, we maintain the same assumptions as in the basic model: the organization can perfectly infer candidates' expectations about their workers and observe the managerial quality of candidates and then selects a manager to maximize the expected final output in period 2.[13]

The expected period-2 output of a team led by a biased specialist manager with characteristics $(\theta_M, m^{\mathrm{IN}})$ is $Q_B(\theta_M, m^{\mathrm{IN}})$ as in (5). For a professional manager with managerial quality m^{OUT}, it is

$$Q_{UB}(m^{\mathrm{OUT}}) = (1-\lambda)n\rho\mathbb{E}[\overline{q}_{UB} + \eta\theta - 1] + \lambda m^{\mathrm{OUT}} = (1-\lambda)n\rho\left\{\frac{\rho}{1-\rho}\left(\eta\mathbb{E}[\theta]-1\right) + \eta\mathbb{E}[\theta] - 1\right\} + \lambda m^{\mathrm{OUT}}.$$

The organization hires the external manager if and only if $Q_{UB}(m^{\mathrm{OUT}}) > Q_B(\theta_M, m^{\mathrm{IN}})$, which reduces to

$$m^{\mathrm{OUT}} - m^{\mathrm{IN}} > \underbrace{\frac{(1-\lambda)}{\lambda}n\frac{\rho^2}{1-\rho}\eta\alpha}_{=|\mathrm{MRTS}|}(\theta_M - \mathbb{E}[\theta]).$$

The relative importance of ability as a worker versus managerial quality is measured by |MRTS| as in the basic model. If the internal candidate has a higher θ_M than $\mathbb{E}[\theta]$, the organization appoints the professional manager only if the managerial quality gap $m^{\mathrm{OUT}} - m^{\mathrm{IN}}$ exceeds the motivational advantage of the specialist manager's high expectations. Otherwise, promotion follows the Peter principle. Conversely, if the internal candidate has a lower θ_M than $\mathbb{E}[\theta]$, the professional manager is hired even with lower managerial quality.

This result offers implications for the choice between internal promotion and external hiring. The specialist manager's motivational advantage increases with team size n, worker sensitivity to expectations (high η and ρ), and the relative importance of workers' effort compared to managerial quality (low λ). These conditions are more likely to be met in teams where output depends primarily on individual effort—such as sales or engineering teams—rather than on coordination and leadership. Indeed, this result identifies when

[13] We also assume that if the internal candidate is promoted, the organization hires one additional worker from outside, so that the number of non-promoted workers is identical under both options.

promoting top performers can be justified as "the antithesis of hiring a general manager, someone whose expertise is really at managing," as quoted in the introduction.

4.2 Incentive contract

Since the motivational effect on workers drives the Peter principle in the basic model, a natural conjecture is that introducing incentive pay for workers *reduces* the benefit of promoting high-θ_M managers. To examine this, we consider a setting where individual performance q_i is contractible and each worker receives an exogenously given identical piece rate $\beta \geq 0$. We focus on a biased manager ($\alpha > 0$) and assume that β is not too large—$\beta < \min\{\frac{1}{\overline{\theta}}, \frac{1-\lambda}{2}\}$—to facilitate the analysis.

The worker's payoff is given by

$$U = w + \beta q - \frac{1}{\theta}e + \eta \ln(1 + q - \overline{q}).$$

Differentiating U with respect to e for the insensitive worker ($\eta = 0$) yields

$$\frac{\partial U}{\partial e} = \beta - \frac{1}{\theta}.$$

Since we assume $\beta < 1/\overline{\theta}$, for any θ, the insensitive worker chooses $e = 0$.

The FOC for the sensitive worker ($\eta > 0$) is given by

$$\frac{\partial U}{\partial e} = \beta - \frac{1}{\theta} + \frac{\eta}{1 + e - \overline{q}} = 0.$$

Solving yields

$$e = \overline{q} - 1 + \frac{\eta}{\frac{1}{\theta} - \beta} = \overline{q} - 1 + \frac{\eta\theta}{1 - \beta\theta}.$$

From this, we obtain $\frac{\partial e}{\partial \beta} = \frac{\eta\theta^2}{(1-\beta\theta)^2} > 0$ and $\frac{\partial^2 e}{\partial\beta\partial\theta} = \frac{2\eta\theta}{(1-\beta\theta)^3} > 0$. Thus, the incentive effect of the bonus is stronger for higher θ.

The biased manager's expectation is given by

$$\overline{q}_B(\theta_M) = \frac{\rho}{1-\rho}\left\{\eta\mathbb{E}\left[\frac{\hat{\theta}}{1-\beta\hat{\theta}} \mid \theta_M\right] - 1\right\}.$$

Finally, we consider how the bonus affects the MRTS. The expected period-2 profit of a team led by a biased manager is given by

$$\begin{aligned}\Pi_B(\theta_M, m) &= (1-\lambda)n\rho\mathbb{E}[e \mid \theta_M] + \lambda m - n\rho\beta\mathbb{E}[e \mid \theta_M] - nw \\ &= (1-\lambda-\beta)n\rho\left\{\frac{\rho}{1-\rho}\left(\eta\mathbb{E}\left[\frac{\hat{\theta}}{1-\beta\hat{\theta}} \mid \theta_M\right] - 1\right) + \eta\mathbb{E}\left[\frac{\theta}{1-\beta\theta}\right] - 1\right\} + \lambda m - nw.\end{aligned}$$

Since $\beta < 1-\lambda$, the bonus is lower than the marginal product of effort, implying that profits strictly increase with worker's effort. Differentiation yields

$$\frac{\partial\Pi_B(\theta_M, m)}{\partial\theta_M} = (1-\lambda-\beta)n\rho\frac{\partial\mathbb{E}[e(\hat{\theta}) \mid \theta_M]}{\partial\theta_M} = (1-\lambda-\beta)n\frac{\rho^2}{1-\rho}\eta\frac{\partial}{\partial\theta_M}\mathbb{E}\left[\frac{\hat{\theta}}{1-\beta\hat{\theta}} \mid \theta_M\right].$$

Let $b(\hat{\theta}) = \frac{\hat{\theta}}{1-\beta\hat{\theta}}$ and $y = \frac{\hat{\theta}-\alpha\theta_M}{1-\alpha}$, and then we have

$$\mathbb{E}\left[\frac{\hat{\theta}}{1-\beta\hat{\theta}} \mid \theta_M\right] = \int_{\underline{\theta}(\theta_M)}^{\overline{\theta}(\theta_M)} b(\hat{\theta})\hat{f}(\hat{\theta} \mid \theta_M)d\hat{\theta} = \int_{\underline{\theta}}^{\overline{\theta}} b((1-\alpha)y + \alpha\theta_M)f(y)dy.$$

Differentiation yields

$$\frac{\partial \mathbb{E}\left[\frac{\hat{\theta}}{1-\beta\hat{\theta}} \mid \theta_M\right]}{\partial \theta_M} = \int_{\underline{\theta}}^{\overline{\theta}} \frac{\partial}{\partial \theta_M} b((1-\alpha)y + \alpha\theta_M)dF(y) = \int_{\underline{\theta}(\theta_M)}^{\overline{\theta}(\theta_M)} \frac{\alpha}{(1-\beta\hat{\theta})^2} d\hat{F}(\hat{\theta} \mid \theta_M) = \alpha\mathbb{E}\left[\frac{1}{(1-\beta\hat{\theta})^2} \mid \theta_M\right] > 0.$$

Thus, with the incentive contract, the MRTS is given by

$$\text{MRTS} = \frac{dm}{d\theta_M} = -\frac{\frac{\partial \Pi_B}{\partial \theta_M}}{\frac{\partial \Pi_B}{\partial m}} = -\frac{(1-\lambda-\beta)}{\lambda} n \frac{\rho^2}{1-\rho} \eta\alpha\mathbb{E}\left[\frac{1}{(1-\beta\hat{\theta})^2} \mid \theta_M\right] < 0, \tag{7}$$

which reduces to (6) when $\beta = 0$.

Proposition 2. *Consider a bonus $\beta \in \left[0, \min\{\frac{1}{\overline{\theta}}, \frac{1-\lambda}{2}\}\right)$. If the candidates exhibit interpersonal projection bias, the organization promotes the worker with (θ_M, m) who maximizes period-2 profit $\Pi_B(\theta_M, m)$. Moreover, the MRTS is negative and its magnitude $|MRTS|$—measuring the relative importance of ability as a worker θ_M compared to managerial quality m in the promotion criterion—is increasing in θ_M. Furthermore, the effect of a higher bonus β on $|MRTS|$, $\frac{\partial |MRTS|}{\partial \beta}$, is nondecreasing in θ_M.*

With bonus β, the incentive effect of promoting a higher-θ_M candidate becomes:

$$\frac{\partial \mathbb{E}[e(\hat{\theta}) \mid \theta_M]}{\partial \theta_M} = \frac{\partial \overline{q}_B(\theta_M)}{\partial \theta_M} = \frac{\rho}{1-\rho}\eta\frac{\partial}{\partial \theta_M}\mathbb{E}\left[\frac{\hat{\theta}}{1-\beta\hat{\theta}} \mid \theta_M\right] = \frac{\rho}{1-\rho}\eta\alpha\mathbb{E}\left[\frac{1}{(1-\beta\hat{\theta})^2} \mid \theta_M\right] > 0,$$

which is increasing in both θ_M and β. Thus, the incentive effect strengthens as θ_M and β increase (the former making isoquant curves concave in θ_M). The reason is that the bonus has a stronger effect on higher types ($\frac{\partial^2 e}{\partial \beta \partial \theta} > 0$), and a higher-$\theta_M$ manager perceives workers as higher types, thus holding higher $\overline{q}_B(\theta_M)$, further strengthening the effect of the bonus on higher types. On the other hand, the organization retains a smaller share of output as β increases. Whether the former effect dominates the latter depends on how the MRTS changes with β. To see this formally,

$$\frac{d|\text{MRTS}|}{d\beta} = \frac{1}{\lambda} n \frac{\rho^2}{1-\rho}\eta\alpha\frac{\partial}{\partial \beta}\mathbb{E}\left[\frac{1-\lambda-\beta}{(1-\beta\hat{\theta})^2} \mid \theta_M\right] = \frac{1}{\lambda} n \frac{\rho^2}{1-\rho}\eta\alpha\mathbb{E}\left[\frac{2(1-\lambda)\hat{\theta}-\beta\hat{\theta}-1}{(1-\beta\hat{\theta})^3} \mid \theta_M\right]$$

where $\frac{\partial}{\partial \hat{\theta}}\frac{2(1-\lambda)\hat{\theta}-\beta\hat{\theta}-1}{(1-\beta\hat{\theta})^3} = \frac{2(1-\lambda-2\beta)+4\beta(1-\lambda-\frac{1}{2}\beta)\hat{\theta}}{(1-\beta\hat{\theta})^4} > 0$ by $\beta < \frac{1-\lambda}{2}$. Thus, since $\hat{F}(\cdot \mid \theta_M)$ first-order stochastically dominates $\hat{F}(\cdot \mid \theta'_M)$ for $\theta_M > \theta'_M$, $\mathbb{E}\left[\frac{2(1-\lambda)\hat{\theta}-\beta\hat{\theta}-1}{(1-\beta\hat{\theta})^3} \mid \theta_M\right]$ is nondecreasing in θ_M.[14] Note that $\mathbb{E}\left[\frac{2(1-\lambda)\hat{\theta}-\beta\hat{\theta}-1}{(1-\beta\hat{\theta})^3} \mid \theta_M\right]$ tends to be positive when $\hat{F}(\cdot \mid \theta_M)$ has most probability mass above $\frac{1}{2(1-\lambda)-\beta}$, which is more likely for a higher-θ_M candidate.

Therefore, the conjecture—higher β reduces the benefit of promoting a higher-θ_M candidate—tends to

[14] See Observation 1 of Gagnon-Bartsch et al. (2021).

be false when θ_M is high. The incentive effect of promoting a higher-θ_M candidate is magnified when θ_M is higher: a higher-θ_M biased manager perceives workers as higher types. With higher β, since the bonus's incentive effect is stronger for higher types, the manager's expectations rise, potentially increasing output at a rate that exceeds the reduction in the organization's share $(1-\beta)$. Thus, when bonuses are offered, the promotion criterion may become polarized as β increases: θ_M's importance rises for high-θ_M candidates but falls for low-θ_M candidates.

Incentive pay and Peter-principle promotions can therefore complement each other. This result generates a prediction that refines the finding of Benson et al. (2019), who document empirically that Peter-principle promotions are less prevalent when incentive pay is stronger. Our model suggests that this substitution pattern should be weaker—and may reverse—in settings where workers are strongly motivated by meeting their manager's expectations and where high-performing workers tend to hold higher expectations once promoted.

4.3 Quits

Our basic model highlights the benefits of high managerial expectations. However, expectations can be too high, causing excessive turnover. To examine this, consider workers' quit decisions at the start of period 2, after learning their type and their manager's identity, and hence expectations. Because a sensitive worker's expected utility from staying increases with his ability, and excessive expectations do not affect insensitive workers, we consider a situation where only sensitive worker i with $\theta_i < \theta^* \in (\underline{\theta}, \overline{\theta})$ quits, while all insensitive workers stay.[15] We focus on fully biased managers ($\alpha = 1$). This greatly facilitates the analysis because fully biased managers with θ_M hold the same expectation $\overline{q}_{FB}(\theta_M)$ about worker performance independent of worker quits. They simply believe that all workers have the same θ as they have.

Consider a sensitive worker's decision whether to stay after learning the manager's expectation. From (2) and the FOC for sensitive worker's effort derived in section 3.1, the expected payoff upon staying for the worker with θ is given by

$$U = w - \frac{1}{\theta}(\overline{q} + \eta\theta - 1) + \eta \ln(\eta\theta).$$

Suppose w yields a threshold $\theta^* \in (\underline{\theta}, \overline{\theta})$ such that only worker i with $\theta_i \geq \theta^*$ stays:

$$w = \eta + \frac{1}{\theta^*}(\overline{q} - 1) - \eta \ln(\eta\theta^*) + \overline{U}. \tag{8}$$

The right-hand side of (8) increases with $\overline{q}$ and decreases with θ^*; thus θ^* must increase with $\overline{q}$.

At the end of period 1, the organization decides which workers to promote. Since all workers receive the same wage w and only sensitive workers with $\theta < \theta^*$ quit the job, the expected period-2 profit of a team led by a fully-biased manager with characteristics (θ_M, m) is given by

$$\begin{aligned}\Pi_{FB}(\theta_M, m) &= (1-\lambda)n\rho[1-F(\theta^*)]\mathbb{E}[\overline{q}_{FB}(\theta_M) + \eta\theta - 1 \mid \theta \geq \theta^*] + \lambda m - n[1-\rho F(\theta^*)]w \\ &= (1-\lambda)n\rho[1-F(\theta^*)]\left\{\frac{\rho}{1-\rho}(\eta\theta_M - 1) + \eta\mathbb{E}[\theta \mid \theta \geq \theta^*] - 1\right\} + \lambda m - n[1-\rho F(\theta^*)]w\end{aligned}$$

[15] The Online Appendix provides a sufficient condition for this equilibrium.

where θ^* is given by (8); that is,

$$w = \eta + \frac{1}{\theta^*}(\overline{q}_{FB}(\theta_M) - 1) - \eta \ln(\eta\theta^*) + \overline{U} = \eta + \frac{\rho\eta\theta_M - 1}{(1-\rho)\theta^*} - \eta \ln(\eta\theta^*) + \overline{U}.$$

The organization promotes the worker with (θ_M, m) who generates the highest period-2 profit $\Pi_{FB}(\theta_M, m)$. We have

$$\frac{d\mathbb{E}[\theta \mid \theta_i \geq \theta^*]}{d\theta^*} = \frac{d}{d\theta^*}\frac{\int_{\theta^*}^{\overline{\theta}} \theta dF(\theta)}{1-F(\theta^*)} = \frac{f(\theta^*)}{1-F(\theta^*)}\left(\mathbb{E}[\theta \mid \theta_i \geq \theta^*] - \theta^*\right) > 0,$$

$$\frac{d\theta^*}{d\theta_M} = -\frac{\frac{\partial w}{\partial \theta_M}}{\frac{\partial w}{\partial \theta^*}} = \frac{\frac{\rho\eta}{(1-\rho)\theta^*}}{\frac{\rho\eta\theta_M - 1}{(1-\rho)\theta^{*2}} + \frac{\eta}{\theta^*}} = \frac{\rho\eta\theta^*}{\rho\eta\theta_M - 1 + (1-\rho)\eta\theta^*} > 0.$$

Using these two equations, we have

$$\begin{aligned}
\frac{d\Pi_{FB}}{d\theta_M} &= \frac{\partial \Pi_{FB}}{\partial \theta_M} + \frac{\partial \Pi_{FB}}{\partial \theta^*}\frac{d\theta^*}{d\theta_M} \\
&= (1-\lambda)n\frac{\rho^2}{1-\rho}[1-F(\theta^*)]\eta \\
&\quad + (1-\lambda)n\rho\left\{-f(\theta^*)\left[\frac{\rho}{1-\rho}(\eta\theta_M - 1) + \eta\mathbb{E}[\theta \mid \theta \geq \theta^*] - 1\right] + \eta[1-F(\theta^*)]\frac{d\mathbb{E}[\theta \mid \theta_i \geq \theta^*]}{d\theta^*} + \frac{f(\theta^*)w}{1-\lambda}\right\}\frac{d\theta^*}{d\theta_M} \\
&= (1-\lambda)n\frac{\rho^2}{1-\rho}[1-F(\theta^*)]\eta - (1-\lambda)n\rho f(\theta^*)\left\{\frac{\rho}{1-\rho}(\eta\theta_M - 1) - 1 + \eta\theta^* - \frac{w}{1-\lambda}\right\}\frac{d\theta^*}{d\theta_M} \\
&= (1-\lambda)n\frac{\rho^2}{1-\rho}[1-F(\theta^*)]\eta - (1-\lambda)n\rho f(\theta^*)\left\{\frac{\rho\eta\theta_M - 1 + (1-\rho)\eta\theta^*}{1-\rho} - \frac{w}{1-\lambda}\right\}\frac{\rho\eta\theta^*}{\rho\eta\theta_M - 1 + (1-\rho)\eta\theta^*} \\
&= n\frac{\rho^2}{1-\rho}\eta\left\{(1-\lambda)[1-F(\theta^*) - f(\theta^*)\theta^*] + \frac{f(\theta^*)\theta^* w}{\frac{\rho\eta\theta_M - 1}{1-\rho} + \eta\theta^*}\right\}.
\end{aligned}$$

Thus, the MRTS is given by

$$\text{MRTS} = \frac{dm}{d\theta_M} = -\frac{\frac{d\Pi_{FB}}{d\theta_M}}{\frac{\partial \Pi_{FB}}{\partial m}} = -\frac{1}{\lambda}n\frac{\rho^2}{1-\rho}\eta\underbrace{\left\{(1-\lambda)[1-F(\theta^*)] - (1-\lambda)\theta^* f(\theta^*) + \frac{\theta^* f(\theta^*)}{\frac{\rho\eta\theta_M - 1}{1-\rho} + \eta\theta^*}w\right\}}_{\equiv \Psi}. \quad (9)$$

When $\Psi = 1-\lambda$, (9) reduces to (6) in the basic model. We have

$$\frac{\partial \Psi}{\partial \theta^*} = (1-\lambda)[-2f(\theta^*) - f'(\theta^*)\theta^*] + \frac{\left\{[f'(\theta^*)\theta^* + f(\theta^*)]\left(\frac{\rho\eta\theta_M - 1}{1-\rho} + \eta\theta^*\right) - \eta f(\theta^*)\theta^*\right\}w}{\left(\frac{\rho\eta\theta_M - 1}{1-\rho} + \eta\theta^*\right)^2},$$

$$\frac{d\Psi}{d\theta_M} = \underbrace{\frac{\partial \Psi}{\partial \theta_M}}_{<0} + \frac{\partial \Psi}{\partial \theta^*}\underbrace{\frac{d\theta^*}{d\theta_M}}_{>0}.$$

If the density f does not decrease too rapidly (f' can be negative but its magnitude is not too large) and w is sufficiently small, then $\frac{\partial \Psi}{\partial \theta^*} < 0$, which implies $\frac{d\Psi}{d\theta_M} < 0$. Therefore, the MRTS increases with θ_M.

Proposition 3. *Suppose that, at the beginning of period 2, only sensitive workers with* $\theta_i < \theta^* \in (\underline{\theta}, \overline{\theta})$

quit after learning their manager's expectations. If the candidates are fully projection-biased, the organization promotes the worker with (θ_M, m) who generates the highest period-2 profit $\Pi_{FB}(\theta_M, m)$. Under full projection bias, θ^* increases with θ_M and decreases with w; furthermore, MRTS increases with θ^* and with θ_M—isoquant curves are convex—provided the density f does not decrease too rapidly and the wage w is sufficiently small.

As in the basic model, promoting a fully biased candidate with higher θ_M increases period-2 output, because higher manager's expectations induce greater effort. When workers can quit after learning their manager's expectations, this output gain is reduced, as seen in the first term within the curly brackets in (9), because a fraction $F(\theta^*)$ of sensitive workers quits. The increase in quits not only scales down output gains, but also has a direct effect on period-2 output, and the magnitude of this reduction increases with $\theta^* f(\theta^*)$, as seen in the second term within the curly brackets in (9). The reason is that the marginal worker with ability θ^* quits with likelihood $f(\theta^*)$ in response to higher expectations, and the forgone output is larger when $f(\theta^*)\theta^*$ is higher. Moreover, lower retention saves on wages, which is captured by the last term within the curly brackets in (9).

To understand why MRTS increases with θ_M, note that the sign of MRTS equals the sign of $-\Psi$ in (9), where Ψ depends on θ^* and measures the importance of θ_M relative to m in promotion. As θ_M rises, θ^* increases (more quits), reducing the motivational gain from higher θ_M. However, higher θ^* may also increase profit through greater wage savings (proportional to w) and a change in output loss from the marginal quitter (proportional to $\theta^* f'(\theta^*)$). If w is sufficiently small and $f'(\theta^*)$ is not too negative, these positive effects are outweighed, yielding $\partial\Psi/\partial\theta^* < 0$, implying $d\Psi/d\theta_M < 0$. Consequently, as θ_M increases, the MRTS rises, implying convex isoquant curves: the relative importance of θ_M compared to m falls with θ_M.

Thus, if θ_M is sufficiently high or w is sufficiently low, θ^* rises enough that the MRTS may turn positive, implying θ_M receives a *negative* marginal score: additional ability is detrimental because promoting a higher-θ_M candidate induces too many quits (including among high-ability workers), so output loss from quits outweighs wage savings and output gains from greater effort by those who stay.

Conversely, if θ_M is not too high or w not too low, output loss from quits is outweighed by wage savings and output gains from greater effort by those who stay, so promoting a higher-θ_M candidate is beneficial—Peter-principle promotions survive.

5 Concluding remarks

We have developed a model of promotions of workers to managerial positions. Organizations may optimally promote high-performing workers—even those with limited managerial aptitude ("Peter-principle" promotions)—because they hold high expectations about subordinates, spurring motivation to meet their manager's expectations. Critically, our theory requires guilt aversion and interpersonal projection bias—both supported by extensive empirical evidence.

Our basic model highlights the benefits of high managerial expectations for organizational performance. However, as shown in Section 4.3, managers' expectations can be too high, causing excessive turnover. Kamphorst and Swank (2013) offer two other drawbacks of managers' excessive expectations: (i) inducing workers to pursue excessively risky projects and (ii) reducing workers' incentive to impress managers, thereby lowering effort. These effects do not arise in our model because workers do not make any project choice and ability is common knowledge when choosing effort.

Our theory helps explain divergence in promotion practices across firms. Provided that managers' expectations do not become so high as to trigger excessive turnover, it predicts that Peter-principle promotions are more likely in organizations or cultures where people care more about meeting others' expectations, where interpersonal projection bias is stronger, and where managerial quality matters less for organizational performance than workers' motivation. Future empirical work can consider these factors in explaining the variation in promotion practices. Crucial elements of our theory can also be tested in laboratory experiments, where subjects can be put in the role of workers managed by former workers. Such experiments could test whether workers infer that managers who performed better in the past hold higher expectations, whether this improves worker performance, and whether managers' beliefs correlate with their own performance. We view the evidence in Bushong and Gagnon-Bartsch (2024) as a very encouraging first step.

References

Avery, Derek R, Scott Tonidandel, Kristin H Griffith, and Miguel A Quinones (2003) "The impact of multiple measures of leader experience on leader effectiveness: New insights for leader selection," *Journal of Business Research*, 56 (8), 673–679.

Baker, George, Michael Gibbs, and Bengt Holmstrom (1994) "The internal economics of the firm: Evidence from personnel data," *Quarterly Journal of Economics*, 109 (4), 881–919.

Baker, George P, Michael C Jensen, and Kevin J Murphy (1988) "Compensation and incentives: Practice vs. theory," *Journal of Finance*, 43 (3), 593–616.

Battigalli, Pierpaolo and Martin Dufwenberg (2007) "Guilt in games," *American Economic Review*, 97 (2), 170–176.

——— (2022) "Belief-dependent motivations and psychological game theory," *Journal of Economic Literature*, 60 (3), 833–882.

Belzil, Christian and Michael Bognanno (2008) "Promotions, demotions, halo effects, and the earnings dynamics of American executives," *Journal of Labor Economics*, 26 (2), 287–310.

Bénabou, Roland and Jean Tirole (2003) "Intrinsic and extrinsic motivation," *Review of Economic Studies*, 70 (3), 489–520.

Benson, Alan, Danielle Li, and Kelly Shue (2019) "Promotions and the Peter Principle," *Quarterly Journal of Economics*, 134 (4), 2085–2134.

Bohns, Vanessa K and Francis J Flynn (2013) "Guilt by design: Structuring organizations to elicit guilt as an affective reaction to failure," *Organization Science*, 24 (4), 1157–1173.

Buchanan, Joy A (2020) "My reference point, not yours," *Journal of Economic Behavior and Organization*, 171, 297–311.

Bushong, Benjamin and Tristan Gagnon-Bartsch (2024) "Failures in Forecasting: An Experiment on Interpersonal Projection Bias," *Management Science*, 70, 8735–8752.

Cartwright, Edward (2019) "A survey of belief-based guilt aversion in trust and dictator games," *Journal of Economic Behavior and Organization*, 167, 430–444.

Cordes, Simon and Max Müller (2026) "Do Firms Know What Workers Want?" *Mimeo*, University of Bonn.

Daniels, Michael A and Sandra L Robinson (2019) "The shame of it all: A review of shame in organizational life," *Journal of Management*, 45 (6), 2448–2473.

Dickinson, David L and Marie Claire Villeval (2012) "Job allocation rules and sorting efficiency: experimental outcomes in a Peter Principle environment," *Southern Economic Journal*, 78 (3), 842–859.

Eden, Dov (1992) "Leadership and expectations: Pygmalion effects and other self-fulfilling prophecies in organizations," *Leadership Quarterly*, 3 (4), 271–305.

Engelmann, Dirk and Martin Strobel (2012) "Deconstruction and reconstruction of an anomaly," *Games and Economic Behavior*, 76 (2), 678–689.

Epley, Nicholas, Boaz Keysar, Leaf Van Boven, and Thomas Gilovich (2004) "Perspective taking as egocentric anchoring and adjustment.," *Journal of Personality and Social Psychology*, 87 (3), 327.

Fahn, Matthias and Nicolas Klein (2025) "The Dynamics of Exploiting Overconfident Workers," *Journal of Labor Economics*, forthcoming.

Fairburn, James A and James M Malcomson (2001) "Performance, promotion, and the Peter Principle," *Review of Economic Studies*, 68 (1), 45–66.

Fang, Dawei and Thomas Noe (2022) "Less competition, more meritocracy?" *Journal of Labor Economics*, 40 (3), 669–701.

Gagnon-Bartsch, Tristan, Marco Pagnozzi, and Antonio Rosato (2021) "Projection of private values in auctions," *American Economic Review*, 111 (10), 3256–3298.

Gagnon-Bartsch, Tristan and Antonio Rosato (2024) "Quality is in the eye of the beholder: taste projection in markets with observational learning," *American Economic Review*, 114 (11), 3746–3787.

Goodall, Amanda H and Agnes Bäker (2014) "A theory exploring how expert leaders influence performance in knowledge-intensive organizations," in *Incentives and performance: Governance of research organizations*, 49–67: Springer.

Howlett, Peter (2004) "The internal labour dynamics of the Great Eastern Railway Company, 1870–1913," *The Economic History Review*, 57 (2), 396–422.

Ishida, Junichiro (2006) "Optimal promotion policies with the looking-glass effect," *Journal of Labor Economics*, 24 (4), 857–877.

Kamphorst, Jurjen JA and Otto H Swank (2013) "When Galatea cares about her reputation: How having faith in your workers reduces their motivation to shine," *European Economic Review*, 60, 91–104.

Khalmetski, Kiryl, Axel Ockenfels, and Peter Werner (2015) "Surprising gifts: Theory and laboratory evidence," *Journal of Economic Theory*, 159, 163–208.

Kierein, Nicole M. and Michael A. Gold (2000) "Pygmalion in work organizations: a meta-analysis," *Journal of Organizational Behavior*, 21 (8), 913–928.

Krueger, Joachim and Russell W Clement (1994) "The truly false consensus effect: an ineradicable and egocentric bias in social perception.," *Journal of Personality and Social Psychology*, 67 (4), 596–610.

Lazear, Edward P. (2004) "The Peter Principle: A Theory of Decline," *Journal of Political Economy*, 112 (S1), S141–S163.

Légeret, Matthieu, Christian Zehnder, and Benjamin Tur (2025) "How good can bad leaders be? The opportunity costs of leader selection," *Leadership Quarterly*, 36 (2), 101856.

Marks, Gary and Norman Miller (1987) "Ten years of research on the false-consensus effect: An empirical and theoretical review," *Psychological Bulletin*, 102 (1), 72–90.

Milgrom, Paul R and John Roberts (1992) *Economics, Organization and Management*: Prentice-hall Englewood Cliffs, NJ.

Morita, Kimiyuki, Akitoshi Muramoto, and Takeharu Sogo (2025) "Optimal contracts under interpersonal projection," *Games and Economic Behavior*, 150, 356–364.

Mujcic, Redzo and Andrew J. Oswald (2025) "The growth and collapse of autonomy at work," *Proceedings of the National Academy of Sciences*, 122 (27), e2504419122.

Peter, Laurence J and Raymond Hull (1969) *The Peter Principle*: William Morrow.

Ross, Lee, David Greene, and Pamela House (1977) "The "false consensus effect": An egocentric bias in social perception and attribution processes," *Journal of Experimental Social Psychology*, 13 (3), 279–301.

Schleu, Joyce Elena and Joachim Hüffmeier (2021) "Simply the best? A systematic literature review on the predictive validity of employee performance for leader performance," *Human Resource Management Review*, 31 (2), 100777.

Swank, Otto H. and Bauke Visser (2007) "Motivating through delegating tasks or giving attention," *Journal of Law, Economics, and Organization*, 23 (3), 731–742.

Van Boven, Leaf and George Loewenstein (2003) "Social projection of transient drive states," *Personality and Social Psychology Bulletin*, 29 (9), 1159–1168.

Van Boven, Leaf, George Loewenstein, and David Dunning (2003) "Mispredicting the endowment effect: Underestimation of owners' selling prices by buyer's agents," *Journal of Economic Behavior and Organization*, 51 (3), 351–365.

Weidmann, Ben, Joseph Vecci, Farah Said, David J Deming, and Sonia R Bhalotra (2026) "How Do You Find a Good Manager?" *Quarterly Journal of Economics*, 141 (2), 1581–1633.

6 Online Appendix

6.1 Quadratic Cost Function Without Insensitive Workers

If all workers are sensitive ($\rho = 1$) and the cost function is linear, then $\frac{\partial e}{\partial \overline{q}} = 1$ holds, which prevents the manager's rational expectation. However, even if insensitive workers do not exist ($\rho = 1$), a rational expectation exists if $\frac{\partial e}{\partial \overline{q}} < 1$. To demonstrate, suppose $\rho = 1$ and cost function $c(e;\theta) = \frac{e^2}{2\theta}$. We show that (i) a unique rational expectation of a manager exists, and (ii) a biased manager's expectation is nondecreasing in θ_M.

The worker's utility function is given by

$$U = w - \frac{e^2}{2\theta} + \eta \ln(1 + q - \overline{q}).$$

The first-order condition is

$$-\frac{e}{\theta} + \frac{\eta}{1 + e - \overline{q}} = 0.$$

Solving yields the optimal effort $e(\theta \mid \overline{q})$ for type θ given $\overline{q}$:

$$e(\theta \mid \overline{q}) = \frac{-(1-\overline{q}) + \sqrt{(1-\overline{q})^2 + 4\eta\theta}}{2} > 0,$$

which is increasing in both θ and $\overline{q}$. For the latter, the implicit function theorem gives

$$\frac{\partial e(\theta \mid \overline{q})}{\partial \overline{q}} = \frac{e(\theta \mid \overline{q})}{[1 + e(\theta \mid \overline{q}) - \overline{q}] + e(\theta \mid \overline{q})} \in (0,1).$$

In what follows, for a manager with θ_M and $\alpha \in [0,1]$ (unbiased, biased, or fully biased), we prove the unique existence of her rational expectation. The manager's expectation is defined by the solution to

$$H(\overline{q}) \equiv \mathbb{E}[e(\hat{\theta} \mid \overline{q}) \mid \theta_M] - \overline{q} = 0.$$

We first show $H'(\overline{q}) < 0$. Since $\frac{\partial e(\theta|\overline{q})}{\partial \overline{q}} \in (0,1)$ for any θ, we have

$$H'(\overline{q}) = \mathbb{E}\left[\frac{\partial e(\hat{\theta} \mid \overline{q})}{\partial \overline{q}} \mid \theta_M\right] - 1 < 0.$$

Thus, to prove the unique existence of the manager's rational expectation, it suffices to show $H(0) > 0$ and $H(\overline{q}) < 0$ for sufficiently large $\overline{q}$.

At $\overline{q} = 0$, we have

$$H(0) = \mathbb{E}[e(\hat{\theta} \mid 0) \mid \theta_M] - 0 = \mathbb{E}\left[\frac{-1 + \sqrt{1 + 4\eta\hat{\theta}}}{2} \mid \theta_M\right] > 0.$$

We now show $H(\overline{q}) < 0$ for sufficiently large $\overline{q}$. Since $e(\theta \mid \overline{q})$ is increasing in θ and $\hat{\theta}$ distributes over $(\underline{\theta}(\theta_M), \overline{\theta}(\theta_M)) \equiv ((1-\alpha)\underline{\theta} + \alpha\theta_M, (1-\alpha)\overline{\theta} + \alpha\theta_M)$, the following inequality holds:

$$H(\overline{q}) = \mathbb{E}\left[e(\hat{\theta} \mid \overline{q}) \mid \theta_M\right] - \overline{q} < e\left(\overline{\theta} \mid \overline{q}\right) - \overline{q}.$$

Thus, we have

$$H(\overline{q}) \leq e\left(\overline{\theta} \mid \overline{q}\right) - \overline{q} = -\frac{(1+\overline{q})}{2} + \frac{1}{2}\sqrt{(\overline{q}-1)^2 + 4\eta\overline{\theta}} < -\frac{(1+\overline{q})}{2} + \frac{1}{2}\left((\overline{q}-1) + \frac{2\eta\overline{\theta}}{(\overline{q}-1)}\right) = -1 + \frac{\eta\overline{\theta}}{\overline{q}-1}$$

where the last inequality holds because $\sqrt{(\overline{q}-1)^2 + 4\eta\overline{\theta}} < (\overline{q}-1) + \frac{2\eta\overline{\theta}}{(\overline{q}-1)}$:

$$\sqrt{(\overline{q}-1)^2 + 4\eta\overline{\theta}} - (\overline{q}-1) = \frac{\left\{\sqrt{(\overline{q}-1)^2 + 4\eta\overline{\theta}}\right\}^2 - (\overline{q}-1)^2}{\sqrt{(\overline{q}-1)^2 + 4\eta\overline{\theta}} + (\overline{q}-1)} = \frac{4\eta\overline{\theta}}{\sqrt{(\overline{q}-1)^2 + 4\eta\overline{\theta}} + (\overline{q}-1)} < \frac{4\eta\overline{\theta}}{2(\overline{q}-1)}.$$

Hence, $H(\overline{q}) < 0$ for sufficiently large $\overline{q}$, and hence a unique rational expectation exists.

Finally, we show $\frac{\partial \overline{q}_B(\theta_M)}{\partial \theta_M} \geq 0$ for $\alpha \in (0, 1]$. The implicit function theorem yields

$$\frac{\partial \overline{q}_B(\theta_M)}{\partial \theta_M} = -\frac{\partial H(\overline{q}_B(\theta_M))/\partial \theta_M}{H'(\overline{q}_B(\theta_M))}.$$

Since $H'(\overline{q}) < 0$ as shown above, it suffices to show $\partial H(\overline{q}_B(\theta_M))/\partial \theta_M \geq 0$ to prove $\frac{\partial \overline{q}_B(\theta_M)}{\partial \theta_M} \geq 0$. We have

$$\left.\frac{\partial H(\overline{q})}{\partial \theta_M}\right|_{\overline{q}=\overline{q}_B(\theta_M)} = \left.\frac{\partial}{\partial \theta_M}\mathbb{E}\left[e(\hat{\theta} \mid \overline{q}) \mid \theta_M\right]\right|_{\overline{q}=\overline{q}_B(\theta_M)} \geq 0$$

where the inequality holds because $e(\theta \mid \overline{q})$ is increasing in θ for any $\overline{q}$ and $\hat{F}(\cdot \mid \theta_M)$ first-order stochastically dominates $\hat{F}(\cdot \mid \theta'_M)$ for $\theta_M > \theta'_M$.

6.2 Model with Quits

In the main text, we focus on the equilibrium where only sensitive worker i with $\theta_i < \theta^* \in (\underline{\theta}, \overline{\theta})$ quits, while all insensitive workers stay. In this appendix, we provide a sufficient condition for this equilibrium: a condition under which all insensitive workers stay whenever some sensitive workers stay.

Each worker stays if and only if their utility upon staying exceeds reservation utility $\overline{U}$. Since each insensitive worker chooses $e = 0$ and each sensitive worker with θ chooses $e(\theta) \equiv \overline{q} + \eta\theta - 1$ upon staying, all insensitive workers stay if and only if

$$w \geq \overline{U},$$

and a sensitive worker with θ stays if and only if

$$w - \frac{e(\theta)}{\theta} + \eta \ln(\eta\theta) = w - \eta - \frac{1}{\theta}(\overline{q}-1) + \eta \ln(\eta\theta) \geq \overline{U}.$$

Thus, all insensitive workers stay whenever some sensitive workers stay if for any $\theta \in [\underline{\theta}, \overline{\theta}]$,

$$h(\theta, \eta) \equiv -\frac{e(\theta)}{\theta} + \eta \ln(\eta\theta) = -\eta - \frac{1}{\theta}(\overline{q}-1) + \eta \ln(\eta\theta) \leq 0.$$

Since $\eta\theta > 1$ by Assumption 1, we have

$$\frac{\partial h(\theta, \eta)}{\partial \theta} = \frac{\overline{q} - 1 + \eta\theta}{\theta^2} > 0 \quad \text{and} \quad \frac{\partial h(\theta, \eta)}{\partial \eta} = \ln(\eta\theta) > 0.$$

Since h increases with θ, the above condition reduces to $h(\overline{\theta}, \eta) \leq 0$. In words, this condition ensures that the effort cost $\frac{e(\overline{\theta})}{\overline{\theta}}$ is not outweighed by the utility $\eta \ln(\eta\overline{\theta})$ from exceeding the manager's expectation for the most able sensitive type. This holds if η is not too large relative to the cost parameter $1/\overline{\theta}$ (but not too small, so that Assumption 1 is satisfied).