# Hiding Lemons among Peaches: Optimal Retention and Promotion Policy Design

Andrew B. Choi and Takuma Habu*

13 December 2024

– Click here for latest version –

**Abstract**

How should an employer design its retention and promotion policy when there is competition for its employees? In an environment where the incumbent employer learns more about its workers than other potential employers, we show that the incumbent's optimal policy always over-retains workers—retaining workers that it would have let go if all employers had learnt about workers symmetrically—to dampen the positive signalling effect of retention. In contrast, the optimal policy may over- or under-promote workers. The incumbent's incentive to distort its policy is driven by the technologies of competing firms in the labour market. We demonstrate the extent to which the firm can implement the optimal policy without the ability to commit and study the incentive for the firm to manipulate the signalling effect via designing jobs. Our results shed light on the role of (possibly vacuous) job titles and provide a novel rationale for the Peter Principle.

*Andrew B. Choi: Department of Economics, University of Michigan, andybchoi@gmail.com. Takuma Habu: Department of Economics, Cornell University, takumahabu@cornell.edu. We thank Marco Battaglini, Raphael Boleslavsky, Bart de Koning, Thomas Jungbauer, Emir Kamenica, Michael Gibbs, Robert Gibbons, Rohit Lamba, Xiao Lin, Elliot Lipnowski, Doron Ravid, Collin Raymond, Lars Stole, Michael Waldman, and Kai Hao Yang for helpful comments.

# 1 Introduction

How should a firm retain or promote workers who can leave for another employer? When an employer has private information about its workers, retention or promotion serves as a positive signal about workers' productivity to other potential employers, driving up the wage.[1] In this paper, we explore whether a firm would dampen the positive signalling effect by retaining or promoting some "lemon" workers along with some "peach" workers, despite the costs from doing so. After all, workplaces abound with stories of managers who do not deserve to be managers. We show that the firm always finds it optimal to over-retain; i.e., retain workers that it would have let go under complete information; however, whether it would over- or under-promote workers is ambiguous.

Specifically, we consider an incumbent firm with a unit mass of junior workers. The firm must decide who to retain and, among those retained, who to promote to a senior position. The workers differ in their abilities, and the output from a worker in a given position increases linearly in ability. The output from the senior position is more sensitive to the worker's ability, and a worker produces more in the senior position if and only if the worker's ability is above a threshold.

The labour market also consists of heterogeneous outside firms with linear production technologies. To retain a worker (either in a junior or a senior position), the incumbent firm pays a wage equal to the worker's outside option—the highest wage that the worker can command from outside firms. To focus on the signalling effects of the incumbent's retention/promotion decisions, we assume that outside firms only observe whether each worker was retained and, if so, whether the worker was promoted. The assumptions together imply that a worker's wage is an increasing and convex function of the average ability of all workers in the position.

The incumbent firm's problem is to maximize its profits by designing an HR policy that specifies retention and promotion probabilities that are increasing in ability. We are interested in how the firm's optimal HR policy differs from the benchmark HR policy that the firm would have adopted under complete information. In fact, due to competition, the complete-information benchmark is equivalent to the efficient allocation of workers that maximises output across all firms. We focus on the case in which it is efficient for the firm to retain workers if and only if their ability is above a certain threshold.

[1] While the labour literature has long studied the signalling effect of promotions since Waldman (1984), little attention has been paid to the signalling effect of retention let alone on the joint signalling effects of retention and promotions.

Our first main finding is that any optimal HR policy must over-retain workers relative to the benchmark (Theorem 1); i.e., the firm always retains some workers whose output would be higher in outside firms. Retaining a marginal worker is beneficial for the firm because it lowers the average ability, and thus the wage cost, of all retained workers; however, doing so is also costly because the worker's output would be lower than the wage based on the average ability. We show that the benefit strictly dominates the cost for the worker for whom it is just efficient for the firm to retain and that this, in turn, implies the firm would promote workers with lower abilities.

This result is driven by the convexity of the worker's outside option arising from the heterogeneity in the production technologies of outside firms. Intuitively, because the firm pays workers their outside options that depend on the average ability of retained workers (which must be greater than the marginally retained worker's ability), convexity ensures it is worthwhile for the incumbent to lower the average by over-retaining workers.

Our second main finding is that an optimal HR policy may exhibit over-promotion, under-promotion, or even both if the policy is non-deterministic (Theorem 2). The same rationale for over-retention means that over-promoting always increases the firm's profits from the promoted workers. However, whereas over-retention does not affect the firm's profits from workers who are let go, over-promotion always decreases the firm's profits from workers in the junior position because the marginally promoted worker was the best among them. The comparison between these two opposing effects can go either way depending on the parameters of the model.

We demonstrate that inefficient promotions also arise from the heterogeneity in production technologies among outside firms by showing that there is a sense in which the likelihood of over-promotion is affected by changes in the convexity of the worker's outside options but not by changes in the relative marginal product of workers at junior and senior positions. We also show that over-promotion is more likely with a better pool of workers in the sense of the distribution of workers' abilities increasing in the likelihood ratio order.

One interpretation of our environment is that the firm commits to the optimal HR policy by, for example, adopting a rule-based HR department to make its retention and promotion decisions (Milgrom and Roberts, 1988) or by establishing reputation (Waldman, 2003). Indeed, because commitment is valuable, the firm has an incentive to find ways to commit. Alternatively, we show that it is enough for the firm to be able to commit to a Blackwell experiment about workers' abilities, which can be interpreted as a test, an exam, or an objective review of the workers' abilities.[2] Moreover, when the optimal HR policy is deterministic, it can be implemented even if the firm does not have commitment power, as long as the outside firms can observe the numbers of retained and promoted workers.

Our results can be thought of as having a normative implication for firms when they design career paths for workers while taking into account the fact workers may leave firms after promotions. The results can also be viewed as a benchmark to assess alternative HR policies. Our observation that firms can over-promote rationalizes the Peter Principle (Peter and Hull, 1969)—which states that employees are promoted to their level of incompetence—as a way for the firm to lower wage costs by attenuating the positive signalling effect of promotions.[3]

We also consider job design—interpreted as the firm merging positions—as an alternative way for the firm to attenuate the signalling effect of retention (with and without promotions). We show firms cannot strictly benefit from merging positions; that is, firms prefer workers to specialise. We also discuss when the firm may benefit from creating vacuous job titles.

Methodologically, we solve the firm's problem of designing an optimal HR policy by transforming the problem into an equivalent information design game. This allows us to use the tools of information design (Dworczak and Martini, 2019; Arieli et al., 2023) to characterise the firm's optimal HR policy.

**Related literature** The labour literature has long recognised the signalling effect of promotions starting with Waldman (1984)—indeed, the term "promotion as signals" has become widely used (Gibbons and Roberts, 2013). This literature has shown, in particular, that firms have the incentive to under-promote workers in the presence of signalling effects (Waldman, 1984; Gibbons and Katz, 1991; Bernhardt, 1995; Ishida, 2004; Owan, 2004; Gibbons and Waldman, 2006; DeVaro and Waldman, 2012). However, the literature has not focused on the signalling effect of retention. Instead, beginning with Greenwald (1986), the literature has predominantly studied the adverse selection effect arising from the employer's strategic decision to only retain workers who are "good enough" following

[2]We note that the firm can implement any Blackwell experiment using contracts via an intermediary that can be thought of as a line manager of workers (Deb, Pai and Said, 2023).

[3]We follow the interpretation of the principle by Lazear (2004). Others, such as Joao Ricardo (2000); Fairburn and Malcomson (2001); Benson, Li and Shue (2019), interpret the Peter Principle as referring to "the tendency to promote workers who excel at their current jobs while downplaying or ignoring their aptitude for management." They also focus on the role of promotions to provide incentives.

attempts to poach workers by outside firms.[4]

We contribute to the labour literature in two ways. First, we distinguish retention from promotion decisions and jointly study the signalling effects of these decisions.[5] Second, we study the firm's incentive to over-retain or over-promote workers as a way to attenuate the positive signalling effect of retention and promotion decisions to minimise wage costs.[6] We do so by allowing the firm to commit to retention and promotion policies, which can be viewed as a strengthening of the assumption in Ghosh and Waldman (2010) that allows firms to (only) commit to letting go of workers.[7]

As noted by Baker, Jensen and Murphy (1988), retentions and promotions can also be used to provide incentives directly (e.g., in the form of tournaments). We do not focus on this aspect of retention and promotions, and analyse instead how asymmetric information shapes the firm's incentives to assign workers to the jobs for which they are best suited. DeVaro and Waldman (2012) provide empirical evidence in support of the signalling effect of promotions, and a number of papers find evidence for asymmetric learning in labour markets (e.g., Gibbons and Katz, 1991; Schonberg, 2007; Pinkston, 2009; Kahn, 2013). However, we are not aware of empirical studies that test whether firms over- or under-retain or promote.

The firm's problem we study is an example of a "signalling game with commitment" in which the sender (i.e., the firm) has the ability to commit to signalling strategies (i.e., HR policies). Like Boleslavsky and Shadmehr (2023),[8] who study such a game, we adopt a belief-based approach in solving the game.[9] However, because our game has a continuum of states, their approach cannot be applied in our environment. Instead, we exploit the linear structure of labour technology and appeal to results from information design (Dworczak and Martini, 2019; Arieli et al., 2023).

The case in which the firm cannot commit to HR policies but can disclose the distri-

[4]Gibbons and Katz (1991) is an exception—their model captures both the signalling and adverse selection effects of retention. However, they do not consider promotions.

[5]In doing so, we abstract from the possible adverse selection effects that can arise with respect to both retention and promotions.

[6]The possibility that over-promotions can benefit workers is briefly mentioned in the conclusion of Bernhardt (1995).

[7]Following Waldman (1984), we do not introduce adverse selection (Greenwald, 1986; Gibbons and Katz, 1991) or allow the workers to possess private information (Costa, 1988).

[8]We note that a later version of Boleslavsky and Shadmehr (2023) also makes the same observation as we do that a signalling game with commitment can be studied as an information design game in which the sender chooses action sequentially rationally after learning about the state from a public experiment.

[9]See also Koessler, Laclau and Tomala (2024) who study a signalling game without commitment using the belief-based approach.

bution of assigned positions is equivalent to adding a credibility constraint (Lin and Liu, 2024) to the firm's problem. However, because we study a game and the firm's payoff depends on the state (i.e., the workers' abilities), the results from Lin and Liu (2024) do not apply immediately to our environment. We are nevertheless able to solve the firm's constrained problem directly.

We emphasise that while we focus on an employer's strategic incentive to retain and/or promote its workers, our approach to solving a signalling game with commitment as well as the relaxation of the commitment assumption are both generalisable to environments beyond the one we study.

# 2 Set up

Consider a firm that employs a unit mass of junior workers. Each worker has a *type*, which we interpret as the worker's ability, and the distribution of workers' types is given by a cumulative distribution function $F$ with full support on $\Theta := [0,1]$ with an associated probability density function $f$. The firm can *let go*, *retain*, or *promote* its workers. Following Costa (1988), we assume that each worker's output is affine and strictly increasing in the worker's type. Letting $j = 1$ and 2 denote the junior and senior positions, respectively, we denote the output of a type-$\theta$ worker in position $j \in \{1,2\}$ by $v_j : \Theta \to \mathbb{R}$ defined as

$$v_j(\theta) := s_j + k_j\theta,$$

where $s_j \in \mathbb{R}$ and $k_j \in \mathbb{R}_{++}$. The intercept $s_j$ can be interpreted as firm-and-position-specific human capital, and the slope $k_j$ as the marginal product of the worker's ability in position $j$. We assume that the more senior position has a higher marginal product but not a uniformly higher total product; i.e., $k_2 > k_1$ and $v_2(\theta) \geq v_1(\theta)$ if and only if $\theta \geq \overline{\theta}_1$ for some $\overline{\theta}_1 \in (0,1)$.[10] We say that the firm assigns position $j = 0$ to mean that the firm lets go of the worker, in which case there is zero output, $v_0(\theta) := 0$.

In addition to working for the firm, workers have access to a competitive labour market consisting of a compact set $I$ of *outside firms* with distinct production technologies that are also affine and strictly increasing in $\theta$. We denote the technology of an outside firm $i \in I$ as $c^i : \Theta \to \mathbb{R}$ and let $c(\theta) := \sup\{c^i(\theta) : i \in I\}$ to be type-$\theta$ worker's highest output across

[10] More precisely, we assume that $\overline{\theta}_1 \equiv \frac{s_1 - s_2}{k_2 - k_1} \in (0,1)$.

all outside firms.[11] As a supremum of strictly increasing affine functions, $c(\cdot)$ is strictly increasing and convex. To ease the interpretation and presentation of results, we assume that $c(\cdot)$ is, in fact, strictly convex and differentiable. We refer to the outside firm that maximises the output of a type-$\theta$ worker as a $\theta$-*outside firm*. Note that $\theta$-outside firm's production technology, which we denote as $c_\theta(\cdot)$, is given by the tangent to $c(\cdot)$ at $\theta$ and its worker's marginal product is given by $c'(\theta)$. Thus, the convexity of $c(\cdot)$ reflects the fact that $\theta$-outside firm's output is more sensitive to workers' abilities when $\theta$ is higher—we will demonstrate that this property is the key driver of our results. We further assume $c(\cdot)$ to be such that, for some $\overline{\theta}_0 \in (0, \overline{\theta}_1)$,

$$\max\{c(\theta), v_1(\theta), v_2(\theta)\} = \begin{cases} c(\theta) & \text{if } \theta \in [0, \overline{\theta}_0), \\ v_1(\theta) & \text{if } \theta \in [\overline{\theta}_0, \overline{\theta}_1], \\ v_2(\theta) & \text{if } \theta \in [\overline{\theta}_1, 1]. \end{cases} \tag{1}$$

For an allocation of workers to firms to be efficient, each type-$\theta$ worker must be employed at a firm (and in a position) that maximises the worker's output. Thus, equation (1) means that it is efficient for the firm to let go of workers whose type is below $\overline{\theta}_0$, retain workers whose type lies in $[\overline{\theta}_0, \overline{\theta}_1]$ at the junior level, and to retain and promote workers whose type is above $\overline{\theta}_1$.[12]

A worker's payoff is simply the wage. The firm's payoff from paying a wage $w$ to a type-$\theta$ worker in position $j$ is $v_j(\theta) - w$. As we explain in more detail below, we follow Waldman (1984) and think of the firm as paying its workers just enough to ensure that they will not be poached by other firms.[13] Thus, $c(\cdot)$ can be interpreted as the wage cost for the firm of a type-$\theta$ worker when the labour market is able to observe $\theta$. We define $V_j := v_j - c$ as the firm's profit in such a case, and assume that $\max V_2 > \max V_1$ to capture the idea that seniors can potentially be more profitable than junior workers.

We give a typical example of the firm's production technology that satisfies our assumptions in Figure 1. The top figure in Figure 1 shows a typical output of workers in junior and

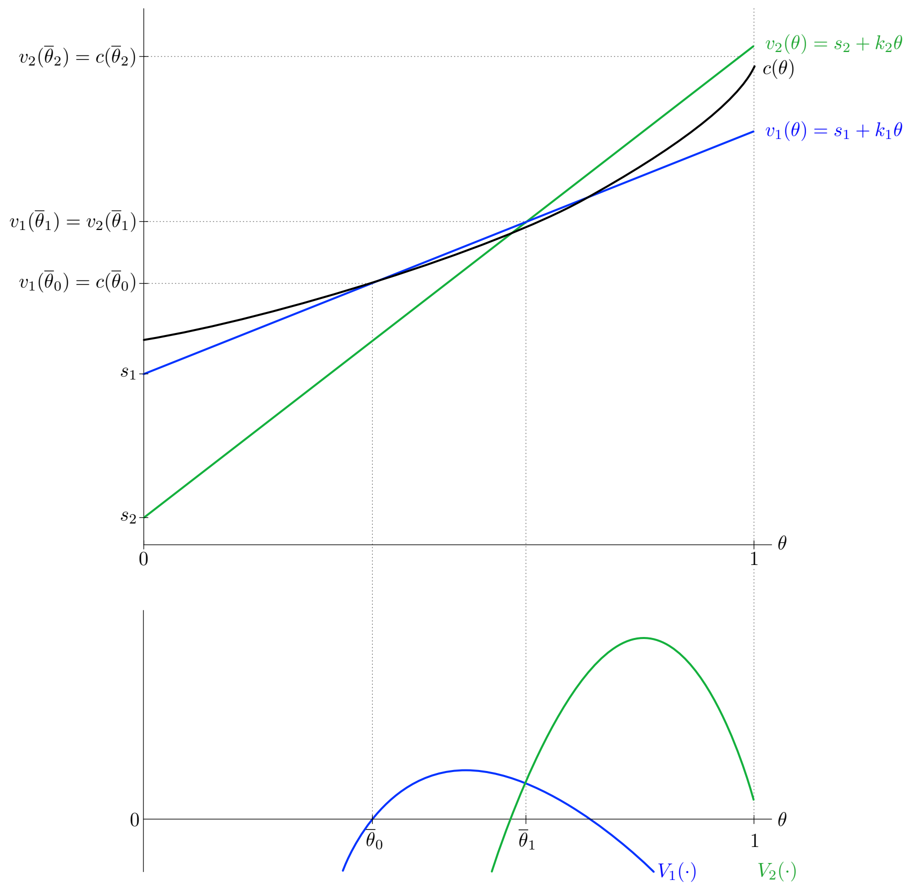
[11] We assume, without loss, that each outside firm has one position for its workers.

[12] The assumption that it is efficient for the firm to retain workers whose types are above $\overline{\theta}_0$ implies that the incumbent firm values the highest type worker more than the labour market; i.e., $v_2(1) \geq c(1)$. In section 5.4, we discuss the case in which it may be efficient for the firm to let go of workers above some threshold $\overline{\theta}_2 \in (\overline{\theta}_1, 1]$.

[13] To focus on the signalling effect of retention and promotion decisions, we do not allow the firm to strategically let go of its workers after receiving a poaching offer as in Greenwald (1986) or allow other firms to offer a menu of wages as in Costa (1988).

senior positions, $v_1(\cdot)$ and $v_2(\cdot)$, respectively, as well as the maximum output of workers across all outside firms, $c(\cdot)$. The bottom figure in Figure 1 shows the firm's typical profit from retaining workers in junior and senior positions when the market observes workers' types, $V_1$ and $V_2$, respectively.

Figure 1: Firm's production technology.




## 2.1 The firm's problem

We study the firm's problem of designing an *HR policy*, $\varphi : \Theta \to \Delta(\{0,1,2\})$, that, possibly stochastically, lets go, retains at the current position, or promotes each type of worker. We account for the interaction of the firm's HR policy and the labour market by assuming that the market can distinguish between workers based only on the positions that they are assigned. Thus, to retain a worker in position $j \in \{1,2\}$, it suffices for the incumbent employer to just outbid the market and pay a wage equal to the market's expectation of the types of worker in position $j$.[14] We denote the expected type of a worker in position $j \sim \varphi$ as $\mathbb{E}_{\varphi}[\theta|j] := \frac{\int_{\Theta} \theta \varphi(j|\theta) \mathrm{d}F(\theta)}{\int_{\Theta} \varphi(j|\theta) \mathrm{d}F(\theta)}$ and let $\Psi := (\Delta(\{0,1,2\}))^{\Theta}$ denote the set of HR policies. We first analyse the case in which the firm is able to commit to any HR policy. Thus, the firm's problem is given by:

$$\max_{\varphi \in \Psi} \int_{\Theta} \sum_{j=1}^{2} \left[ v_j(\theta) - c\left( \mathbb{E}_{\varphi}\left[ \tilde{\theta} | j \right] \right) \right] \varphi(j|\theta) \, \mathrm{d}F(\theta). \tag{2}$$

We also study two further variations of the firm's problem: (i) the case in which the firm cannot commit to HR policies, but the resulting marginal distribution of positions is observable by the labour market, and (ii) the case in which the firm makes a retention and promotion decisions in a sequentially rational manner after learning the type of each worker.

## 2.2 Simplifying the firm's problem

An HR policy $\varphi$ is a signalling strategy—a mapping from the firm's private information, $\theta \in \Theta$, to a costly action, $j \in \{0,1,2\}$, of assigning a position to each worker type. However, it is not clear how one goes about directly solving (2) that does not involve sequential rationality constraint on the part of the firm (i.e., the sender). We therefore solve the firm's problem indirectly by studying a related, payoff-equivalent information design problem in which the firm designs information about workers' types instead of an HR policy. More precisely, in this game, the firm first publicly chooses a Blackwell experiment about the workers' productivity, the result of which is public. We think of such an experiment as a test of whether a worker has reached a particular milestone or passed an exam.

Letting $M$ denote the set of payoff-irrelevant possible realisations of experiments with $|M| \geq 3$, we let $\Pi := (\Delta M)^{\Theta}$ denote the set of possible Blackwell experiments that the firm can choose. We assume that the firm has no private information about workers' types. Consequently, the firm's choice of a Blackwell experiment, $\pi \in \Pi$, together with the prior distribution $F$, completely determine how the firm and the labour market symmetrically learn about the types of workers. Having observed the outcome of the experiment, $m \sim \pi(\cdot|\theta)$, the firm assigns positions, $j \in \{0,1,2\}$, to each of its workers in a sequentially rational manner. The labour market, having observed $(\pi, m, j)$ for each worker, makes poaching offers in a sequentially rational manner.

[14] We implicitly assume a Bertrand-type competition for the workers so that $c(\cdot)$, strictly speaking, represents the second highest output of workers among firms in the labour market (see, for example, Blume, 2003). We also assume that workers break ties in favour of the incumbent employer.

In setting up the firm's problem in the related game, we take a belief-based approach. Because production technologies are affine in $\theta$, both the firm's and the market's decisions depend only on the posterior mean after observing $(\pi, m)$ or $(\pi, m, j)$. Moreover, because the firm and the labour market are symmetrically informed about workers' types, the assignment of positions provides no more information to the market than that provided through the result of the experiment. Thus, if the realised posterior mean is $E \in \Theta$, then the firm knows that it must pay $c(E)$ to keep its workers no matter the position it assigns so that the firm chooses the position $j \in \{1,2\}$ to maximise the average payoff of workers in position $j$, $v_j(E)$. However, if the average payoff from optimally not letting go is negative, the firm should simply let go of the workers. Thus, we let $V : \Theta \to \mathbb{R}$ be the firm's value function, i.e., its payoff from inducing posterior mean $E$, and define it as

$$V(E) := \left[ \max_{j \in \{1,2\}} v_j(E) - c(E) \right]_+ ,$$

where $[\cdot]_+ \equiv \max\{\cdot, 0\}$. Finally, because of the well-known fact (Gentzkow and Kamenica, 2016; Kolotilin, 2018) that the firm can induce any distribution of posterior means that is a mean-preserving contraction of $F$ (and the converse is also true), we can write the firm's information design problem as

$$\widehat{V}^* := \max_{G \in \mathrm{MPC}(F)} \int_\Theta V(E) \, \mathrm{d}G(E), \tag{3}$$

where $\mathrm{MPC}(F)$ denotes the set of all mean-preserving contractions of $F$.

We first establish an equivalence between the firm's original problem of choosing an optimal HR policy and the firm's problem of choosing an optimal distribution of posterior means in the related information design game.

**Lemma 1.** *The value of the firm's problem,* (2)*, in the original game equals the value of the firm's problem,* (3)*, in the related game; i.e.,* $V^* = \widehat{V}^*$*. Moreover, if* $G^*$ *solves* (3)*, then there exists an HR policy* $\varphi^* \in \Psi$ *that solves the firm's original problem,* (2)*, and induces* $G^*$ *as the distribution of posterior means.*

We further simplify the firm's information design problem by showing that it is without loss to only consider distributions of posterior means from a parametric family of CDFs. Towards this goal, let us call a tuple $(\theta_0, \theta_1, E_2) \in \Theta^3$ as *feasible* if $\theta_0 \in [0, \overline{\theta}_0]$, $\theta_1 \in [\theta_0, 1]$, and $\mathbb{E}[\theta | \theta > \theta_0] \leq E_2 \leq \mathbb{E}[\theta | \theta \geq \theta_1]$. Given any feasible tuple $(\theta_0, \theta_1, E_2)$, we define a

CDF $G_{\theta_0,\theta_1,E_2}:\mathbb{R}\to[0,1]$, to be such that (i) $G\in \mathrm{MPC}(F)$; (ii) $G_{\theta_0,\theta_1,E_2}$ and $F$ coincide on $[0,\overline{\theta}_0]$, and (iii) the mass on $[\overline{\theta}_0,1]$ from $F$ is put on $\mathbb{E}[\theta|\theta_0<\theta<\theta_1]$ and on $E_2$. We say that a distribution $G$ is *feasible* if there exists a feasible tuple $(\theta_0,\theta_1,E_2)$ such that $G=G_{\theta_0,\theta_1,E_2}$ and, in that case, we also say that the tuple $(\theta_0,\theta_1,E_2)$ is *associated* with $G$.

**Lemma 2.** *There exists a feasible $G^*$ that solves the information design problem,* (3)*; i.e.,*

$$\widehat{V}^* = \max_{(\theta_0,\theta_1,E_2)} \int_{\Theta} V(E)\,\mathrm{d}G_{\theta_0,\theta_1,E_2}(\theta) \;\; \textit{s.t.} \;\; (\theta_0,\theta_1,E_2) \;\textit{is feasible}. \tag{4}$$

We prove Lemma 2 using results from Arieli et al. (2023). In particular, they show that information design problems, such as the firm's problem, (3), always have a solution that belongs in a family of CDFs that contains the set of feasible CDFs. We then use the structure of $V$ to show that the firm's problem is always solved by a feasible CDF. Importantly, because (4) is a maximisation problem with respect to three real-valued parameters, Lemma 2 tells us that we can solve the firm's information design problem, (3), and characterise optimal distributions of posterior beliefs using the first-order approach.[15]

Once we derive an optimal distribution of posterior means that solve (4), Lemma 1 tells us that we can obtain an HR policy that solves the original problem. However, in general, many information structures can induce a particular distribution of posterior means.[16] This also implies that there is a multiplicity of optimal HR policies.

Among optimal HR policies, we sometimes focus our attention on HR policies that have the property that higher-type workers are more likely to be assigned to a higher position. We refer to such HR policies as being *monotone*.[17] Monotone HR policies have the desirable property that, in an enriched environment in which worker types are endogen-

[15]We note that the objective function need not be concave nor differentiable everywhere so that the first-order conditions are not sufficient conditions for optima. Nevertheless, we are able to solve the problem using the first-order approach by using the special properties of our environment and results from Dworczak and Martini (2019) and Arieli et al. (2023).

[16]See Example 3 in Arieli et al. (2023).

[17]To formally define this property, note that an HR policy $\varphi$, when viewed as an information structure (i.e., with $M=\{0,1,2\}$ and $\varphi=\pi$), induces three possible posterior means, $\{E_j\}_{j=0}^2$, where $E_j$ is the posterior mean after observing $j$. Let $\widehat{\varphi}(\theta)\in\Delta(\{E_j\}_{j=0}^2)$ denote the distribution of posterior means such that

$$\widehat{\varphi}(E|\theta) = \begin{cases} \varphi(j|\theta) & \text{if } \exists j\in\{0,1,2\},\ E=E_j, \\ 0 & \text{otherwise.} \end{cases}$$

With this notation, an HR policy $\varphi$ is *monotone* if, for any $\theta'\geq\theta$, $\widehat{\varphi}(\theta')$ first-order stochastically dominates $\widehat{\varphi}(\theta)$. Observe that if $\varphi$ is deterministic, then $\mathrm{supp}(\widehat{\varphi}(\cdot))$ is a singleton, and thus, the monotonicity condition requires that the posterior means are higher for higher types of workers.

ous, higher-type workers would not have the incentive to pretend to be lower-type workers (for example, by shirking).[18] As we demonstrate later, it will also turn out to be related to the types of HR policies that can be implemented even when firms do not have “full” commitment power.

Importantly, any feasible distribution of posterior means $G$ can be induced by a monotone HR policy (Arieli et al., 2023). For example, given a feasible tuple $(\theta_0, \theta_1, E_2)$, the following HR policy $\varphi$ induces $G_{(\theta_0,\theta_1,E_2)}$: $\varphi$ lets go of the workers whose types lie in $[0, \theta_0]$ with probability one, promotes workers whose types lie in $[\theta_1, 1]$ with probability one, and randomise between promoting and not promoting (but retaining) workers whose types lie in $[\theta_0, \theta_1]$. Thus, we interpret $\theta_0$ as a *retention threshold* and $\theta_1$ as a *always-promote threshold*, where we make the distinction to account for the possible randomisation. However, if $E_2 = \mathbb{E}[\theta|\theta \geq \theta_1]$, then $\theta_0$ can be interpreted as the promotion threshold (i.e., workers are promoted if and only if their type is above $\theta_1$). Moreover, whenever $E_2 \in \{\mathbb{E}[\theta|\theta > \theta_0], \mathbb{E}[\theta|\theta \geq \theta_1]\}$, the associated feasible $G$ can be induced by a monotone policy that is deterministic.[19]

# 3 Inefficiencies of optimal HR policies

Our main result concerns the nature of inefficiencies associated with the firm’s optimal retention and promotion decisions. To be precise, we say that an HR policy $\varphi$ exhibits *over-retention* (resp. *under-retention*) if $\varphi$ retains workers whose types are below (resp. above) the efficient retention cutoff, $\overline{\theta}_0$, with positive probability.[20] Under this definition, an HR policy exhibits efficient retention if and only if it does not exhibit over-retention or under-retention. We also emphasise that an HR policy can exhibit both over- and under-retention if it over-retains some types of workers and under-retains other types of workers. Analogously, we say that $\varphi$ exhibits *over-promotion* (resp. *under-promotion*) if $\varphi$ promotes workers whose types are below (resp. above) the efficient promotion cutoff, $\overline{\theta}_1$, with positive probability. We study how the firm trades off its desire not to retain or promote workers whose productivities are lower than the wage the firm must pay to ensure they do not leave,

[18] Goldstein and Leitner (2018) and Mensch (2021) provide discussions on why and when the monotonicity property might be desirable.

[19] If $E_2 = \mathbb{E}[\theta|\theta > \theta_0]$, workers whose types are above $\theta_0$ are all retained as juniors or are all promoted to be seniors.

[20] We can equivalently interpret there to be a continuum of workers of each type such that the distribution of types corresponds to $F$. Under such an interpretation, $\varphi$ can be thought of as a deterministic mapping from $(i, \theta) \in [0,1] \times \Theta$ to a position in $j \in \{0,1,2\}$.

against its desire to over-retain or over-promote to attenuate the positive signalling effect of retention and promotions. We first establish that over-retention is optimal for the firm except in a knife-edge case where the optimal policy retains workers efficiently.

**Theorem 1.** *An optimal HR policy exists with efficient retention if and only if the efficient retention threshold* $\overline{\theta}_0$ *satisfies*

$$V_2'(\mathbb{E}[\theta|\theta \geq \overline{\theta}_0]) \geq V_1'(\overline{\theta}_0) \tag{5}$$

*and*

$$V_2(\mathbb{E}[\theta|\theta \geq \overline{\theta}_0]) = V_2'(\mathbb{E}[\theta|\theta \geq \overline{\theta}_0])(\mathbb{E}[\theta|\theta \geq \overline{\theta}_0] - \overline{\theta}_0). \tag{6}$$

*Moreover, under-retention is never optimal. Therefore, if conditions* (5) *or* (6) *do not hold, then any optimal HR policy exhibits over-retention.*

Theorem 1 says that an optimal HR policy—whether monotone or not—exhibits over-retention. We emphasise that conditions (5) and (6) hold only in knife-edge cases.[21] Thus, we interpret Theorem 1 as saying that we should generally expect firms to over-retain workers. The underlying intuition for over-retention is most clearly demonstrated by considering the firm's expected payoff under a feasible distribution $G$ that can be implemented with a deterministic monotone HR policy. Recall that deterministic monotone HR policies are characterised by a retention threshold, $\theta_0$, and a promotion threshold, $\theta_1$, so that the firm's expected payoff is given by:

$$\begin{aligned}\widehat{V}\left(\theta_0, \theta_1\right) = &\int_{\theta_0}^{\theta_1} \left(v_1\left(\theta\right) - c\left(\mathbb{E}\left[\tilde{\theta}|\theta_0 < \tilde{\theta} < \theta_1\right]\right)\right) f\left(\theta\right) \mathrm{d}\theta \\ &+ \int_{\theta_1}^{1} \left(v_2\left(\theta\right) - c\left(\mathbb{E}\left[\tilde{\theta}|\tilde{\theta} \geq \theta_1\right]\right)\right) f\left(\theta\right) \mathrm{d}\theta,\end{aligned} \tag{7}$$

where the first and second lines are the firm's profits from workers in position $j = 1$ and 2, respectively.

Let us consider whether the firm would find it optimal to choose the efficient retention threshold $\overline{\theta}_0$ by analysing the derivative of $\widehat{V}$ with respect to $\theta_0$ evaluated at the efficient retention cutoff $\overline{\theta}_0$ while assuming that $E_1 \equiv \mathbb{E}[\theta|\overline{\theta}_0 < \theta < \theta_1] \in (\overline{\theta}_0, \overline{\theta}_1)$ (which can be

[21] In words, the conditions require the tangent to $V$ at $\mathbb{E}[\theta|\theta \geq \overline{\theta}_0]$ to have a slope greater than $V_1'(\overline{\theta}_0)$ and to cut the $x$-axis at exactly $\overline{\theta}_0$.

shown to hold in the relevant cases). Leibniz rule gives that

$$\left.\frac{\mathrm{d}\widehat{V}(\cdot)}{\mathrm{d}\theta_0}\right|_{\theta=\overline{\theta}_0} = -f\left(\overline{\theta}_0\right)\left(v_1\left(\overline{\theta}_0\right) - c\left(E_1\right)\right) - \int_{\overline{\theta}_0}^{\theta_1} c'\left(E_1\right)\frac{\mathrm{d}E_1}{\mathrm{d}\theta_0} f(\theta)\mathrm{d}\theta. \tag{8}$$

Because the firm must pay $c(E_1)$ to retain a type-$\overline{\theta}_0$ worker that produces $v_1(\overline{\theta}_0)$ in output, the first term in (8) is the direct effect on the firm's profit from letting go of a type-$\overline{\theta}_0$ worker. The second term reflects the signalling effect of a change in the retention threshold: $c'(E_1)\frac{\mathrm{d}E_1}{\mathrm{d}\theta_0}$ reflects the change in the wage rate the firm must pay to retain each worker in position $j=1$, and the integral gives the change in the wage cost associated with all workers assigned to position $j=1$.

We note that the first term in (8) is strictly positive because the labour market's willingness to pay for a worker is increasing in the worker's type (i.e., $c$ is strictly increasing) and a type-$\overline{\theta}_0$ worker, by definition, is equally valued by the firm and the labour market (i.e., $c(\overline{\theta}_0) = v_1(\overline{\theta}_0)$). Thus, letting go of a type-$\overline{\theta}_0$ worker allows the firm to avoid the loss associated with retaining such a worker. However, doing so also increases the wage cost for all other workers in position $j=1$ due to the signalling effect—an increase in the retention threshold leads to an increase in the posterior mean of workers' types in position $j=1$ ($\frac{\mathrm{d}E_1}{\mathrm{d}\theta_0} > 0$), which, in turn, increases the labour market's willingness to pay for workers in position $j=1$ ($c'(E_1) > 0$). Thus, *a priori* the sign of the derivative seems ambiguous.

However, it turns out that the convexity of $c$ (which arises from the fact that $\theta$-outside firm's technology is more sensitive to workers' abilities when $\theta$ is higher) implies that the signalling effect dominates the direct effect on firm profits. To see this, we rewrite the derivative (8) as follows:

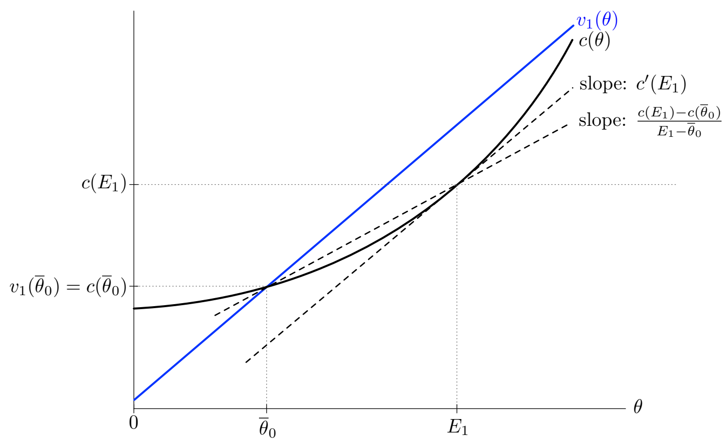
$$\left.\frac{\mathrm{d}\widehat{V}(\cdot)}{\mathrm{d}\theta_0}\right|_{\theta=\overline{\theta}_0} = f\left(\overline{\theta}_0\right)\left(E_1 - \overline{\theta}_0\right)\left(\frac{c\left(E_1\right) - c\left(\overline{\theta}_0\right)}{E_1 - \overline{\theta}_0} - c'\left(E_1\right)\right) < 0. \tag{9}$$

As can be seen from Figure 2, the derivative is strictly negative because the strict convexity of $c'$ implies the difference inside the last parentheses is strictly negative. In words, the signalling effect dominates the direct effect when $c(\cdot)$ is strictly convex because the wage cost for the firm is then "more" sensitive to changes in the posterior mean $E_1$.[22] We there-

[22] Observe that the term $\frac{c(E_1)-c(\overline{\theta}_0)}{E_1-\overline{\theta}_0}$ is the marginal product of a hypothetical outside firm with technology, $\tilde{c}(\cdot)$, such that (i) type-$\overline{\theta}_0$ worker's output is the same as that of the firm, i.e., $\tilde{c}(\overline{\theta}_0) = v_1(\overline{\theta}_0)$, and (ii) the average productivity of workers in position $j=1$, $\tilde{c}(E_1)$, equals $c(E_1)$. For the signalling effect to domin-

fore conclude that choosing the efficient retention threshold is suboptimal for the firm and the firm would, in this case, strictly prefer to over-retain workers by setting the retention threshold to be strictly below the efficient one $\overline{\theta}_0$.

Figure 2: Over-retention.




In proving Theorem 1, we also show that there is over-retention even when the optimal $G$ is associated with $(\theta_0, \theta_1, E_2)$ for which $E_2 < \mathbb{E}[\theta|\theta \geq \theta_1]$. In this case, the effect of changing the retention threshold is more complex, because changing the retention threshold also affects the mass of workers in position $j = 2$.[23] Nevertheless, the same forces described above lead the firm to over-retain workers.

We now ask: Does the firm's optimal promotion decision also always lead to over-promotion? Even if we focus on optimal monotone policies, it turns out that the nature of the inefficiency associated with promotion decisions is more nuanced.

**Theorem 2.** *An optimal deterministic monotone HR policy exhibits either over-promotion or under-promotion but not both. If no deterministic monotone HR policy is optimal, then an optimal monotone HR policy could exhibit only over-promotion, only under-promotion, or both.*

ate, it must be the case that $E_1$-outside firm's technology exhibits greater marginal product than that of the hypothetical outside firm.

[23] To see why, in general, the firm's expected payoff from a feasible $G$ associated with $(\theta_0, \theta_1, E_2)$ is

$$(1 - F(\theta_0))\left[V(\mathbb{E}[\theta|\theta_0 < \theta < \theta_1])\,\alpha(\theta_0, \theta_1, E_2) + V(\mathbb{E}[\theta|\theta \geq \theta_1])(1 - \alpha(\theta_0, \theta_1, E_2))\right],$$

where $\alpha(\theta_0, \theta_1, E_2)$ represents the proportion of retained workers who are assigned to position $j = 1$. Hence, changing the retention threshold $\theta_0$ now also affects the mass of workers who are retained in position $j = 2$.

Once again, the intuition for why the results differ between retention and promotion decisions is most clear when considering feasible $G$ that can be implemented with a deterministic monotone HR policy. Thus, consider again the firm's expected payoff from such a $G$ given by equation (7). Observe that a change in the always-promote threshold $\theta_1$ affects workers in positions $j = 1$ *and* 2. In contrast, changing the retention threshold only affects workers in position $j = 1$ because the firm's average profit-per-worker from letting go of workers is unaffected by its retention decisions—it always equals zero.

Let us consider the derivative of $\widehat{V}$ with respect to $\theta_1$ evaluated at the efficient promotion threshold $\overline{\theta}_1$ assuming that $E_1 \equiv \mathbb{E}[\theta|\theta_0 < \theta < \overline{\theta}_1] \in (\overline{\theta}_0, \overline{\theta}_1)$ and $E_2 \equiv \mathbb{E}[\theta|\theta \geq \overline{\theta}_1] \in (\overline{\theta}_1, 1)$ (which can be shown to hold in the relevant cases):

$$\left.\frac{\mathrm{d}\widehat{V}(\cdot)}{\mathrm{d}\theta_1}\right|_{\theta=\overline{\theta}_1} = f\left(\overline{\theta}_1\right)\left(\overline{\theta}_1 - E_1\right)\left(\frac{v_1\left(\overline{\theta}_1\right) - c\left(E_1\right)}{\overline{\theta}_1 - E_1} - c'\left(E_1\right)\right) + f\left(\overline{\theta}_1\right)\left(E_2 - \overline{\theta}_1\right)\left(\frac{c\left(E_2\right) - v_2\left(\overline{\theta}_1\right)}{E_2 - \overline{\theta}_1} - c'\left(E_2\right)\right), \tag{10}$$

where the first line represents the direct and signalling effects of a change in the promotion threshold $\overline{\theta}_1$ on workers in position $j = 1$, and the second line represents the direct and signalling effects on workers in position $j = 2$. By exactly the same argument as in the case of retention, the second line is negative. In contrast, the first term is strictly positive. To see this, observe that $\frac{v_1(\overline{\theta}_1) - c(E_1)}{\overline{\theta}_1 - E_1} > \frac{c(\overline{\theta}_1) - c(E_1)}{\overline{\theta}_1 - E_1}$ because it is efficient for the firm to assign position $j = 1$ to workers whose types lie in $[\overline{\theta}_0, \overline{\theta}_1]$. Convexity of $c$, in turn, implies that the first line is strictly positive.[24] Thus, whether the firm prefers to over- or under-promote is ambiguous.

When no deterministic monotone HR policy is optimal, the firm must randomise at least some assignment of types of workers to positions. However, because many monotone HR policies can induce the same distribution of posterior means, it is difficult to characterise when there will be over- and/or under-promotion. In contrast, when a deterministic monotone HR policy is optimal, it is unique. In the appendix, we provide the exact (cumbersome-to-define) conditions under which there is over- or under-promotion in the case where a deterministic monotone HR policy is optimal. We show that there are two conditions under which over-promotion can arise. These two conditions roughly say the

[24]We note that convexity of $c$ is not necessary to ensure that the first term is strictly positive. Indeed, the first term would be strictly positive even if $c$ was affine on $[E_1, \overline{\theta}_1]$.

following.

**Condition A** There is a sufficient mass of workers with very high abilities (i.e., $\theta$ close to one) so that promoting according to the efficient threshold results in wage costs that are "too high" for the firm; or

**Condition B** Condition A does not hold, but there is a sufficient mass of workers with abilities close to but below the efficient promotion threshold $\overline{\theta}_1$, so that the effect of lowering the threshold on the profits from workers in position $j = 1$ is smaller.

**Comparative statics** To avoid ambiguity, we focus on the case in which an optimal monotone HR policy is deterministic so that the optimal monotone HR policy either exhibits over-promotion or under-promotion but not both. We also focus on changes that do not affect the levels of efficient retention and promotion thresholds.

We say that tuples of primitives $(v_2, c, F)$ and $(\tilde{v}_2, \tilde{c}, \tilde{F})$ (satisfying our assumptions) are *comparable* if condition A holds under both set of primitives and they both imply common efficient retention and promotion thresholds. Given such comparable tuples of primitives, we say that over-promotion is *more likely* under $(v_2, c, F)$ than under $(\tilde{v}_2, \tilde{c}, \tilde{F})$ if the change can result in a switch from the optimal monotone HR policy exhibiting over-promotion to under-promotion but not the other way around.[25]

We first contemplate a reduction in the marginal product of the $\overline{E}_2$-outside firm, $c'(\overline{E}_2)$, which determines the strength of the signalling effect of changing the promotion threshold from the efficient one. To isolate this signalling effect, we ensure $c(\overline{E}_2)$ and $c(1)$ remain unchanged so that the change does not directly alter the wage cost for the firm from promoting efficiently, and we also leave $c$ unchanged on $[0, \overline{\theta}_1]$ to avoid directly affecting the firm's choice of retention threshold.

**Proposition 1.** *Suppose no deterministic monotone HR policy is optimal. Consider comparable $(v_2, c, F)$ and $(v_2, \tilde{c}, F)$ such that $c(\theta) = \tilde{c}(\theta)$ for $\theta \in [0, \overline{\theta}_1] \cup \{\overline{E}_2, 1\}$. Then, over-promotion is more likely under $(v_2, \tilde{c}, F)$ than under $(v_2, c, F)$ if $\tilde{c}'(\overline{E}_2) < c'(\overline{E}_2)$.*

[25]More concretely, we say that over-promotion is more likely under $(v_2, c, F)$ than under $(\tilde{v}_2, \tilde{c}, \tilde{F})$ if the derivative $\frac{\mathrm{d}\widehat{V}}{\mathrm{d}\theta_1}$ evaluated at $(\overline{\theta}_{-1}, \overline{\theta}_1)$ is lower under $(v, c, F)$ than under $(\tilde{v}_2, \tilde{c}, \tilde{F})$, where $\overline{\theta}_{-1}$ is such that the first-order condition with respect to $\theta_0$ holds. That is, $\frac{\mathrm{d}\widehat{V}}{\mathrm{d}\theta_0}$ evaluated at $(\overline{\theta}_{-1}, \overline{\theta}_1)$ is zero if $\overline{\theta}_{-1} > 0$ or strictly negative if $\overline{\theta}_{-1} = 0$. Whether there is over- or under-promotion also depends on whether condition A holds for reasons unrelated to the sign of $\frac{\mathrm{d}\widehat{V}}{\mathrm{d}\theta_1}$ evaluated at $(\overline{\theta}_{-1}, \overline{\theta}_1)$. We therefore only consider changes in the primitives that do not affect whether condition A holds.

In words, the above proposition sys over-promotion is more likely with a lower marginal product of $\overline{E}_2$-outside firm. At first sight, this result may seem surprising because a lower marginal product of $\overline{E}_2$-outside firm implies a smaller reduction in wage cost with respect to workers in position $j=2$ due to the signalling effect of over-promotion. However, notice that the choice of the promotion threshold also affects the choice of the retention threshold, and the latter alters the firm's profits from workers in position $j=1$ by affecting the average productivity of such workers. It turns out that, when taking both effects into account, the firm is more likely to over-promote when the marginal product of $\overline{E}_2$-outside firm is lower.

To demonstrate that the incentive to over- or unde-promote is largely driven by the technologies of the outside firm, we also consider a change in $k_2$ (with a corresponding change in $s_2$ to ensure that the efficient promotion threshold remains unchanged).

**Proposition 2.** *Suppose no deterministic monotone HR policy is optimal. Consider comparable* $(v_2, c, F)$ *and* $(\tilde{v}_2, c, F)$. *Then, over-promotion is no more likely under* $(\tilde{v}_2, c, F)$ *than under* $(v_2, c, F)$.

We interpret Propositions 1 and 2 as reinforcing the insight that whether there is over-promotion or under-promotion is driven by the technologies of outside firms, rather than the technology of the incumbent employer.

We can also use our model to study how the nature of inefficiency associated with optimal HR policies changes when the workers have different characteristics that are publicly observable; e.g., workers may have different education levels or professional qualifications. To that end, suppose there are two sets of workers and their prior distribution of types is ordered according to their likelihood ratios.[26] We show that over-promotion is more likely for the set of workers whose distribution dominates another in the likelihood-ratio order.

**Proposition 3.** *Suppose no deterministic monotone HR policy is optimal. Consider comparable* $(v_2, c, F)$ *and* $(v_2, c, \tilde{F})$. *Then, over-promotion is more likely under* $(v_2, c, \tilde{F})$ *than under* $(v_2, c, F)$ *if* $F$ *dominates* $\tilde{F}$ *in the likelihood ratio order.*

For example, to the extent that education levels are publicly observable, over-promotion may be more likely among workers who are less educated.

[26] Recall that, given two absolutely continuous CDFs with support $[0,1]$, $F$ and $H$, the CDF $H$ dominates $F$ in the likelihood ratio order if and only if $\frac{h(\cdot)}{f(\cdot)}$ is increasing.

# 4 Relaxing the firm's commitment power

As described in the introduction, there are a number of ways in which the firm can commit HR policies, e.g., by adopting a rule-based HR department to make its retention and promotion decisions (Milgrom and Roberts, 1988) or by establishing reputation (Waldman, 2003). Alternatively, the information design game we use to solve the firm's original problem itself represents a way for the firm to implement optimal HR policies. That is, by having a clear career path for workers in which progress depends on the workers passing certain milestones (or tests), the firm can implement optimal HR policies. Indeed, in many industries, firms disclose to potential workers how they can progress through the ranks.[27] Alternatively, the firm could also implement the optimal information structure using "indirect contracts" à la Deb, Pai and Said (2023).[28] In this light, the resistance from HR departments that line managers often face in trying to get their managees promoted "faster" can be viewed as an optimal response for the firm to manipulate the signalling effects of promotion. Nevertheless one may wonder: To what extent do our results depend on the firm's ability to commit?

To further demonstrate how else the firm can implement optimal HR policies, we consider the case in which the firm lacks the power to commit to HR policies (i.e., it can only implement incentive-compatible HR policies), but the labour market can observe the distribution of positions held by its workers after it has assigned them. That is, in addition to observing the firm's choice of HR policy $\varphi$ and the position of the workers $j \sim \varphi(\cdot|\theta)$, the labour market also observes the distribution of positions after the firm has made its retention and promotion decisions. Firms sometimes disclose such information voluntarily,[29] or disclose them as mandated by accounting rules or laws,[30] alternatively, they are available publicly through websites such as LinkedIn or through services that provide intelligence

[27]Career fairs at colleges and univer are ripe with employers explaining how graduates can expect to progress through the rank in the company.

[28]The Principal, the Agent, and the Receiver, in their setting, correspond to the firm, a line manager, and the labour market in our setting, where the line manager is an agent of the firm who is able to observe the workers' types. We would add to the set-up of Deb, Pai and Said (2023) that the firm is able to assign positions once it has gathered information about workers from the line manager.

[29]For example, certain types of companies and institutions, such as economic consultancies and academic departments, provide a public list of their staff in various positions.

[30]For example, in the UK, the Companies Act 2006 (section 411) specifies that firms above certain size "must also disclose the average number of persons within each category of persons so employed." Some diversity, equity and inclusion reporting requirements stipulate companies to disclose demographic information about their employees (e.g., The European Union's Directive 2014/95/EU).

about HR structures of companies.[31]

Importantly, public knowledge of the distribution of assigned positions effectively allows the market can detect (and respond to) deviations by the firm from its stated HR policy, $\varphi$, if deviations result in a distribution of positions that is consistent with $\varphi$. Then, following the arguments from Lin and Liu (2024), it follows that the firm is able to effectively commit to HR policies that result in the same marginal distribution of messages as that of the announced one.

To state the firm's problem when the firm has this partial commitment power, let $\Psi := (\Delta(\{0,1,2\}))^\Theta$ denote the set of HR policies and let $D(\varphi)$ denote the set of HR policies that have the same marginal distribution as $\varphi \in \Psi$; i.e.,

$$D(\varphi) := \left\{ \varphi' \in \Psi : \int_\Theta \varphi'(\cdot|\theta)\,\mathrm{d}F(\theta) = \int_\Theta \varphi(\cdot|\theta)\,\mathrm{d}F(\theta) \right\}.$$

The firm's problem is to choose an HR policy that maximises its ex ante expected payoff:

$$\max_{\varphi \in \Psi} \int_\Theta \sum_{j=1}^{2} \left[ v_j(\theta) - c\left(\mathbb{E}_\varphi\left[\tilde{\theta}|j\right]\right)\right] \varphi(j|\theta)\,\mathrm{d}F(\theta) \tag{11}$$

$$\text{s.t. } \varphi \in \arg\max_{\varphi' \in D(\varphi)} \int_\Theta \sum_{j=1}^{2} \left[ v_j(\theta) - c\left(\mathbb{E}_\varphi\left[\tilde{\theta}|j\right]\right)\right] \varphi'(j|\theta)\,\mathrm{d}F(\theta). \tag{12}$$

The constraint (12), which we refer to as the firm's *credibility constraint*, ensures that the firm, after announcing that it will follow an HR policy $\varphi$, would not deviate to some other HR policy that has the same marginal distribution of positions when the labour market uses $\varphi$ to update beliefs about workers. We say that an HR policy $\varphi$ is *credible* if $\varphi$ satisfies the credibility constraint.

We first show that, in solving (11), we can restrict attention to deterministic and monotone HR policies.

**Lemma 3.** *An HR policy is credible if and only if it is deterministic and monotone.*

Note that an HR policy $\varphi$ that is both deterministic and monotone is a simple threshold strategy in which every type up to some threshold $\theta_0$ is let go, every type above some threshold $\theta_1 \geq \theta_0$ is promoted, and every type in between $(\theta_0, \theta_1)$ is retained but not promoted. Thus, Lemma 7 simplifies the problem (2) to one of finding optimal retention and

[31] For example, companies such as Crayon collect "Competitive Intelligence" to "provide the most complete picture of what your competitors are up to."

promotion thresholds. Using this simplification, we prove that the nature of inefficiency remains unchanged.

**Proposition 4.** *Any HR policy $\varphi^*$ that solves* (11) *is a threshold policy with retention and promotion thresholds, $\theta_0^*$ and $\theta_1^* \geq \theta_0^*$ respectively. In addition, $\varphi^*$ exhibits over-retention, and it may exhibit over- or under-promotion but not both.*

Because an optimal HR policy might require randomisation, there can be a gap between the full and partial commitment power. However, the following result establishes that randomisation is the sole source of this gap.

**Proposition 5.** *A solution $G^*$ to the firm's information design problem* (3) *can be implemented under partial commitment if and only if there does not exist a tuple $(\theta_0^b, E_1^b, E_2^b)$ with $\theta_0^b \leq E_1^b \leq E_2^b$ such that:*

*(i)* $c(E_j^b) - v_j(\theta_0^b) = c'(E_j^b)(E_j^b - \theta_0^b)$ *for* $j = \{1,2\}$*;*

*(ii)* $k_2 - k_1 = c'(E_2^b) - c'(E_1^b)$*; and*

*(iii)* $E_1^b \leq \mathbb{E}[\theta|\theta > \theta_0^b] < E_2^b < \mathbb{E}[\theta|\theta \geq \theta_1^b]$.

The existence of a tuple $(\theta_0^b, E_1^b, E_2^b)$ that satisfies the conditions of the above proposition implies the firm's information design problem is solved by a $G$ associated with $(\theta_0^b, \theta_1^b, \mathbb{E}_2^b)$, where $\theta_1^b$ satisfies $\mathbb{E}[\theta|\theta_0^b < \theta < \theta_1^b] = E_1^b$. Moreover, when such a tuple does not exist, the firm's information design problem is solved by a $G$ associated with a deterministic monotone HR policy. Since any deterministic monotone HR policy is credible, it follows that the ability to partially commit is sufficient for the firm to attain the full-commitment payoff.

To better understand when the firm benefits from randomisation, let us consider each condition in Proposition 5 in turn. For each $j \in \{1,2\}$, condition (i) guarantees that, if $\mathbb{E}[\theta|\theta \geq \theta_0^b] = E_j^b$, the threshold strategy in which the firm assigns position $j$ if and only if workers' types are above $\theta_0^b$ maximises the firm's ex ante payoff from choosing among HR policies that assign either position 0 or $j$ to each type of worker.[32] Thus, $E_j^b$ can be interpreted as the average type of retained workers induced by such an HR policy. To understand condition (ii), recall that $c'(E)$ represents the marginal product of the $E$-outside

[32] To see this, recall the derivative of firm's payoff with respect to the retention threshold $\theta_0$ from a feasible $G$ associated with $E_2 = \mathbb{E}[\theta|\theta \geq \theta_1]$, (8). Observe that letting $\theta_1 = 1$ and setting the derivative equal to zero gives condition (i) for each $j \in \{1,2\}$.

firm. Now, for each $j \in \{1,2\}$, denote the group of workers who would be optimally assigned to position $j$ over being let go as $I_j \subseteq \Theta$. Then, condition (ii) requires that the firm and the labour market value the additional marginal product of type of workers in group $I_2$ over group $I_1$ to be the same. Note that conditions (i) and (ii) are conditions on $v_1$, $v_2$ and $c$.[33] The last condition (iii) is a restriction on the prior $F$ that ensures that the firm can induce distribution of posterior means $G^*$ using some information structure. Roughly, it requires a sufficient mass of workers who are "good enough" so that the firm can reduce the wage cost of workers in position $j = 2$ by sometimes promoting workers who are not "good enough."

Because the HR policy $\varphi^*$ that solves the firm's problem under partial commitment, (11), exhibits over-retention (i.e., $\theta_0^* < \overline{\theta}_0$), the firm's profits from some types of junior workers are strictly negative under $\varphi^*$. In fact, $v_1(\theta) - c(\mathbb{E}[\theta|\theta_0^* < \theta < \theta_1^*]) < 0$ for all $\theta \in (\theta_0^*, \frac{c(\mathbb{E}[\theta|\theta_0^* < \theta < \theta_1^*]) - s_1}{k_1}) \supseteq (\theta_0^*, \overline{\theta}_1]$. Thus, assigning $j = 1$ to workers whose types are in $(\theta_0^*, \frac{c(\mathbb{E}[\theta|\theta_0^* < \theta < \theta_1^*]) - s_1}{k_1})$ would not be sequentially rational if the firm had (privately) observed its workers' types before assigning positions. An analogous argument means that over-promoting workers would not be sequentially rational for such a firm either. Thus, some ability to commit to HR policies (or to commit to a Blackwell experiment in the information design game) is needed to ensure that the trade-off that we study is nontrivial.

# 5 Discussions

## 5.1 Optimal HR policies

We focus on the nature of inefficiency associated with the firm's optimal monotone HR policies, which do not require us to characterise the firm's optimal HR policy. Nevertheless, in the appendix, we provide a full characterisation of the firm's optimal monotone HR policies.

## 5.2 Job design

An alternative way to attenuate the positive signalling effect of promotion is to combine the jobs associated with positions 1 and 2 into one role, say to position $j = m$ (for *merged* pos-

[33] In the appendix, we show that a tuple $(\theta_0^b, E_1^b, E_2^b)$ satisfying conditions (i) and (ii) of Proposition 5 if and only if $c(\cdot)$ is sufficiently convex but not too convex. In other words, whether the firm can benefit from randomisation depends on the marginal product of workers in outside firms.

ition) as such merging of positions prevents the labour market from distinguishing workers based on whether they are retained as a junior or promoted to be a senior. However, designing jobs in this manner would presumably be inefficient for the firm,[34] and we capture this inefficiency by assuming that a type-$\theta$ worker's output in the merged position is a weighted average of $v_1$ and $v_2$; i.e.,

$$v_m(\theta) = \alpha v_1(\theta) + (1-\alpha) v_2(\theta),$$

where $\alpha \in (0,1)$.[35] However, the firm cannot benefit from merging jobs in this manner. To see this, define $V_m(E) := [v_m(E) - c(E)]_+$ as the firm's value function $G^*$ denote a solution to the firm's problem when merging the two positions. But because $V_m \leq V$, we have

$$\begin{aligned}\int_\Theta V_m(E)\,\mathrm{d}G^*(E) &\leq \int_\Theta V(E)\,\mathrm{d}G^*(E) \\ &\leq \max_{G \in \mathrm{MPC}(F)} \int_\Theta V(E)\,\mathrm{d}G(E) = \widehat{V}^*.\end{aligned}$$

Hence, the firm cannot do better by merging positions. Thus, while the firm benefits from obfuscating workers' productivity to the labour market by misallocating across positions, it prefers to specialise workers into positions.

## 5.3 Vacuous job titles

In the National Longitudinal Survey of Youth in 1990, Pergamit and Veum (1999) finds that approximately a third of those who are promoted "continued to perform basically the same duties as before." [36] While we would interpret such a finding as suggestive of the fact that firms do indeed over-promote workers, it is also possible that the finding represents evidence of a disconnect in the relationship between job titles and job responsibilities.

In the information design game that we used to solve the firm's original problem, the firm can provide more information using messages ($m \in M$) than possible via the assignment of positions. However, we show that the firm has no incentive to do so; in our model, the firm sends at most one message for each position.

[34] We note that the firm benefits from the ability to secretly assign positions $j \in \{1,2\}$ to attenuate the signalling effect without suffering a cost in terms of worker output.

[35] The results also hold when $\alpha$ depends continuously on $\theta$.

[36] Somewhat more recently, Kosteas (2011) reports a similar finding based on the 1996-2006 waves of the National Longitudinal Surveys of Youth 1979 cohort.

What drives this result is the fact that our set of assumptions implies that the firm's value function, $V$, is not strictly convex on any nondegenerate interval of $\Theta$.[37] If we were to make the alternative assumption that the set of production functions associated with outside firms is finite, then $c$ and, consequently, $V$ will be piecewise affine (but still convex). Consequently, in case the firm finds it optimal to induce posterior means in an affine region of $V$, it would be indifferent between not revealing and (partially or) fully revealing types of workers in the affine region. To the extent that there are other rationales for job titles (e.g., workers value job titles *per se*), the firm may give different job titles (i.e., send different $m$'s) to workers who have the same responsibilities (i.e., assigned the same $j$).

## 5.4 Labour market that values the workers more

In our main setup, we assumed that it is efficient for the current employer of the firm to retain the most productive workers; i.e., $v_2(1) \geq c(1)$. However, it is also possible that high-$\theta$ workers should be employed by other firms that value them more; i.e., $v_2(1) < c(1)$. In this case, define $\overline{\theta}_2 := \max\{\theta \in \Theta : v_2(\theta) = c(\theta)\}$ as the efficient threshold for letting go of workers rather than retaining them in position $j = 2$. Then, when the prior mean about workers' types is sufficiently high, it can now be optimal for the firm to adopt a deterministic HR policy characterised by a retention threshold $\theta_2^* \geq \overline{\theta}_2$ in which all workers of type $\theta < \theta_2^*$ are retained and assigned to position $j = 2$ and all workers of type $\theta \geq \theta_2^*$ are let go.

## 5.5 Beyond two jobs

Our approach to solving for the optimal HR policies extends easily to cases in which the firm can assign workers to more positions. We note that, for additional positions (interpreted as more senior positions in which workers are potentially more roductive than existing ones) to matter, they must have a lower intercept and a greater slope that existing jobs.[38] However, the tradeoffs involved between any two consecutive positions would resemble that between $j = 1$ and 2 as we describe above. Moreover, additional care must be taken to determine whether the firm would find it optimal to assign workers to each position.

[37] Recall that $V$ is a continuous function that is constant and equals zero on $[0, \overline{\theta}_0]$, positive and strictly concave on $[\overline{\theta}_0, \overline{\theta}_1]$ and $[\overline{\theta}_1, 1]$ with a convex kink at $\overline{\theta}_1$.

[38] As Costa (1988) notes, otherwise, the firm will never assign workers to the additional jobs.

## 5.6 Counteroffers

To simplify exposition and to be consistent with Waldman (1984), we do not explicitly model the wage-setting game and assume that the firm commits to offering a wage to each worker so as to ensure that the worker cannot obtain a strictly higher wage from other firms in the labour market. An alternative interpretation of this assumption is that the firm is always willing to make a counteroffer up to the average output of workers in position $j$, $v_j(\mathbb{E}_\varphi[\theta|j])$ if an external firm were to make a poaching offer of say $w$ to a worker in position $j \in \{1,2\}$. This assumption is reasonable in the information design game as the firm learns no more information about the worker than the market. However, to the extent that the firm learns the workers' types before assigning positions to them, the firm may be tempted to let go of workers who are worth less than $w$ to the firm; i.e., let go of workers in position $j$ whose types $\theta$ is such that $v_j(\theta) \leq w$. This type of counteroffer behaviour by the firm introduces an adverse selection problem for those attempting to poach workers from the firm (Greenwald, 1986; Gibbons and Katz, 1991). While a detailed analysis of the model with adverse selection is beyond the scope of this paper, we note that the firm benefits from letting go of workers strategically because doing so lowers the poaching offers the labour market is willing to make for the workers.

## 5.7 Relation with Waldman (1984) and Gibbons and Katz (1991)

Waldman (1984) studies how the firm's decision to promote or not promote a worker affects the wage it must pay in order to retain workers. In Waldman (1984), the firm assigns job $j \in \{1,2\}$ to the worker, where the firm's direct payoff from assigning job $j$ to a type-$\theta$ worker is given by $(s_1,k_1) = (x(1+s),0)$ and $(s_2,k_2) = (0,1+s)$ for some $s > 0$ and $x \in (\mathbb{E}[\theta],1]$. The labour market's value for a type-$\theta$ worker is simply $c(\theta) := \theta$. Thus, the firm's value function in Waldman (1984) is given by

$$V(E) = (1+s)\max\{E,x\} - E = \begin{cases} (1+s)x - \theta & \text{if } \theta < \overline{\theta}_1 \\ s\theta & \text{if } \theta \geq \overline{\theta}_1 \end{cases}, \tag{13}$$

where $\overline{\theta}_1 = x$. Observe that $V(\cdot) > 0$ so that the efficient outcome is for the incumbent firm to retain all workers. Because $V$ is convex, a full revelation is optimal for the firm. However, because $V$ is piecewise linear, the firm can attain the optimal payoff by simply setting the promotion threshold to be $\overline{\theta}_1$. Thus, the optimal promotion rule in Waldman

(1984) is efficient and does not exhibit under- or over-promotion.

Gibbons and Katz (1991) study a model that can capture both the signalling effect of retention à la Waldman (1984) and the adverse-selection effect of retention à la Greenwald (1986). The firm decides whether to retain or let go of the worker. The firm's payoffs from assigning each position are characterised, respectively, by $(s_0, k_0) = (0,0)$ and $(s_1, k_1) = (s, 1)$, where $0 < s < \mathbb{E}[\theta]$. The labour market's value for a type-$\theta$ worker is again $c(\theta) := \theta$. Hence, the firm's value function is simply $V(E) = s$. Observe that it is always efficient for the firm to retain workers of all types.

Observe that both Waldman (1984) and Gibbons and Katz (1991) assume a common production technology for outside firms. The analysis here, therefore, highlights that when the firm can commit to HR policies, inefficiencies in HR policies arise from the fact that heterogeneity in the production technologies of outside firms, which, in turn, leads to convexity in the workers' outside options.

# References


**Arieli, Itai, Yakov Babichenko, Rann Smorodinsky, and Takuro Yamashita.** 2023. "Optimal persuasion via bi-pooling." *Theoretical Economics*, 18(1): 15–36.

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

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

**Bernhardt, Dan.** 1995. "Strategic Promotion and Compensation." *The Review of Economic Studies*, 62(2): 315–339.

**Blume, Andreas.** 2003. "Bertrand without fudge." *Economics Letters*, 78(2): 167–168.

**Boleslavsky, Raphael, and Mehdi Shadmehr.** 2023. "Signaling with Commitment." *Working Paper*.

**Costa, Joan E. Ricart I.** 1988. "Managerial Task Assignment and Promotions." *Econometrica*, 56(2): 449–466.

**Deb, Rahul, Mallesh Pai, and Maher Said.** 2023. "Indirect Persuasion." *Working Paper*.

**DeVaro, Jed, and Michael Waldman.** 2012. "The Signaling Role of Promotion: Further Theory and Empirical Evidence." *Journal of Labor Economics*, 30(1): 91–147.

**Dworczak, Piotr, and Giorgio Martini.** 2019. "The Simple Economics of Optimal Persuasion." *Journal of Political Economy*, 127(5): 1993–2048.

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

**Gentzkow, Matthew, and Emir Kamenica.** 2016. "A Rothschild-Stiglitz Approach to Bayesian Persuasion." *American Economic Review*, 106(5): 597–601.

**Ghosh, Suman, and Michael Waldman.** 2010. "Standard promotion practices versus up-or-out contracts." *The RAND Journal of Economics*, 41(2): 301–325.

**Gibbons, Robert, and John Roberts,** ed. 2013. *The Handbook of Organizational Economics.* . 1 ed., Princeton University Press.

**Gibbons, Robert, and Lawrence F. Katz.** 1991. "Layoffs and Lemons." *Journal of Labor Economics*, 9(4): 351–380.

**Gibbons, Robert, and Michael Waldman.** 2006. "Enriching a Theory of Wage and Promotion Dynamics inside Firms." *Journal of Labor Economics*, 24(1): 59–107.

**Goldstein, Itay, and Yaron Leitner.** 2018. "Stress tests and information disclosure." *Journal of Economic Theory*, 177: 34–69.

**Greenwald, Bruce C.** 1986. "Adverse Selection in the Labour Market." *The Review of Economic Studies*, 53(3): 325–347.

**Ishida, Junichiro.** 2004. "Signaling and strategically delayed promotion." *Labour Economics*, 11(6): 687–700.

**Joao Ricardo, Fario.** 2000. "An Economic Analysis of the Peter and Dilbert Principles." *University of Technology, Sydney Working Paper*.

**Kahn, Lisa B.** 2013. "Asymmetric Information between Employers." *American Economic Journal: Applied Economics*, 5(4): 165–205.

**Koessler, Frédéric, Marie Laclau, and Tristan Tomala.** 2024. "A belief-based approach to signaling."

**Kolotilin, Anton.** 2018. "Optimal information disclosure: A linear programming approach." *Theoretical Economics*, 13(2): 607–635.

**Kosteas, Vasilios D.** 2011. "Job Satisfaction and Promotions." *Industrial Relations: A Journal of Economy and Society*, 50(1): 174–194.

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

**Lin, Xiao, and Ce Liu.** 2024. "Credible Persuasion." *Journal of Political Economy*, 132(7): 2228–2273.

**Mensch, Jeffrey.** 2021. "Monotone persuasion." *Games and Economic Behavior*, 130: 521–542.

**Milgrom, Paul, and John Roberts.** 1988. "An Economic Approach to Influence Activities in Organizations." *American Journal of Sociology*, 94: S154–S179.

**Owan, Hideo.** 2004. "Promotion, Turnover, Earnings, and Firm-Sponsored Training." *Journal of Labor Economics*, 22(4): 955–978.

**Pergamit, Michael R., and Jonathan R. Veum.** 1999. "What is a Promotion?" *Industrial and Labor Relations Review*, 52(4): 581.

**Peter, Laurence J., and Raymond Hull.** 1969. *The Peter Principle.* New York:William Morrow.

**Pinkston, Joshua C.** 2009. "A Model of Asymmetric Employer Learning with Testable Implications." *Review of Economic Studies*, 76(1): 367–394.

**Schonberg, Uta.** 2007. "Testing for Asymmetric Employer Learning." *Journal of Labor Economics*, 25: 651–691.

**Shaked, Moshe, and George J. Shanthikumar,** ed. 2007. "Stochastic Orders." , ed. Moshe Shaked and George J. Shanthikumar . 1 ed. Springer New York, NY.

**Waldman, Michael.** 1984. "Job Assignments, Signalling, and Efficiency." *The RAND Journal of Economics*, 15(2): 255.

**Waldman, Michael.** 2003. "Ex Ante versus Ex Post Optimal Promotion Rules: The Case of Internal Promotion." *Economic Inquiry*, 41(1): 27–41.

# A Appendix

## A.1 Proof of Lemma 1

**Lemma 1.** *The value of the firm's problem,* (2)*, in the original game equals the value of the firm's problem,* (3)*, in the related game; i.e.,* $V^* = \widehat{V}^*$. Moreover, if $G^*$ *solves* (3)*, then there exists an HR policy* $\varphi^* \in \Psi$ *that solves the firm's original problem,* (2)*, and induces* $G^*$ *as the distribution of posterior means.*

*Proof.* We first show that $\widehat{V}^* \leq V^*$. Suppose $\widehat{G}^* \in \text{MPC}(F)$ solves the firm's information design problem, (3). Define

$$\Theta_0 := \left\{E \in \text{supp}\left(\widehat{G}^*\right) : \max_{j\in\{1,2\}} v_j(E) \leq c(E)\right\},$$
$$\Theta_1 := \left\{E \in \text{supp}\left(\widehat{G}^*\right) \backslash \Theta_0 : v_2(E) \leq v_1(E)\right\},$$
$$\Theta_2 := \left\{E \in \text{supp}\left(\widehat{G}^*\right) \backslash \Theta_0 : v_1(E) < v_2(E)\right\},$$

and let

$$E_j^* := \frac{\int_{\Theta_j} E \mathrm{d}\widehat{G}^*(E)}{\widehat{G}^*\left(\Theta_j\right)} \; \forall j \in \{0,1,2\}.$$

By our choice of $E_j^*$'s, there exists $G^* \in \text{MPC}(F)$ so that $\text{supp}(G^*) = \{E_0^*, E_1^*, E_2^*\}$. Then, by construction, and the fact that $v$ is affine and $c$ is convex, we have

$$\widehat{V}^* \leq \int_\Theta V(E)\,\mathrm{d}G^*(E) = \sum_{j=1}^{2}\left[v_j\left(E_j^*\right) - c\left(E_j^*\right)\right] G\left(\left\{E_j^*\right\}\right).$$

Let $\pi^* \in \Pi$ denote an experiment with $\text{supp}(\pi^*) = \{m_0, m_1, m_2\}$ such that

$$\mathbb{E}_{\pi^*}\left[\theta | m_j\right] = E_j^* \; \forall j \in \{0,1,2\},$$

which exists because $G^* \in \mathrm{MPC}(F)$. Then, $\int_\Theta \pi^*(m_j|\theta)\mathrm{d}F(\theta) = G^*(\{E_j^*\})$ and so

$$
\begin{aligned}
\widehat{V}^* &\leq \sum_{j=1}^{2}\left[v_j\left(E_j^*\right)-c\left(E_j^*\right)\right]\int_\Theta \pi^*\left(m_j|\theta\right)\mathrm{d}F\left(\theta\right)\\
&= \sum_{j=1}^{2}\left[s_j+k_j\frac{\int_\Theta \theta\pi^*\left(m_j|\theta\right)\mathrm{d}F\left(\theta\right)}{\int_\Theta \pi^*\left(m_j|\theta\right)\mathrm{d}F\left(\theta\right)}-c\left(E_j^*\right)\right]\int_\Theta \pi^*\left(m_j|\theta\right)\mathrm{d}F\left(\theta\right)\\
&= \int_\Theta \sum_{j=1}^{2}\left[\underbrace{s_j+k_j\theta}_{=v_j(\theta)}-c\left(E_j^*\right)\right]\pi^*\left(m_j|\theta\right)\mathrm{d}F\left(\theta\right)\\
&\leq \max_{\varphi:\Theta\to\Delta\{0,1,2\}}\int_\Theta \sum_{j=1}^{2}\left[v_j\left(\theta\right)-c\left(\mathbb{E}_\varphi\left[\theta|j\right]\right)\right]\varphi\left(j|\theta\right)\mathrm{d}F\left(\theta\right) = V^*,
\end{aligned}
$$

where the last inequality follows because we can choose $\varphi^* : \Theta \to \Delta\{0,1,2\}$ as

$$
\varphi^*\left(j|\cdot\right) = \pi^*\left(m_j|\cdot\right)\ \forall j \in \{0,1,2\} \tag{14}
$$

so that $\mathbb{E}_{\varphi^*}[\theta|j] = E_j^*$.

We now show that $V^* \leq \widehat{V}^*$. Let $\varphi^*$ solve the firm's original problem, (2), and denote $E_j^* := \mathbb{E}_{\varphi^*}[\theta|j]$. Let $\{m_0, m_1, m_2\}$ be a set of three distinct elements of $M$ and construct a Blackwell experiment $\pi^* \in \Pi$ via (14). Let us first show that $v_j(E_j^*) - c(E_j^*) \geq 0$ for $j \in \{1,2\}$. By way of contradiction, suppose for $\widehat{j} \in \{1,2\}$, $v_{\widehat{j}}(E_{\widehat{j}}^*) - c(E_{\widehat{j}}^*) < 0$. Consider $\widehat{\varphi} : \theta \to \mathbb{R}$ such that

$$
\widehat{\varphi}\left(j'|\cdot\right) := \begin{cases}\varphi^*\left(0|\cdot\right)+\varphi^*\left(\widehat{j}|\cdot\right) & \text{if } j'=0,\\ 0 & \text{if } j'=\widehat{j},\\ \varphi^*\left(j'|\cdot\right) & \text{otherwise.}\end{cases}
$$

Letting $\widehat{j'} \in \{1,2\}$ with $\widehat{j'} \neq \widehat{j}$, we have

$$
\begin{aligned}
V\left(\widehat{\varphi}\right)-V\left(\varphi^*\right) &= \int_\theta \sum_{j=1}^{2}\left[v_j\left(\theta\right)-c\left(E_j^*\right)\right]\left[\widehat{\varphi}\left(j|\theta\right)-\varphi^*\left(j|\theta\right)\right]\mathrm{d}F\left(\theta\right)\\
&= -\int_\Theta \left[v_{\widehat{j}}(\theta)-c\left(E_{\widehat{j}}^*\right)\right]\varphi^*\left(\widehat{j}|\theta\right)\mathrm{d}F\left(\theta\right) > 0,
\end{aligned}
$$

which contradicts the optimality of $\varphi^*$. Letting $\pi^*(m_j) \coloneqq \int_\Theta \pi^*(m_j|\theta)\mathrm{d}F(\theta)$, we have

$$\begin{aligned}
V^* &= \int_\Theta \sum_{j=1}^{2} \left[v_j(\theta) - c\left(E_j^*\right)\right] \varphi^*(j|\theta)\,\mathrm{d}F(\theta) \\
&= \sum_{j=1,2} \left[v_j\left(E_j^*\right) - c\left(E_j^*\right)\right] \pi^*\left(m_j\right) \\
&\leq \sum_{j=1,2} \left[\max_{j'\in\{1,2\}} v_{j'}\left(E_j^*\right) - c\left(E_j^*\right)\right] \pi^*\left(m_j\right) \\
&\leq \max_{G\in \mathrm{MPC}(F)} \int_\Theta V(E)\,\mathrm{d}G(E) = \widehat{V}^*,
\end{aligned}$$

where the second follows from the fact that $\pi^*$ induces $G^* \in \mathrm{MPC}(F)$ such that $\mathrm{supp}(G^*) = \{E_0^*, E_1^*, E_2^*\}$. ■

## A.2 Proof of Lemma 2

We first summarise results from Dworczak and Martini (2019) and Arieli et al. (2023) that are relevant in our proofs for convenience.

The following of follows directly from theorems 1 and 2 of Dworczak and Martini (2019).

**Fact 1** (Dworczak and Martini, 2019)**.** *A CDF $G$ is a solution to the firm's information design problem,* (3)*, if there exists a function $p : \Theta \to \mathbb{R}$ such that*

$$p \geq V \text{ and } p \text{ is convex}, \tag{15}$$

$$\mathrm{supp}(G) \subseteq \{E \in \Theta : V(E) = p(E)\}, \tag{16}$$

$$\int_\Theta p(E)\,\mathrm{d}G(E) = \int_\Theta p(\theta)\,\mathrm{d}F(\theta), \tag{17}$$

$$G \in \mathrm{MPC}(F). \tag{18}$$

*Moreover, if $V$ is Lipschitz continuous, then the converse also holds; i.e., if a CDF $G$ solves* (3)*, then there exists a function $p : \Theta \to \mathbb{R}$ that satisfies* (15)–(18).

Given a pair $(G, p)$ that satisfies (15)–(18), we refer to $p$ as a *support function of $G$*.[39]

[39] Dworczak and Martini (2019) refers to $p$ instead as a price function for reasons that are not relevant for our paper.

Following Arieli et al. (2023), we call a mean-preserving contraction $G$ of $F$, $G \in \mathrm{MPC}(F)$, a *bi-pooling* distribution (with respect to $F$) if $G$ can be induced as a distribution of posterior means by an information structure that, for some collection (indexed by $\mathscr{I}$) of pairwise disjoint open intervals $\bigcup_{I\in\mathscr{I}} I \subseteq \Theta$: (i) partitions $\Theta$ into $\bigcup_{I\in\mathscr{I}} I$ and its complement; (ii) whenever $\theta \in \bigcup_{I\in\mathscr{I}} I$, the information structure reveals which interval the state is in and provides an additional binary signal; and (iii) the sender fully reveals the state whenever the state is in the complement of $\bigcup_{I\in\mathscr{I}} I$. If a nontrivial binary signal is provided in an interval $I \in \mathscr{I}$, we call the interval $I$ a *bi-pooling* interval; if only a trivial binary signal is provided, then we call $I$ a *pooling* interval. We call an interval $I \in \mathscr{I}$ in which the sender fully reveals the state a *separating* interval. If all intervals are pooling intervals, we refer to $G$ as a *pooling* distribution.

The following is a combination of Lemma 1, Theorem 1 and Proposition 3 in Arieli et al. (2023).

**Fact 2** (Arieli et al., 2023)**.**

*(i) A bi-pooling distribution $G$ solves* (3)*.*

*(ii) Suppose $V$ is upper semicontinuous and there exists a collection of $J \in \mathbb{N}$ points $\{\overline{\theta}_j\}_{j=-1}^{J}$, with $0 = \overline{\theta}_{-1} < \overline{\theta}_0 < \cdots < \overline{\theta}_J = 1$, such that $V$ is either concave or convex in the interval $(\overline{\theta}_{j-1}, \overline{\theta}_j)$ for every $j \in \{0,1,\ldots,J\}$. Let $p : \Theta \to \mathbb{R}$ denote the support function of an optimal bi-pooling distribution $G$. Then, there exists a collection of functions, $\{p_j : [\theta_{j-1}, \theta_j] \to \mathbb{R}\}_{j=0}^{J}$, where $0 = \theta_{-1} < \theta_0 < \ldots < \theta_m = 1$ for some $m \leq 2(J+1)-1$,*[40] *such that $p|_{[\theta_{j-1},\theta_j]} = p_j$ for all $j \in \{0,1,\ldots,J\}$, and each $p_j$ satisfies exactly one of the following conditions:*

*(a) $p_j = V|_{[\theta_{j-1},\theta_j]}$;*

*(b) $p_j \geq V|_{[\theta_{j-1},\theta_j]}$, and $p_j$ is affine and tangential to $V$ at $\mathbb{E}[\theta|\theta_{j-1} < \theta < \theta_j]$; or*

*(c) $p_j \geq V|_{[\theta_{j-1},\theta_j]}$, and $p_j$ affine and tangential to $V$ at two distinct points $\underline{E}_j$ and $\overline{E}_j$ such that*

$$\theta_{j-1} \leq \underline{E}_j \leq \mathbb{E}\left[\theta|\theta_{j-1} < \theta < \theta_j\right] \leq \overline{E}_j \leq \mathbb{E}\left[\theta|\theta'_{j-1} < \theta < \theta_j\right] \leq \theta_j, \quad (19)$$

*where $\theta'_{j-1}$ is defined implicitly via $\mathbb{E}[\theta|\theta_{j-1} < \theta < \theta'_{j-1}] = \underline{E}_j$.*

[40] While Proposition 3 in Arieli et al. (2023) states $m \leq 2(J+1)$ instead, the argument in their proof in fact implies $m \leq 2(J+1)-1$.

For each of the cases (a) to (c), we call the interval $[\theta_{j-1},\theta_j]$ a *separating*, *pooling* and *bi-pooling* interval, respectively. Moreover, in case $[\theta_{j-1},\theta_j]$ is a bi-pooling interval, we refer to $p_j$ (extended to $\mathbb{R}$) as the *bi-tangent* of $V$ on the interval $[\theta_{j-1},\theta_j]$.

We first establish the necessary and sufficient conditions for $V$ to have a bi-tangent on $[\overline{\theta}_0,1]$. We note that, because $V$ is constant and equals zero, $V$ always has a bi-tangent on any subinterval of $[0,\overline{\theta}_0]$.)

**Lemma 2.** *$V$ has a bi-tangent on $[\overline{\theta}_0,1]$ if and only if (i) tangent to $V_1$ at $\overline{\theta}_0$ lies above $V_2$ on $[\overline{\theta}_0,1]$ and (ii) tangent to $V_2$ at 1 lies above $V_1$ on $[\overline{\theta}_0,1]$. Moreover, when it exists, the bi-tangent is unique, and it is tangent to $V$ at some $E_1\in[\overline{\theta}_0,\overline{\theta}_1)$ and $E_2\in[\overline{\theta}_1,1]$.*

Recall $c^E(\cdot)$ denotes the tangent to $c$ at $E$. Then, condition (i) of Lemma 2 holds if and only if

$$c(\theta)-c^{\overline{\theta}_0}(\theta)\geq v_2(\theta)-v_1(\theta)\ \forall\theta\in\left[\overline{\theta}_1,1\right], \tag{20}$$

where we note that $v_2(\cdot)-v_1(\cdot)\geq 0$ on $[\overline{\theta}_1,1]$. Condition (ii) of Lemma 2 holds if and only if

$$v_1(\theta)-v_2(\theta)\geq c^1(\theta)-c(\theta)\ \forall\theta\in\left[\overline{\theta}_0,\overline{\theta}_1\right], \tag{21}$$

where we note that $v_1(\cdot)-v_2(\cdot)\geq 0$ on $[\overline{\theta}_0,\overline{\theta}_1]$. The first condition requires $c$ to exhibit sufficient convexity while the second condition is a limit to how much convexity $c$ can exhibit. Thus, a bi-tangent exists whenever $c$ is sufficiently convex but not too convex; in other words, a bi-tangent exists whenever the increase in the marginal product of labour is not too little or too much.

**Lemma 3.** *There exists a feasible $G^*$ that solves the information design problem,* (3)*; i.e.,*

$$\widehat{V}^*=\max_{(\theta_0,\theta_1,E_2)}\int_\Theta V(E)\,\mathrm{d}G_{\theta_0,\theta_1,E_2}(\theta)\ \textit{s.t.}\ (\theta_0,\theta_1,E_2)\ \textit{is feasible}. \tag{22}$$

*Proof.* By Fact 2(i), a bi-pooling $G$ solves (3). Because $V$ is continuous, linear on $[\overline{\theta}_{-1}:=0,\overline{\theta}_0]$, strictly convex on the intervals $[\overline{\theta}_0,\overline{\theta}_1]$ and $[\overline{\theta}_1,\overline{\theta}_2:=1]$, and has convex kinks at $\overline{\theta}_0$ and $\overline{\theta}_1$, by Fact 2(ii), any bi-pooling interval can be found by identifying the bi-tangents of $V$.

Consider first the case in which $V$ has a bi-tangent on $[\overline{\theta}_0,1]$. By Lemma 2, bi-tangent is unique and we denote it as $V^b(\cdot)$. Let $\theta_0$ be such that $V^b(\theta_0)=0$. Given the assumptions on $V$, $\theta_0^b<\overline{\theta}_0$. Denote the two bi-tangency points as $E_1\in[\overline{\theta}_0,\overline{\theta}_1)$ and $E_2\in[\overline{\theta}_1,1]$. If $E_1$ and $E_2$ satisfy (19), then $[\theta_0^b,1]$ is a bi-pooling interval. Since $V$ is constant and equals zero on

$[0,\overline{\theta}_0]$, $\max\{V^b(\cdot),0\}$ is function satisfying (15)–(18). Moreover, $G$ that can be supported by such a $p$ is associated with $(\theta_0,\theta_1,E_2)$, where $\theta_1$ satisfies $\mathbb{E}[\theta|\theta_0<\theta<\theta_1]=E_2$; i.e., $G$ is feasible.

If an optimal $G$ is not bi-pooling (either because a bi-tangent does not exist or even if it exists, the tangency points, $E_1$ and $E_2$, do not satisfy (19)), then $G$ must be a pooling distribution. By Fact 2(ii), a support function of an optimal distribution $G$, $p:\Theta\to\mathbb{R}$, is a concatenation of at most 5 affine functions; i.e., there exists a collection $\{p_j:[\theta_{j-1},\theta_j]\to\mathbb{R}\}_{j=0}^4$, where

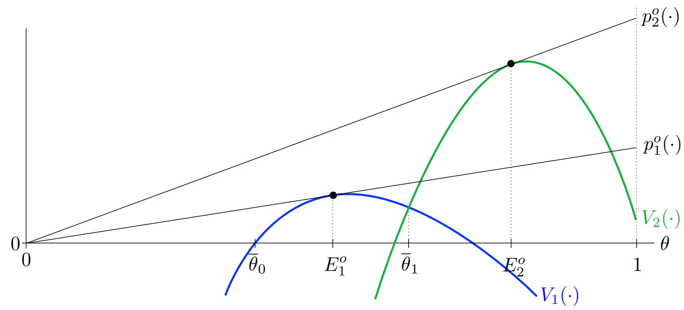
$$0=\theta_{-1}<\theta_0<\overline{\theta}_0<\theta_1<\theta_2<\overline{\theta}_1<\theta_3<\theta_4=1, \tag{23}$$

such that $p|_{[\theta_{j-1},\theta_j]}=p_j$ for each $j\in\{0,1,\ldots,4\}$. Moreover, each $p_j$ is such that either $p_j=V|_{[\theta_{j-1},\theta_j]}$ and $p_j$ is convex (i.e., $[\theta_{j-1},\theta_j]$ is a separating interval), or $p_j\geq V_{|[\theta_{j-1},\theta_j]}$ and $p_j$ is affine and tangent to $V$ at $\mathbb{E}[\theta|\theta_{j-1}\leq\theta\leq\theta_j]$ (i.e., $[\theta_{j-1},\theta_j]$ is a pooling interval). Since $V$ is convex (in fact, linear) on $[0,\overline{\theta}_0]$, $p_0$ must be a constant function equal to zero. Furthermore, to ensure that $p$ is convex and $p\geq V$, $p_1$ must be tangent to $V$ at some $E_1\in[\overline{\theta}_0,\overline{\theta}_1]$, and $p_4$ must be tangent to $V$ at some $E_2\in[\overline{\theta}_1,1]$. However, because $V$ is strictly concave on $[\overline{\theta}_0,\overline{\theta}_1]$ and on $[\overline{\theta}_1,1]$, and has a convex kink at $\overline{\theta}_1$, $p_2$, if it exists, would lie strictly above $V$; contradicting that $p_2$ is tangential to $V$ on $[\theta_1,\theta_2]$. Moreover, any point in which $p_3$ is tangential to $V$ would be a concave kink; contradicting the convexity of $p$. Thus, it follows that $p$ must be a concatenation of at most three affine functions. Moreover, any $G$ that is pooling that can be supported by such a $p$ is associated with $(\theta_0,\theta_1,\mathbb{E}[\theta|\theta\geq\theta_1])$ and so feasible. ■

## A.3 Optimal feasible $G$

As shown in Figure 3, for each $j\in\{1,2\}$, let $p_j^o:\Theta\to\mathbb{R}$ denote the smallest linear function that exceeds $V_j$, and let $\mathbb{E}_j^o$ denote the smallest $E\in\Theta$ such that $p_j^o(E)=V_j(E)$. Observe that while $p_2^o\geq p_1^o$ implies $\max V_2>\max V_1$, the converse is not true.

Figure 3: $p_j^o$ and $E_j^o$'s.




We first consider the case when $p_2^o \geq p_1^o$.

**Proposition 6.** *Suppose $p_2^o \geq p_1^o$ and $\mathbb{E}[\theta] \geq E_2^o$. Then, $G = F$ solves the firm's problem and assigning all workers to position $j = 2$ is optimal.*

*Proof.* Because $V_2$ is concave on $[E_2^o, 1] \subseteq [\overline{\theta}_1, 1]$, then $G = F$ (i.e., providing no information) is optimal for the firm because it can be supported by $p$ given by the tangent to $V$ at $\mathbb{E}[\theta]$. Since $\mathbb{E}[\theta] \in [\overline{\theta}_1, 1]$, the firm maximises its expected payoff by assigning all workers to position $j = 2$. ■

Given a feasible $G$ associated with $(\theta_0, \theta_1, E_2)$, we define the firm's expected payoff as:

$$\begin{aligned}\widehat{V}(\theta_0, \theta_1, E_2) &= \int_{\Theta} V(E)\, \mathrm{d}G_{\theta_0,\theta_1,E_2}(\theta) \\ &= (1 - F(\theta_0)) \left[ V(\mathbb{E}[\theta|\theta_0 < \theta < \theta_1])\, \alpha_{\theta_0,\theta_1,E_2} + V(E_2) \left(1 - \alpha_{\theta_0,\theta_1,E_1}\right) \right],\end{aligned}$$

where

$$\alpha_{\theta_0,\theta_1,E_2} := \begin{cases} 0 & \text{if } \theta_1 = 1 \text{ and } E_2 = \mathbb{E}[\theta|\theta > \theta_0], \\ \frac{E_2 - \mathbb{E}[\theta|\theta > \theta_0]}{E_2 - \mathbb{E}[\theta|\theta_0 < \theta < \theta_1]} & \text{otherwise.} \end{cases}$$

The problem for the firm now, (4), is to maximise $\widehat{V}$ with respect to $(\theta_0, \theta_1, E_2)$ among all feasible tuples. Since $\widehat{V}$ is differentiable, a necessary condition for optimality is that $(\theta_0, \theta_1, E_2)$ satisfies the first-order conditions (while accounting for boundary solutions). We establish sufficiency by constructing the support function for $G$ associated with $(\theta_0, \theta_1, E_2)$ as an upper envelope of tangents to $\widehat{V}$ at $\mathbb{E}[\theta|\theta_0 < \theta < \theta_1]$ and $E_2$, while imposing an additional condition to ensure that the support function is convex. We say that a feasible tuple is *nondegenerate* if it is not associated with $F$.

**Lemma 4.** *A feasible nondegenerate tuple* $(\theta_0^*, \theta_1^*, E_2^*)$ *solves the firm's problem,* (4)*, if and only if it satisfies the first-order conditions and* $V'(\mathbb{E}[\theta|\theta \geq \theta_1^*]) \geq V'(\mathbb{E}[\theta|\theta_0^* < \theta < \theta_1^*])$.

*Proof.* The idea is to construct a supporting function $p^*$ that supports $G_{(\theta_0^*, \theta_1^*, E_2^*)}$ using the first-order conditions to (4). Then, by Fact 2, it would follow that $(\theta_0^*, \theta_1^*, E_2^*)$ must be optimal. To that end, let us first fix $\theta_0 \in (0, \overline{\theta}_0)$ and $\theta_1 \in (\overline{\theta}_0, 1)$ and solve

$$\max_{E_2 \in \Theta} \widehat{V}\left(\theta_0, \theta_1, E_2\right) \text{ s.t. } \mathbb{E}\left[\theta|\theta > \theta_0\right] \leq E_2 \leq \mathbb{E}\left[\theta|\theta \geq \theta_1\right].$$

We denote the maximiser as $E_2^*(\theta_0, \theta_1)$. The derivative of the objective with respect to $E_2$ is

$$\frac{\mathrm{d}\widehat{V}\left(\cdot\right)}{\mathrm{d}E_2} = \left(1 - F\left(\theta_0\right)\right)\left(1 - \alpha_{\theta_0, \theta_1, E_2}\right)\left[V'\left(E_2\right) - \frac{V\left(E_2\right) - V\left(\mathbb{E}\left[\theta|\theta_0 < \theta < \theta_1\right]\right)}{E_2 - \mathbb{E}\left[\theta|\theta_0 < \theta < \theta_1\right]}\right].$$

If $\frac{\mathrm{d}\widehat{V}(\cdot)}{\mathrm{d}E_2}|_{E_2 = \mathbb{E}[\theta|\theta \geq \theta_1]} > 0$, we let $E_2^*(\theta_0, \theta_1) = \mathbb{E}[\theta|\theta \geq \theta_1]$ and if $\frac{\mathrm{d}\widehat{V}(\cdot)}{\mathrm{d}E_2}|_{E_2 = \mathbb{E}[\theta|\theta > \theta_0]} < 0$, we let $E_2^*(\theta_0, \theta_1) = \mathbb{E}[\theta|\theta > \theta_0]$. Otherwise, we let $E_2^* \equiv E_2^*(\theta_0, \theta_1)$ be the solution to

$$V\left(E_2^*\right) - V\left(\mathbb{E}\left[\theta|\theta_0 < \theta < \theta_1\right]\right) = V'\left(E_2^*\right)\left(E_2^* - \mathbb{E}\left[\theta|\theta_0 < \theta < \theta_1\right]\right).$$

Next, consider the program:

$$\max_{\theta_0 \in \left[0, \overline{\theta}_0\right],\ \theta_1 \in \left[\overline{\theta}_0, 1\right]} \widehat{V}\left(\theta_0, \theta_1, E_2^*\left(\theta_0, \theta_1\right)\right) \text{ s.t. } \theta_1 \geq \theta_0.$$

Let $\theta_0^*$ and $\theta_1^*$ be the maximisers. There are three cases to consider:

(i) $\mathbb{E}[\theta|\theta > \theta_0^*] < E_2^*(\theta_0^*, \theta_1^*) < \mathbb{E}[\theta|\theta \geq \theta_1^*]$.

(ii) $E_2^*(\theta_0^*, \theta_1^*) = \mathbb{E}[\theta|\theta \geq \theta_1^*]$ because $\frac{\mathrm{d}\widehat{V}(\theta_0^*, \theta_1^*, \mathbb{E}[\theta|\theta \geq \theta_1^*])}{\mathrm{d}E_2} > 0$.

(iii) $E_2^*(\theta_0^*, \theta_1^*) = \mathbb{E}[\theta|\theta > \theta_0^*]$ because $\frac{\mathrm{d}\widehat{V}(\theta_0^*, \theta_1^*, \mathbb{E}[\theta|\theta \geq \theta_1^*])}{\mathrm{d}E_2} < 0$.

We consider case (i) first. Letting $E_2^* \equiv E_2^*(\theta_0, \theta_1)$ and $\alpha^*_{\theta_0, \theta_1} \equiv \alpha_{\theta_0, \theta_1, E_2^*(\theta_0, \theta_1)}$, the

Envelope theorem gives us that

$$\begin{aligned}
\frac{\mathrm{d}\widehat{V}\left(\theta_1,\theta_1,E_2^*\right)}{\mathrm{d}\theta_1} &= \left.\frac{\partial\widehat{V}\left(\theta_0,\theta_1,E_2\right)}{\partial\theta_1}\right|_{E_2=E_2^*} \\
&= f\left(\theta_1\right)\left(\theta_1-\mathbb{E}\left[\theta|\theta_0<\theta<\theta_1\right]\right)\frac{1-F\left(\theta_0\right)}{F\left(\theta_1\right)-F\left(\theta_0\right)}\alpha^*_{\theta_0,\theta_1} \\
&\quad\times\left[V'\left(\mathbb{E}\left[\theta|\theta_0<\theta<\theta_1\right]\right)-\frac{V\left(E_2^*\right)-V\left(\mathbb{E}\left[\theta|\theta_0<\theta<\theta_1\right]\right)}{E_2^*-\mathbb{E}\left[\theta|\theta_0<\theta<\theta_1\right]}\right]
\end{aligned}$$

and

$$\begin{aligned}
&\frac{\mathrm{d}\widehat{V}\left(\theta_0,\theta_1,E_2^*\right)}{\mathrm{d}\theta_0} = \left.\frac{\partial\widehat{V}\left(\theta_0,\theta_1,E_2\right)}{\partial\theta_0}\right|_{E_2=E_2^*} \\
&= \frac{f\left(\theta_0\right)}{f\left(\theta_1\right)}\frac{\mathbb{E}\left[\theta|\theta_0<\theta<\theta_1\right]-\theta_0}{\theta_1-\mathbb{E}\left[\theta|\theta_0<\theta<\theta_1\right]}\frac{\mathrm{d}\widehat{V}\left(\theta_0,\theta_1,E_2^*\left(\theta_0,\theta_1\right)\right)}{\mathrm{d}\theta_1} \\
&\quad + f\left(\theta_0\right)\left(\mathbb{E}\left[\theta|\theta>\theta_0\right]-\theta_0\right)\frac{V\left(E_2^*\right)-V\left(\mathbb{E}\left[\theta|\theta_0<\theta<\theta_1\right]\right)}{E_2^*-\mathbb{E}\left[\theta|\theta_0<\theta<\theta_1\right]} \\
&\quad - f\left(\theta_0\right)\left(\mathbb{E}\left[\theta|\theta>\theta_0\right]-\theta_0\right)\frac{V\left(\mathbb{E}\left[\theta|\theta_0<\theta<\theta_1\right]\right)\alpha^*_{\theta_0,\theta_1}+V\left(E_2^*\right)\left(1-\alpha^*_{\theta_0,\theta_1}\right)}{\mathbb{E}\left[\theta|\theta>\theta_0\right]-\theta_0}.
\end{aligned}$$

Thus, at any interior optimal $\theta_0^*\in(0,\theta_0)$ and $\theta_1^*\in(\theta_0^*,1)$, it must be that

$$V'\left(\mathbb{E}\left[\theta|\theta_0^*<\theta<\theta_1^*\right]\right)\left(E_2^*-\mathbb{E}\left[\theta|\theta_0^*<\theta<\theta_1^*\right]\right)=V\left(E_2^*\right)-V\left(\mathbb{E}\left[\theta|\theta_0^*<\theta<\theta_1^*\right]\right),$$

which ensures that the impact on the average profit from workers equals the gains from more workers being assigned to $j=1$ net of losses from fewer workers being assigned to $j=2$; and

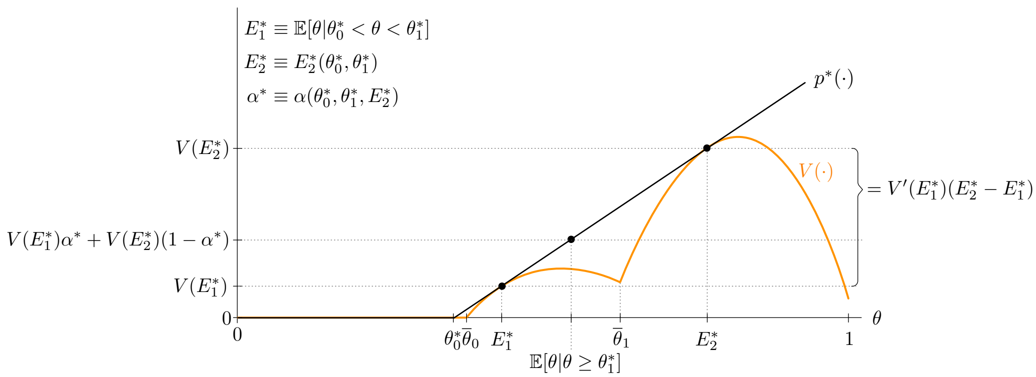
$$\frac{V\left(E_2^*\right)-V\left(\mathbb{E}\left[\theta|\theta_0^*<\theta<\theta_1^*\right]\right)}{E_2^*-\mathbb{E}\left[\theta|\theta_0^*<\theta<\theta_1^*\right]}=\frac{V\left(\mathbb{E}\left[\theta|\theta_0^*<\theta<\theta_1^*\right]\right)\alpha^*_{\theta_0^*,\theta_1^*}+V\left(E_2^*\right)\left(1-\alpha^*_{\theta_0^*,\theta_1^*}\right)}{\mathbb{E}\left[\theta|\theta>\theta_0^*\right]-\theta_0^*},$$

which ensures that changes in profits due to changes in workers assigned to $j=1$ or 2 equals the changes in profits from retaining fewer workers. If an interior solution $(\theta_0^*,\theta_1^*,E_2^*)$ to the first-order condition exists, then we can construct a convex supporting function $p^*$ as follows:

$$p^*(\theta) \coloneqq \left[V'\left(\mathbb{E}\left[\theta|\theta_0^* < \theta < \theta_1^*\right]\right)\left(\theta - \theta_0^*\right)\right]_+ \\ = \begin{cases} V(\theta) = 0 & \text{if } \theta \in [0, \theta_0^*], \\ V'\left(\mathbb{E}\left[\theta|\theta_0^* < \theta < \theta_1^*\right]\right)\left(\theta - \theta_0^*\right) & \text{if } \theta \in (\theta_0^*, 1]. \end{cases}$$

By construction, $p^*$ is convex and $p^* \geq V$. One can also verify that $p^*$ and $G_{\theta_0^*,\theta_1^*,E_2(\theta_0^*,\theta_1^*)}$ satisfy (16)–(18). Hence, we conclude that $G_{\theta_0^*,\theta_1^*,E_2^*(\theta_0^*,\theta_1^*)}$ solves (3). Observe that this corresponds to the case in which the interval $(\theta_0^*, 1)$ is bi-pooled to $\mathbb{E}[\theta|\theta_0^* < \theta < \theta_1^*]$ and $E_2(\theta_0^*, \theta_1^*)$. Figure 4 depicts the typical example of this case.

Figure 4: Case 1: $\mathbb{E}[\theta|\theta \geq \theta_0^*] < E_2^* < \mathbb{E}[\theta|\theta \geq \theta_1^*]$.




We now consider case (ii) in which $E_2^*(\theta_0^*, \theta_1^*) = \mathbb{E}[\theta|\theta \geq \theta_1^*]$. The expected payoff for the firm now is:

$$\widehat{V}\left(\theta_0^*, \theta_1^*, \mathbb{E}\left[\theta|\theta \geq \theta_1^*\right]\right) = V\left(\mathbb{E}\left[\theta|\theta_0^* < \theta < \theta_1^*\right]\right)\left(F\left(\theta_1^*\right) - F\left(\theta_0^*\right)\right) \\ + V\left(\mathbb{E}\left[\theta|\theta \geq \theta_1^*\right]\right)\left(1 - F\left(\theta_1^*\right)\right)$$

Observe that

$$\frac{\mathrm{d}\widehat{V}\left(\theta_0, \theta_1, \mathbb{E}\left[\theta|\theta \geq \theta_1\right]\right)}{\mathrm{d}\theta_0} \\ = f(\theta_0)\left[\mathbb{E}\left[\theta|\theta_0 < \theta < \theta_1\right] - \theta_0\right]\left(V'\left(\mathbb{E}\left[\theta|\theta_0 < \theta < \theta_1\right]\right) - \frac{V\left(\mathbb{E}\left[\theta|\theta_0 < \theta < \theta_1\right]\right)}{\mathbb{E}\left[\theta|\theta_0 < \theta < \theta_1\right] - \theta_0}\right),$$

and

$$
\begin{aligned}
&\frac{\mathrm{d}\widehat{V}\left(\theta_0,\theta_1,\mathbb{E}\left[\theta|\theta\geq\theta_1\right]\right)}{\mathrm{d}\theta_1}\\
&=f\left(\theta_1\right)\big(V\left(\mathbb{E}\left[\theta|\theta_0<\theta<\theta_1\right]\right)+V'\left(\mathbb{E}\left[\theta|\theta_0<\theta<\theta_1\right]\right)\left(\theta_1-\mathbb{E}\left[\theta|\theta_0<\theta<\theta_1\right]\right)\big)\\
&\qquad+f\left(\theta_1\right)V'\left(\mathbb{E}\left[\theta|\theta\geq\theta_1\right]\right)\left(\mathbb{E}\left[\theta|\theta\geq\theta_1\right]-\theta_1\right)\\
&\qquad-f\left(\theta_1\right)\left[V\left(\mathbb{E}\left[\theta|\theta\geq\theta_1\right]\right)\right].
\end{aligned}
$$

Suppose $\theta_0^*>0$, then the first-order condition with respect to $\theta_0$ means

$$
V'\left(\mathbb{E}\left[\theta|\theta_0^*<\theta<\theta_1^*\right]\right)\left(\mathbb{E}\left[\theta|\theta_0^*<\theta<\theta_1^*\right]-\theta_0^*\right)=V\left(\mathbb{E}\left[\theta|\theta_0^*<\theta<\theta_1^*\right]\right),
$$

but if $\theta_0^*=0$, then

$$
\frac{\mathrm{d}\widehat{V}\left(0,\theta_1^*,\mathbb{E}\left[\theta|\theta\geq\theta_1^*\right]\right)}{\mathrm{d}\theta_0}=V'\left(\mathbb{E}\left[\theta|\theta<\theta_1^*\right]\right)-\frac{V\left(\mathbb{E}\left[\theta|\theta<\theta_1^*\right]\right)}{\mathbb{E}\left[\theta|\theta<\theta_1^*\right]}<0.
$$

The first-order condition with respect to $\theta_1$ is

$$
\begin{aligned}
V\left(\mathbb{E}\left[\theta|\theta\geq\theta_1^*\right]\right)=V&\left(\mathbb{E}\left[\theta|\theta_0^*<\theta<\theta_1^*\right]\right)\\
&+V'\left(\mathbb{E}\left[\theta|\theta_0^*<\theta<\theta_1^*\right]\right)\left(\theta_1^*-\mathbb{E}\left[\theta|\theta_0^*<\theta<\theta_1^*\right]\right)\\
&+V'\left(\mathbb{E}\left[\theta|\theta\geq\theta_1^*\right]\right)\left(\mathbb{E}\left[\theta|\theta\geq\theta_1^*\right]-\theta_1^*\right)
\end{aligned}
$$

Letting $E_1^*\equiv\mathbb{E}[\theta|\theta_0^*<\theta<\theta_1^*]$ and $E_2^*\equiv\mathbb{E}[\theta|\theta\geq\theta_1^*]$, construct $p^*$ as follows:

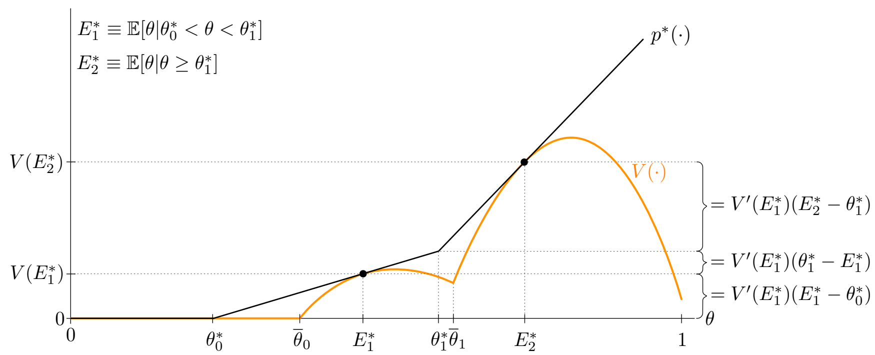
$$
\begin{aligned}
p^*\left(\theta\right)&:=\max\Big\{\left[V\left(E_1^*\right)+V'\left(E_1^*\right)\left(\theta-E_1^*\right)\right]_+,V\left(E_2^*\right)+V'\left(E_2^*\right)\left(\theta-E_2^*\right)\Big\}\\
&=\begin{cases}\left[V\left(E_1^*\right)+V'\left(E_1^*\right)\left(\theta-E_1^*\right)\right]_+ & \text{if } \theta\in[0,\theta_1^*],\\ V\left(E_2^*\right)+V'\left(E_2^*\right)\left(\theta-E_2^*\right) & \text{if } \theta\in(\theta_1^*,1].\end{cases}
\end{aligned}
\tag{24}
$$

In case $\theta_0^*>0$, the first-order condition with respect to $\theta_0$ means that $V'(E_1^*)>0$ so that $p^*$ is convex on $[0,\theta_1^*)$. If $\theta_0^*=0$, then $p^*$ is affine (and thus convex) on $[0,\theta_1^*)$. To ensure that $p^*$ is convex on $[0,1]$, we further require that $V'(E_2^*)\geq V'(E_1^*)$. By construction, $p^*\geq V$. Moreover, $p^*$ and $G_{\theta_0^*,\theta_1^*,E_2^*(\theta_0^*,\theta_1^*)}$ satisfy (16)–(18) and so we conclude that $G_{\theta_0^*,\theta_1^*,E_2^*(\theta_0^*,\theta_1^*)}$ solves (3). Observe that, in this case, the interval $(\theta_0^*,\theta_1^*)$ is pooled to $E_1^*)$ and the interval $[\theta_1^*,1]$ is pooled to $E_2^*$. Figure 5 shows the typical example of the case when $E_2^*(\theta_0^*,\theta_1^*)=$

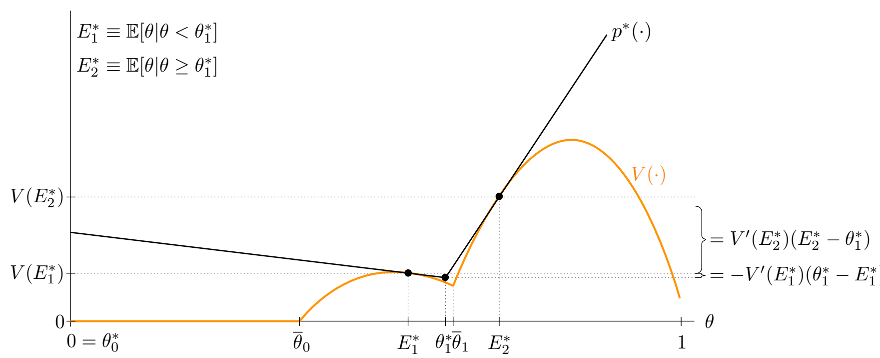
$\mathbb{E}[\theta|\theta \geq \theta_1^*]$ and either $\theta_0^* \in (0, \overline{\theta}_0)$ (Figure 5a) or $\theta_0^* = 0$ (Figure 5b).

Figure 5: Case 2: $E_2^*(\theta_0^*, \theta_1^*) = \mathbb{E}[\theta|\theta \geq \theta_1^*]$.

(a) Case 2a: $\theta_0^* \in (0, \overline{\theta}_0)$.




(b) Case 2b: $\theta_0^* = 0$.




Finally, we consider case (iii) in which $E_2^*(\theta_0^*, \theta_1^*) = \mathbb{E}[\theta|\theta > \theta_0^*]$. The expected payoff for the firm is now:

$$\widehat{V}\left(\theta_0^*, \theta_1^*, \mathbb{E}\left[\theta|\theta > \theta_0^*\right]\right) = V\left(\mathbb{E}\left[\theta|\theta > \theta_0^*\right]\right)\left(1 - F\left(\theta_0^*\right)\right).$$

Since the payoff does not depend on $\theta_1$, we set it to $\theta_1^* = 1$. Observe that

$$\frac{\mathrm{d}\widehat{V}\left(\theta_0, \theta_1, \mathbb{E}\left[\theta|\theta > \theta_0\right]\right)}{\mathrm{d}\theta_0} = f\left(\theta_0\right)\left[\mathbb{E}\left[\theta|\theta > \theta_0\right] - \theta_0\right]\left(V'\left(\mathbb{E}\left[\theta|\theta > \theta_0\right]\right) - \frac{V\left(\mathbb{E}\left[\theta|\theta > \theta_0\right]\right)}{\mathbb{E}\left[\theta|\theta > \theta_0\right] - \theta_0}\right).$$

Thus, the first-order condition then requires that, if $\theta_0^* \in (0, \theta_0^*)$ and $\theta_1^* \in (\theta_0^*, 1)$,

$$V'\left(\mathbb{E}\left[\theta|\theta > \theta_0^*\right]\right)\left(\mathbb{E}\left[\theta|\theta > \theta_0^*\right] - \theta_0^*\right) = \frac{V\left(\mathbb{E}\left[\theta|\theta > \theta_0^*\right]\right)}{\mathbb{E}\left[\theta|\theta > \theta_0^*\right] - \theta_0^*};$$

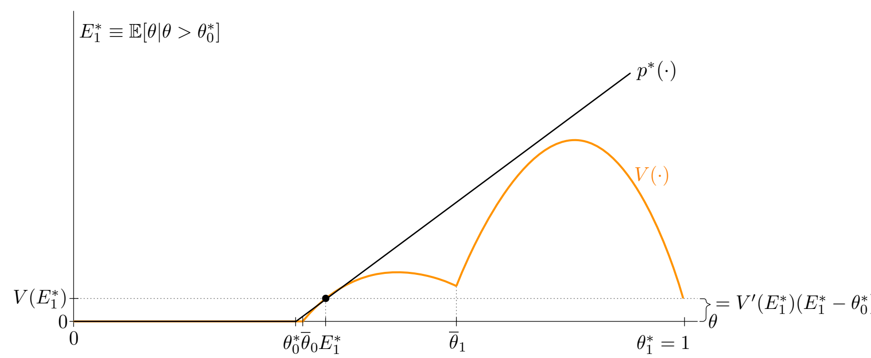
i.e., the tangent to $V$ at $\mathbb{E}[\theta|\theta > \theta_0^*]$ crosses the $x$-axis at $\theta_0^*$. Construct $p^*$ now as follows:

$$p^*(\theta) = \begin{cases} V(\theta) = 0 & \text{if } \theta \in [0, \theta_0^*], \\ V'\left(\mathbb{E}\left[\theta|\theta \geq \theta_0^*\right]\right)\left(\theta - \theta_0^*\right) & \text{if } \theta \in (\theta_0^*, 1]. \end{cases}.$$

Observe that $p^*$ is convex, $p^* \geq V$. Moreover, $p^*$ and $G_{\theta_0^*,\theta_1^*,E_2^*(\theta_0^*,\theta_1^*)}$ satisfy (16)—(18) so that $G_{\theta_0^*,\theta_1^*,E_2^*(\theta_0^*,\theta_1^*)}$ solves (3). In this case, the interval $(\theta_0^*, 1]$ is pooled to $\mathbb{E}[\theta|\theta > \theta_0^*]$.

Figure 6: Case 3: $E_2^*(\theta_0^*, \theta_1^*) = \mathbb{E}[\theta|\theta > \theta_0^*]$.

(a) Case 3a: $E_1^* \in (\overline{\theta}_0, \overline{\theta}_1)$.




(b) Case 3b: $E_1^* \in (\overline{\theta}_1, 1)$.

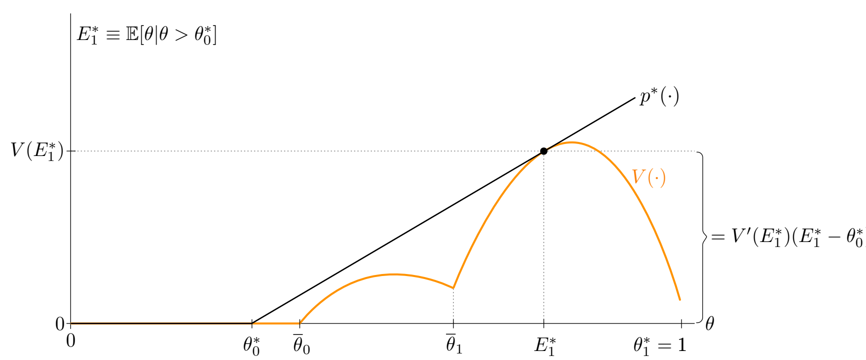
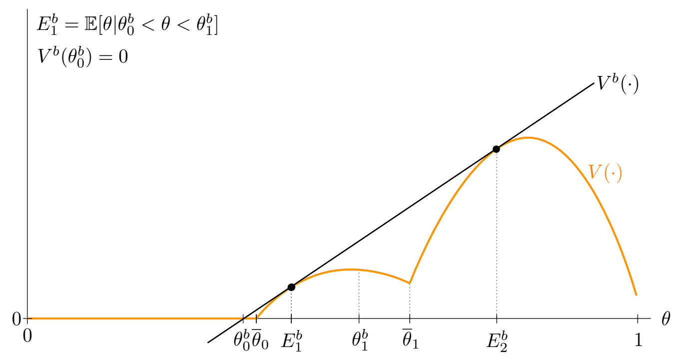



Figure 6 shows the typical examples of the case when $E_2^*(\theta_0^*, \theta_1^*) = \mathbb{E}[\theta|\theta > \theta_0^*]$. ■

We now characterise feasible $G$ that solves the firm's information design problem, (4). Towards this goal, in case $V$ has a bi-tangent, we denote it as $V^b(\cdot)$, and let $E_1^b$ and $E_2^b$ denote the two tangency points, where $E_2^b > E_1^b$, and let $\theta_0^b$ and $\theta_1^b$ be points such that $V^b(\theta_0^b) = 0$ and $\mathbb{E}[\theta|\theta_0^b < \theta < \theta_1^b] = E_1^b$. Figure 7 shows how $E_1^b$, $E_2^b$, $\theta_0^b$ and $\theta_1^b$ are obtained.

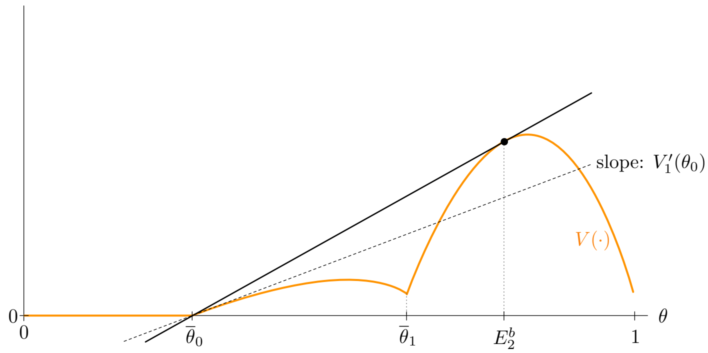
Figure 7: $V$ has a bi-tangent, $V^b$.



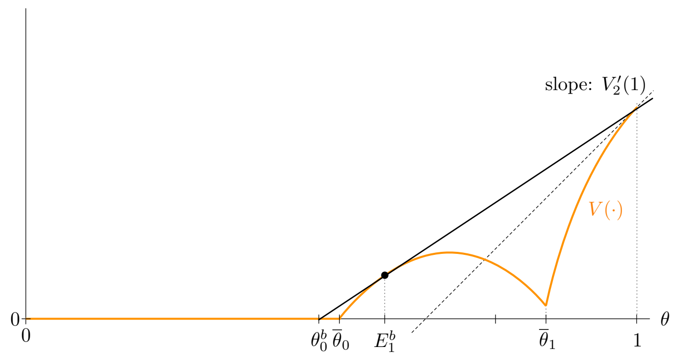

In case $V$ does not have a bi-tangent, then at least one of (20) and (21) does not hold. If (20) does not hold, then let $E_2^b \in [\overline{\theta}_1, 1]$ be the point such that the tangent to $V_2$ at $E_2^b$ goes through the point $(\overline{\theta}_0, 0)$. If (21) does not hold, then let $E_1^b \in [\overline{\theta}_0, \overline{\theta}_1]$ be the point such that the tangent to $V_1$ at $E_1^b$ goes through the point $(1, V_2(1))$, and let $\theta_0^b$ be the point such that the tangents equals zero.

Figure 8: $V$ does not have a bi-tagent.

(a) Case: (20) does not hold.




(b) Case: (21) does not hold




**Theorem 3.** *Suppose $p_2^o \geq p_1^o$ and $\mathbb{E}[\theta] < E_2^o$. Suppose a bi-tangent of $V$ exists.*

*(i) If*

$$E_1^b \leq \mathbb{E}[\theta|\theta > \theta_0^b] \leq E_2^b \leq \mathbb{E}[\theta|\theta \geq \theta_1^b], \tag{25}$$

*then an optimal $G$ exists that bi-pools the interval $[\theta_0^b, 1]$ to $E_1^b$ and $E_2^b$.*

*(ii) If* (25) *does not hold and*

$$E_2^b > \mathbb{E}[\theta|\theta \geq \theta_1^b], \tag{26}$$

*then an optimal $G$ exists that pools the interval $(\theta_0^*, \theta_1^*)$ to $\mathbb{E}[\theta|\theta_0^* < \theta < \theta_1^*] > E_1^b$, and pools the interval $[\theta_1^*, 1]$ to $E[\theta|\theta \geq \theta_1^*] < E_2^b$, where $\theta_0^* \in [0, \theta_0^b)$ and $\theta_1^* \in (\theta_0^*, 1)$.*

*(iii) If (25) does not hold and*

$$E_1^b > \mathbb{E}[\theta|\theta > \theta_0^b], \tag{27}$$

*then an optimal $G$ exists that pools the interval $(\theta_0^*, 1]$ to $\mathbb{E}[\theta|\theta > \theta_0^*]$.*

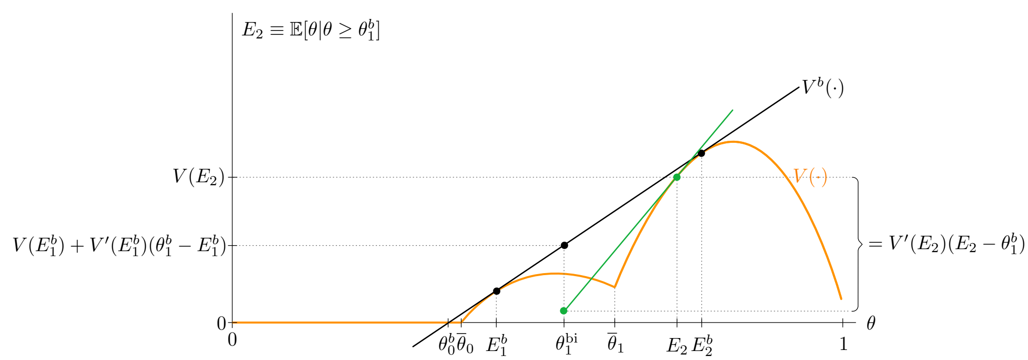
*Suppose $V$ does not have a bi-tangent: If* (20) *does not hold, then optimal $G$ has the same features as case (iii) above if $\mathbb{E}[\theta|\theta > \overline{\theta}_0] \geq E_2^b$ and otherwise $G$ has the same features as case (ii).* (21) *does not hold, then optimal $G$ has the same features as case (iii) above if $\theta_0^b < \mathbb{E}[\theta|\theta > \theta_0^b]$ and otherwise $G$ has the same features as case (ii).*

*Proof.* (i) By Fact 2(ii), the interval $[\theta_0^b, 1]$ can be bi-pooled at $E_1^b$ and $E_2^b$ if and only if (25) holds. This case corresponds to case 1 from Lemma 4 with $\theta_0^* = \theta_0^b$, $\theta_1^* = \theta_1^b$ and $E_2^* = E_2^b$.
(ii) This case includes the case in which (25) does not hold and $E_2^b > \mathbb{E}[\theta|\theta > \theta_0^b]$ because $\theta_1^b \geq \theta_0^b$. Consider first the case $E_2 \geq \overline{\theta}_1$. Letting denote $E_2 \equiv \mathbb{E}[\theta|\theta > \theta_0^b]$, we first argue that

$$V(E_2) < V(E_1^b) + V'(E_1^b)(\theta_1^b - E_1^b) + V'(E_2)(E_2 - \theta_1^b). \tag{28}$$

Toward a contradiction, suppose the inequality does not hold. Then, there must be a line through the point $(\theta_1^b, V(E_1^b) + V'(E_1^b)(\theta_1^b - E_1^b))$ that is tangent to $V$ at $\mathbb{E}[\theta|\theta \geq \theta_1^b]$. However, by the choice of $\theta_1^b$, $V$ has a unique tangent that goes through the point $(\theta_1^b, V(E_1^b) + V'(E_1^b)(\theta_1^b - E_1^b))$ given by the tangent at $E_2^b$—a contradiction. Figure 9 depicts a typical example of this case.

Figure 9: Case (ii): Positive difference.




For any $\theta_0 \in [0, \theta_0^b]$, let $E_1(\theta_0)$ be given by

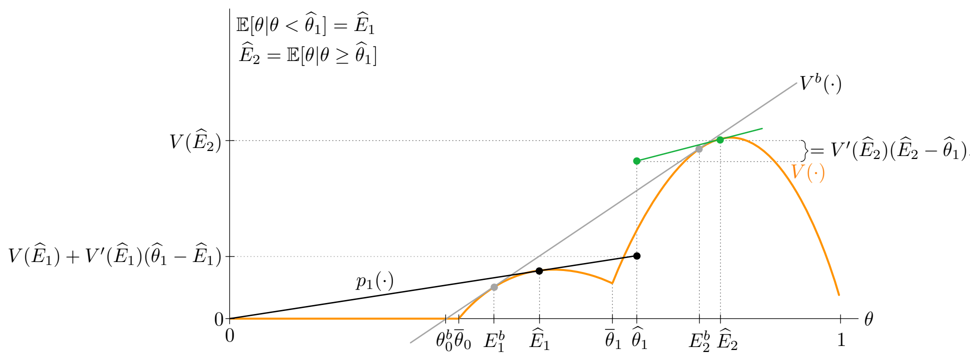
$$\theta_0 + V'(E_1(\theta_0))\, E_1(\theta_0) = V(E_1(\theta_0));$$

i.e., $E_1(\theta_0)$ is the point such that the line through $(\theta_0, 0)$ and $(E_1(\theta_0), V(E_1(\theta_0)))$ is tangent to $V$ at $E_1(\theta_0)$. Because $E_1^b \in (\overline{\theta}_0, \overline{\theta}_1)$, concavity of $V$ on $[\overline{\theta}_0, \overline{\theta}_1]$ means that $E_1(\theta_0)$ is continuous and strictly decreasing. Moreover, $\theta_1(\theta_0)$, defined implicitly by $\mathbb{E}[\theta|\theta_0 < \theta <$

$\theta_1(\theta_0)] = E_1(\theta_0)$, is also continuous and strictly decreasing. Now, let $p_1 : \theta \to \mathbb{R}$ denote the smallest linear function that exceeds $V_1$ and let $\widehat{E}_1$ denote the smallest $E \in \theta$ such that $p_1(E) = V_1(E)$. Let $\widehat{\theta}_1$ be such that $\mathbb{E}[\theta|\theta < \widehat{\theta}_1] = \widehat{E}_1$. Suppose, as shown in Figure 10,

$$V(\mathbb{E}[\theta|\theta \geq \widehat{\theta}_1]) > V(\widehat{E}_1) + V'(\widehat{E}_1)(\widehat{\theta}_1 - \widehat{E}_1) + V'(\mathbb{E}[\theta|\theta \geq \widehat{\theta}_1])(\mathbb{E}[\theta|\theta \geq \widehat{\theta}_1] - \widehat{\theta}_1). \quad (29)$$

Figure 10: Case (ii.a): Negative difference.



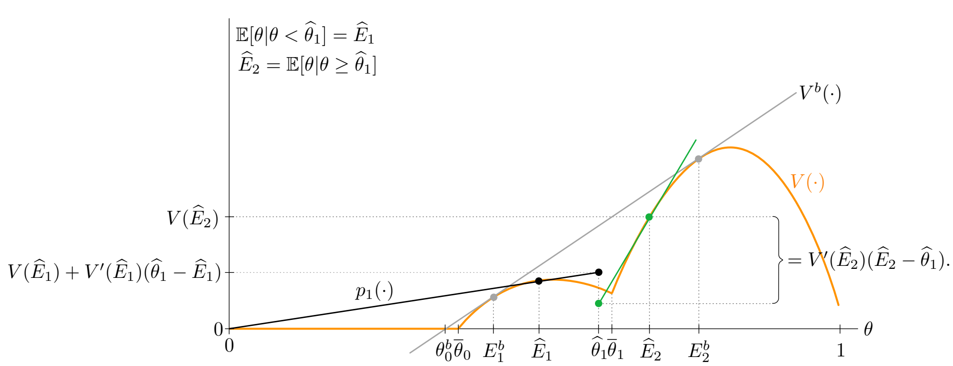

In this case, by the intermediate value theorem, there exists a $\theta_0 \in (0, \overline{\theta}_0)$ such that

$$\begin{aligned} V\left(\mathbb{E}\left[\theta|\theta \geq \theta_1\left(\theta_0\right)\right]\right) = V\left(E_1\left(\theta_0\right)\right) + V'\left(E_1\left(\theta_0\right)\right)\left(\theta_1\left(\theta_0\right) - E_1\left(\theta_0\right)\right) \\ + V'\left(\mathbb{E}\left[\theta|\theta \geq \theta_1\left(\theta_0\right)\right]\right)\left(\mathbb{E}\left[\theta|\theta \geq \theta_1\left(\theta_0\right)\right] - \theta_1\left(\theta_0\right)\right). \end{aligned}$$

Moreover, observe that such a $\theta_0$ is unique due to the concavity of $V$ on $[\overline{\theta}_0, \overline{\theta}_1]$ and on $[\overline{\theta}_1, 1]$. This case corresponds to the case 2a in the proof of Lemma 4.

Suppose now (29) does not hold as shown in Figure 11.

Figure 11: Case (ii.b): Positive difference.



Now let $\widetilde{\theta}_1 > \widehat{\theta}_1$ be such that $p_1(\widetilde{\theta}_1|\widetilde{\theta}_1) = V(\widetilde{\theta}_1)$, where

$$p_1\left(\theta'|\theta_1\right) := V\left(\mathbb{E}\left[\theta|\theta < \theta_1\right]\right) + V'\left(\mathbb{E}\left[\theta|\theta < \theta_1\right]\right)\left(\theta' - \mathbb{E}\left[\theta|\theta < \theta_1\right]\right).$$

We will argue that (see Figure 12):

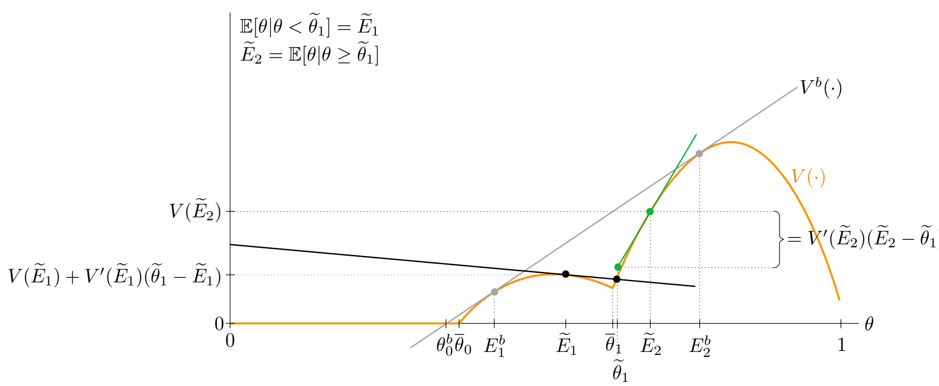
$$\begin{aligned} V(\mathbb{E}[\theta|\theta \geq \widetilde{\theta}_1]) > V(\mathbb{E}[\theta|\theta < \widetilde{\theta}_1]) + V'(\mathbb{E}[\theta|\theta < \widetilde{\theta}_1])(\widetilde{\theta}_1 - \mathbb{E}[\theta|\theta < \widetilde{\theta}_1]) \\ + V'(\mathbb{E}[\theta|\theta \geq \widetilde{\theta}_1])(\mathbb{E}[\theta|\theta \geq \widetilde{\theta}_1] - \widetilde{\theta}_1) \end{aligned}$$

By construction, $\widetilde{\theta}_1 \in (\overline{\theta}_1, 1)$ is such that the tangent to $V$ at $\mathbb{E}[\theta|\theta < \widetilde{\theta}_1] \in (\overline{\theta}_1, 1)$ intersects $V$ at $\widetilde{\theta}_1$. Since $V$ is concave on $[\overline{\theta}_1, 1]$, $V'(\cdot)$ is strictly decreasing on $(\widetilde{\theta}_1, 1]$, for any $E_2 \in (\widetilde{\theta}_1, E_2^b) \subseteq (\widetilde{\theta}_1, 1]$,

$$V'(E_2)(E_2 - \widetilde{\theta}_1) > V(E_2) - V(\tilde{\theta}_1).$$

Letting $E_2 = \mathbb{E}[\theta|\theta < \widetilde{\theta}_1]$ gives the desired inequality.

Figure 12: Case (ii.b): Negative difference.




By the intermediate value theorem, there exists $\theta_1 \in (\widehat{\theta}_1, \widetilde{\theta}_1)$ such that

$$\begin{aligned} V\left(\mathbb{E}\left[\theta|\theta \geq \theta_1\right]\right) = V\left(\mathbb{E}\left[\theta|\theta < \theta_1\right]\right) + V'\left(\mathbb{E}\left[\theta|\theta < \theta_1\right]\right)\left(\theta_1 - \mathbb{E}\left[\theta|\theta < \theta_1\right]\right) \\ + V'\left(\mathbb{E}\left[\theta|\theta \geq \theta_1\right]\right)\left(\mathbb{E}\left[\theta|\theta \geq \theta_1\right] - \theta_1\right). \end{aligned}$$

Moreover, observe that such a $\theta_0$ is unique due to the concavity of $V$ on $[\overline{\theta}_0, \overline{\theta}_1]$ and on $[\overline{\theta}_1, 1]$. This case corresponds to case 2b in the proof of Lemma 4.

Now suppose that $\mathbb{E}[\theta|\theta \geq \theta_1^b] < \overline{\theta}_1$. Because $V$ is concave on $[\overline{\theta}_0, \overline{\theta}_1]$, $\mathbb{E}[\theta|\theta \geq \theta_1^b] > E_1^b$ means that $V'(\mathbb{E}[\theta|\theta \geq \theta_1^b]) < V'(E_1^b)$. Hence, a function constructed as in (24) would not be convex. As already argued, choosing any $\theta_0 \in (0, \theta_0^b)$ means $E_1(\theta_0) > E_1^b$ and

$\theta_1(\theta_0) > \theta_1^b$ and $E_1(\theta_0) \geq \mathbb{E}[\theta|\theta \geq \theta_1^b]$. So long as $E_1(\theta_0) < \overline{\theta}_1$, a function constructed as in (24) would not be convex. If $E_1(\theta_0) \geq \overline{\theta}_1$ then we can proceed as above with $\theta_0$ replacing $\overline{\theta}_0$ so we are once again in either case 2a or 2b of Lemma 4.

(iii) Let $\widetilde{\theta}_0$ be such that $\mathbb{E}[\theta|\theta > \widetilde{\theta}_0] = E_1^b$ and note that $\widetilde{\theta}_0 > \overline{\theta}_0$ because of condition (27). Define $\gamma\colon [\theta_0^b, \widetilde{\theta}_0] \to \mathbb{R}$ as

$$\gamma(\theta_0) := V'(\mathbb{E}[\theta|\theta > \theta_0]) - \frac{V(\mathbb{E}[\theta|\theta > \theta_0])}{\mathbb{E}[\theta|\theta > \theta_0] - \theta_0}.$$

Since the line segment connecting $\theta_0^b$ and $E_1^b$ is tangent to $V$ at $E_1^b$, $\widetilde{\theta}_0 > \theta_0^b$, and $V$ is concave on $[\overline{\theta}_0, \overline{\theta}_1]$, the condition (27) implies that

$$\gamma(\theta_0^b) > 0 > \gamma(\widetilde{\theta}_0).$$

Since $\gamma$ is continuous, by the intermediate value theorem, there exists $\theta_0 \in (\theta_0^b, \widetilde{\theta}_0)$ such that $\gamma(\theta_0^*) = 0$. Observe that this case corresponds to case (iii) in the proof of Lemma 4.

Now, suppose that a bi-tangent does not exist. Consider the first the case in which (20) does not hold. If $\mathbb{E}[\theta|\theta > \overline{\theta}_0] \geq E_2^b$. Since $\mathbb{E}[\theta] \leq \overline{E}$, there exists $\theta_0 \in [0, \overline{\theta}_0]$ such that

$$V\left(\mathbb{E}\left[\theta|\theta > \theta_0\right]\right) = V'\left(\mathbb{E}\left[\theta|\theta > \theta_0\right]\right)\left(\mathbb{E}\left[\theta|\theta > \theta_0\right] - \theta_0\right),$$

which is case (iii) in the proof of Lemma 4. Now suppose $\mathbb{E}[\theta|\theta > \overline{\theta}_0] < \overline{E}_0$. Observe that neither case (iii) nor case (i) in the proof of Lemma 4 is possible; i.e., only case (ii) can arise.

Finally, suppose (21) does not hold. If $\mathbb{E}[\theta|\theta > \theta_0^b] < E_1^b$, then any tangency to $V$ on $[\overline{\theta}_0, 1]$ cannot be on $[\overline{\theta}_1, 1]$ so that this case corresponds to case (iii) in the proof of Lemma 4. If $\mathbb{E}[\theta|\theta > \theta_0^b] \geq E_1^b$, neither case (iii) nor case (i) in the proof of Lemma 4 is possible; i.e., only case (ii) can arise. ■

We now characterise the optimal $G$ when $p_1^o \geq p_2^o$.

**Theorem 4.** *Suppose $p_1^o \geq p_2^o$. Then, $V$ has a bi-tangent on $[\overline{\theta}_0, 1]$. Let $E_1^b$ and $E_2^b$ denote the two bi-tangency points.*

*(i) If $\mathbb{E}[\theta] \in [0, E_1^o]$, then $G$ associated with $(\theta_0, 1, 1)$ for some $\theta_0 \in [0, \overline{\theta}_0)$ solves the firm's problem. Moreover, no worker is ever assigned position $j = 2$.*

*(ii)* *If* $\mathbb{E}[\theta] \in [E_1^o, E_1^b]$, *then* $G = F$ *solves the firm's problem and assigning all workers to position* $j = 1$ *is optimal.*

*(iii)* *If* $\mathbb{E}[\theta] \in [E_1^b, E_2^b]$*...*

*(a)* *... and* (25) *holds with* $\theta_0^b = 0$,[41] *then an optimal* $G$ *bi-pools the* $\Theta$ *to* $E_1^b$ *and* $E_2^b$.

*(b)* *... but* (25) *does not hold, then an optimal* $G$ *pools the interval* $(0, \theta_1^*)$ *to* $\mathbb{E}[\theta|\theta < \theta_1^*] > E_1^b$, *and pools the interval* $[\theta_1^*, 1]$ *to* $E[\theta|\theta \geq \theta_1^*] < E_2^b$ *for some* $\theta_1^* \in \Theta$.

*(iv)* *If* $\mathbb{E}[\theta] \in [E_2^b, 1]$, *then* $G = F$ *solves the firm's problem and assigning all workers to position* $j = 2$ *is optimal.*

*Proof.* The result is immediate from previous arguments while observing that the slope of $p_1^o$ must be greater than the tangent so that $\theta_1^b = 0$. ■

## A.4 Inefficiencies

### A.4.1 Retention

Formally, we say that an HR policy $\varphi : \Theta \to \Delta(\{0,1,2\})$ exhibits:

- *over-retention* if $\int_0^{\overline{\theta}_0} \varphi(0|\theta) \mathrm{d}F(\theta) < F(\overline{\theta}_0)$;
- *under-retention* if $\int_{\overline{\theta}_0}^1 \varphi(0|\theta) \mathrm{d}F(\theta) > 0$;
- *efficient retention* if $\int_0^{\overline{\theta}_0} \varphi(0|\theta) \mathrm{d}F(\theta) = F(\overline{\theta}_0)$ and $\int_{\overline{\theta}_0}^1 \varphi(0|\theta) \mathrm{d}F(\theta) = 0$.

Note that these definitions allow for an HR policy to exhibit both over-retention and under-retention.

Toward proving Theorem 1, we first characterise the case in which efficient retention is possible.

**Lemma 5.** *An optimal HR policy exists with efficient retention if and only if the efficient retention threshold* $\overline{\theta}_0$ *satisfies* $V_2'(\mathbb{E}[\theta|\theta \geq \overline{\theta}_0]) \geq V_1'(\overline{\theta}_0)$ *and* $V_2(\mathbb{E}[\theta|\theta \geq \overline{\theta}_0]) = V_2'(\mathbb{E}[\theta|\theta \geq \overline{\theta}_0])(\mathbb{E}[\theta|\theta \geq \overline{\theta}_0] - \overline{\theta}_0)$.

[41] As before, $\theta_1^b$ is given via $\mathbb{E}[\theta|\theta < \theta_1^b] = E_1^b$

*Proof.* Let $G$ be a solution to the firm's information design problem and suppose it can be induced by an HR policy that exhibits efficient retention. By Fact 1, the support function $p$ of $G$ equals zero on $[0, \overline{\theta}_0]$. Since $p \geq V$, $p$ must have a kink with a right-derivative greater than $V_1'(\overline{\theta}_0)$. Since $\text{supp}(G) \subseteq [\overline{\theta}_0, 1]$, by Fact 2(ii), $p$ must be tangent to $V$ at some $E \in (\overline{\theta}_0, 1]$. Because $V_1$ is concave, it must be that $E \in (\overline{\theta}_1, 1]$. The concavity of $V_2$ means that such a point $E$, if it exists, must be unique. Finally, observe that condition $V_2(\mathbb{E}[\theta|\theta \geq \overline{\theta}_0]) = V_2'(\mathbb{E}[\theta|\theta \geq \overline{\theta}_0])(\mathbb{E}[\theta|\theta \geq \overline{\theta}_0] - \overline{\theta}_0)$ ensures $E = \mathbb{E}[\theta|\theta \geq \overline{\theta}_0]$. That the converse also holds follows from Fact 1. ■

Let us now prove the key lemma that delivers 1.

**Lemma 6.** *No optimal HR policy exhibits under-retention.*

*Proof.* Let $\varphi$ be an optimal HR policy that exhibits under-retention. Then, there exists a subset $S \subseteq [\overline{\theta}_0, 1]$ who are let go with strictly positive probability. In the associated information design game, consider the $G$ that is induced by $\varphi$. Suppose we modify $G$ by fully revealing the types of workers in $S$ whenever they are to be let go. Since $V(\theta) > 0$ for all $\theta \in S$, the firm can assign them to positions that maximise their output to obtain a strictly positive payoff, rather than a payoff of zero from letting go of them. Moreover, the firm's profits from retained workers are unaffected by this modification of $G$. Thus, the firm's payoff must be strictly higher with this modification. By Lemma 1, the firm can attain at least this payoff from an HR policy, which contradicts the optimality of $\varphi$. ■

We now prove that any optimal HR policy (monotone or otherwise) exhibits over-retention.

**Theorem 1.** *An optimal HR policy exists with efficient retention if and only if the efficient retention threshold $\overline{\theta}_0$ satisfies*

$$V_2'(\mathbb{E}[\theta|\theta \geq \overline{\theta}_0]) \geq V_1'(\overline{\theta}_0) \tag{30}$$

*and*

$$V_2(\mathbb{E}[\theta|\theta \geq \overline{\theta}_0]) = V_2'(\mathbb{E}[\theta|\theta \geq \overline{\theta}_0])(\mathbb{E}[\theta|\theta \geq \overline{\theta}_0] - \overline{\theta}_0). \tag{31}$$

*Moreover, under-retention is never optimal. Therefore, if conditions* (30) *and* (31) *do not hold, then optimal HR policy exhibits over-retention.*

*Proof.* The first two statements are simply Lemma 5 and 6, respectively. If conditions (30) and (31) do not hold, then no HR policy with efficient retention is optimal. That is, any optimal HR policy exhibits inefficient retention. It must also exhibit no under-retention by Lemma 6. By definitions of efficient retention and under-retention, it follows that any optimal HR policy must exhibit over-retention. ■

### A.4.2 Promotion

We say that an HR policy $\varphi : \Theta \to \Delta(\{0,1,2\})$ exhibits:

- *over-promotion* if $\int_0^{\overline{\theta}_1} \varphi(2|\theta)\mathrm{d}F(\theta) > 0$;
- *under-promotion* if $\int_{\overline{\theta}_1}^1 \varphi(2|\theta)\mathrm{d}F(\theta) < 1 - F(\overline{\theta}_1)$;
- *efficient promotion* if $\int_0^{\overline{\theta}_1} \varphi(2|\theta)\mathrm{d}F(\theta) = 0$ and $\int_{\overline{\theta}_1}^1 \varphi(2|\theta)\mathrm{d}F(\theta) = 1 - F(\overline{\theta}_1)$

Note again that these definitions allow for an HR policy to exhibit both over-retention and under-retention. Recall that any feasible $G$ can be induced by a monotone HR policy and that, if $G$ is pooling (vs bi-pooling), then it can be induced by a deterministic monotone HR policy.

We first note that if the degenerate distribution of posterior means (i.e., $= F$) is optimal, then there is over-promotion (resp. under-promotion) if the firm assigns all workers to $j = 2$ (resp. $j < 2$) under the prior belief. Toward proving theorem Theorem 2, let us consider the case in which $p_2^o > p_1^o$.

**Proposition 7.** *Suppose $p_2^o > p_1^o$. Suppose no optimal monotone HR policy is deterministic. Then, optimal monotone HR policies can exhibit only over-promotion, only under-promotion, or both over- and under-promotion.*

*Proof.* If no optimal monotone HR policy is deterministic, then optimal $G$ must bi-pool the interval $[\theta_0^b, 1]$ to $E[\theta|\theta_0^b < \theta < \theta_1^b]$ and $E_2 \in (\mathbb{E}[\theta|\theta > \theta_0^b], E[\theta|\theta \geq \theta_1^b])$. A monotone HR policy that induces $G$ is one that: (i) assigns $j = 2$ with probability one to workers whose types are in $[\theta_1^b, 1]$, and (ii) assigns $j = 1$ with some probability $\beta$ to workers whose types are in $[\theta_0^b, \theta_1^b]$ and assigns $j = 2$ with complementary probability. If $\overline{\theta}_1 \geq \theta_1^b$, then this monotone HR policy exhibits only over-promotion. However, if $\overline{\theta}_1 < \theta_1^b$, this monotone HR policy exhibits both over-promotion (because it sometimes assigns $j = 2$ to workers whose types are in $[\theta_0^b, \overline{\theta}_1])$ and under-promotion (because it sometimes assigns $j = 1$ to workers whose types are in $[\overline{\theta}_1, \theta_1^b]$). To show that monotone HR policy can exhibit only under-promotion, let $\tilde{t}_1^b$ be such that $\mathbb{E}[\theta|\theta \geq \tilde{t}_1^b] = E_2^b$. If $\mathbb{E}[\theta|\theta_0^b < \theta < \tilde{\theta}_1^b] < E_1^b$, following the same argument as in the proof of Lemma 2 in Arieli et al. (2023), we can induce a bi-pooling distribution by a monotone HR policy that: (i) assigns $j = 1$ with probability one to workers whose types are in $[\theta_0^b, \tilde{\theta}_1^b]$, and (ii) assigns $j = 1$ with some probability $\beta$ to workers whose types are in $[\tilde{\theta}_1^b, 1]$ and assigns $j = 2$ with complementary probability. In this case, if $\overline{\theta}_1 \leq \tilde{t}_0^b$, there is only under-promotion. ■

We now focus on the case in which an optimal HR policy is deterministic. To state the result formally, let $\overline{E}_2 \coloneqq \mathbb{E}[\theta|\theta \geq \overline{\theta}_1]$. When $\overline{E}_2 < E_2^b$, define $\overline{E}_1$ be such that

$$V\left(\overline{E}_1\right) + V'\left(\overline{E}_1\right)\left(\overline{\theta}_1 - \overline{E}_1\right) = V\left(\overline{E}_2\right) + V'\left(\overline{E}_2\right)\left(\overline{\theta}_1 - \overline{E}_2\right). \tag{32}$$

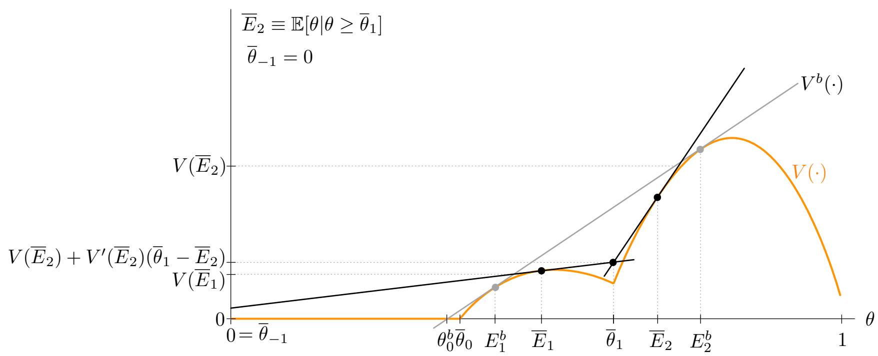
and let

$$\overline{\theta}_{-1} = \begin{cases} 0 & \text{if } V\left(\overline{E}_1\right) \geq V'\left(\overline{E}_1\right)\overline{E}_1, \\ \overline{E}_1 - \frac{V\left(\overline{E}_1\right)}{V'\left(\overline{E}_1\right)} & \text{otherwise.} \end{cases} \tag{33}$$

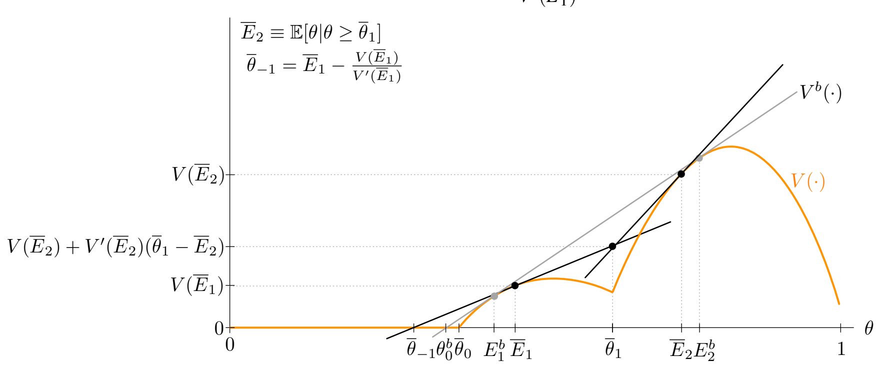
As Figure 13 shows, given $\overline{E}_2 < E_2^b$, $\overline{E}_1$ is the point such that the tangent to $V$ at $\overline{E}_1$ intersects the tangent of $V$ at $\overline{E}_2$ at $(\overline{\theta}_1, V(\overline{E}_2) + V'(\overline{E}_2)(\overline{\theta}_1 - \overline{E}_2))$. Figures 13a and 13b show the cases in which $\overline{\theta}_{-1} = 0$ and $\overline{\theta}_{-1} > 0$, respectively.

Figure 13: Over- vs under-promotion.

(a) Case: $\overline{\theta}_{-1} = 0$.




(b) Case: $\overline{\theta}_{-1} = \overline{E}_1 - \frac{V(\overline{E}_1)}{V'(\overline{E}_1)}$.




Recall that we say that a deterministic monotone HR policy exhibits over-promotion (resp. under-promotion) if the associated promotion cutoff is below (resp. above) the efficient promotion threshold.

**Proposition 8.** *Suppose optimal monotone HR policy, $\varphi^*$, is deterministic.*

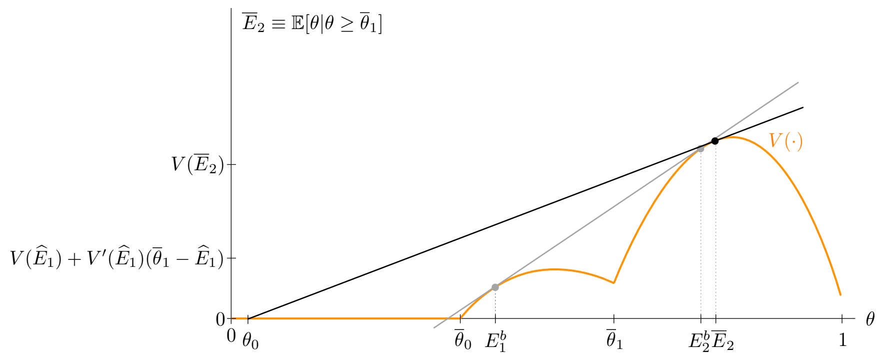
*(i) Suppose $V$ has a bi-tangent. Then, $\varphi^*$ exhibits over-promotion if $E_2^b < \mathbb{E}[\theta|\theta \geq \theta_1^b]$ and either (a) $\overline{E}_2 > E_2^b$ or (b) $\overline{E}_2 < E_2^b$ and $\mathbb{E}[\theta|\overline{\theta}_{-1} < \theta < \overline{\theta}_1] > \overline{E}_1$; otherwise, $\varphi^*$ exhibits under-promotion.*

*(ii) Suppose no bi-tangent of $V$ exists. Then, $\varphi^*$ exhibits over-promotion if $\mathbb{E}[\theta|\overline{\theta}_{-1} < \theta < \overline{\theta}_1] > \overline{E}_1$; otherwise $\varphi^*$ exhibits under-promotion.*

*Proof.* If $V$ has a bi-tangent and a deterministic monotone HR policy is optimal, it must be the case that (25) does not hold and $E_2^b > \mathbb{E}[\theta|\theta \geq \theta_1^b]$.

Suppose $\overline{E}_2 > E_2^b$. Toward a contradiction, suppose $\theta_1^* \geq \overline{\theta}_1$. By definition of $E_2^b$ as (one of) the bi-tangency points of $V$ and the fact that $V|_{[\overline{\theta}_0, \overline{\theta}_1]}$ and $V|_{[\overline{\theta}_1, 1]}$ are concave, the tangent of $V$ at $\overline{E}_2 \in (\overline{\theta}_1, 1)$ must lie strictly above $V$ on $[\overline{\theta}_0, \overline{\theta}_1]$ and equal zero at some $\theta_0 < \overline{\theta}_0$ as can be seen from the figure below.

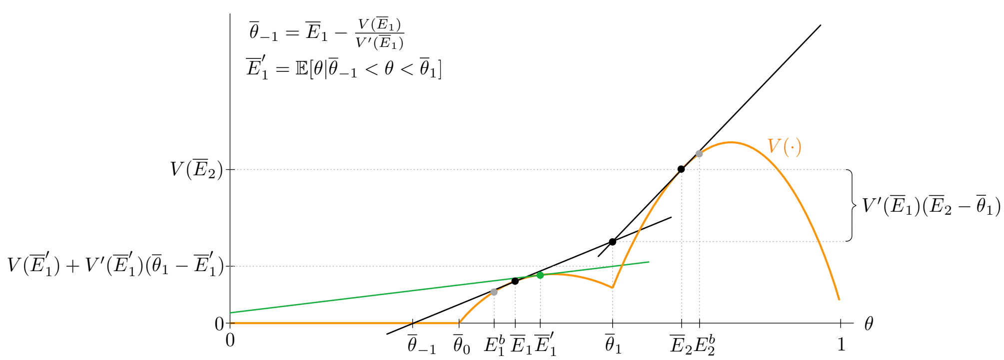
Figure 14: $\overline{E}_2 > E_2^b$: Over-promotion.




Thus, any $G$ supported by $p$ such that $p|_{[\overline{\theta}_1,1]}$ is given by the tangent on $V$ at $\overline{E}_2$ must pool the mass on $[\theta_0, 1]$ at $\mathbb{E}[\theta|\theta \geq \theta_0]$. But because $\mathbb{E}[\theta|\theta \geq \theta_0] < \overline{E}_2 \equiv \mathbb{E}[\theta|\theta \geq \overline{\theta}_1]$, such a support function is not a tangent to $V$ at $\mathbb{E}[\theta|\theta \geq \theta_0]$; i.e., no such $G$ is optimal. Hence, we must have $\theta_1^* < \overline{\theta}_1$.

Now suppose $\overline{E}_2 < E_2^b$. Consider first the case $\overline{E}_1' := \mathbb{E}[\theta|\overline{\theta}_{-1} < \theta < \overline{\theta}_1] > \overline{E}_1$. Then, concavity of $V|_{[\overline{\theta}_0,\overline{\theta}_1]}$ implies that the tangent to $V$ at $\overline{E}_1'$ evaluated at $\overline{\theta}_1$ must lie strictly below the intersection of the tangents to $V$ at $\overline{E}_1$ and $\overline{E}_2$as can be seen in the figure below.

Figure 15: $\overline{E}_2 < E_2^b$ and $\overline{E}_1' > \overline{E}_1$: Over-promotion.

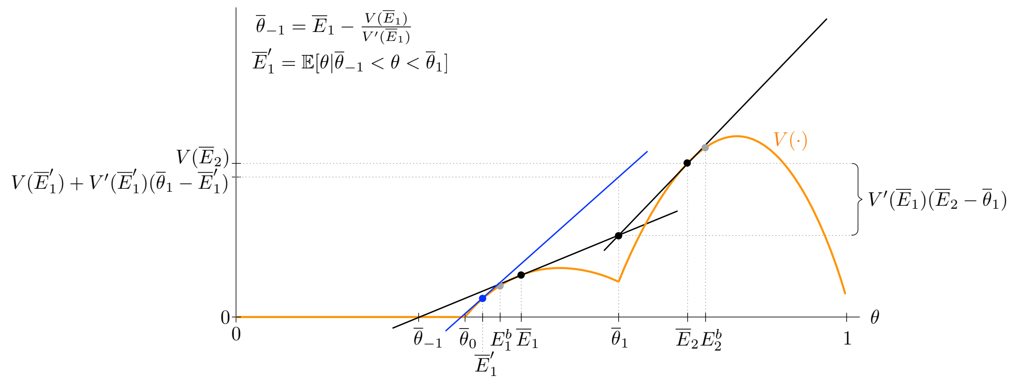



But this means that

$$\begin{aligned}&\frac{\mathrm{d}\widehat{V}\left(\overline{\theta}_{-1}, \overline{\theta}_1, \overline{E}_2\right)}{\mathrm{d}\theta_1} \\ &= f\left(\overline{\theta}_1\right)\left(V\left(\overline{E}_1'\right) + V'\left(\overline{E}_1'\right)\left(\overline{\theta}_1 - \overline{E}_1'\right) + V'\left(\overline{E}_2\right)\left(\overline{E}_2 - \overline{\theta}_1\right) - V\left(\overline{E}_2\right)\right) \qquad (34)\end{aligned}$$

is strictly negative and so $\theta_1^* < \overline{\theta}_1$; i.e., there is over-promotion. If $\overline{E}_1' < \overline{E}_1$, analogous arguments yields that $\theta_1^* > \overline{\theta}_1$ (see figure below) so that there is under-promotion.

Figure 16: $\overline{E}_2 < E_2^b$ and $\overline{E}_1' < \overline{E}_1$: Under-promotion.




The case for when $E_2^b < \mathbb{E}[\theta|\theta \geq \theta_1^b]$ follows immediately from Theorem 3.

For the case in which $V$ does not have a bi-tangent, observe that the argument for why there is over- or under-promotion when $\overline{E}_2 < E_2^b$ did not rely on bi-tangents. Hence, the same argument still goes through even when $V$ does not have a bi-tangent. ■

We note that the condition $\overline{E}_2 > E_2^b$ holds if there is a sufficient mass of workers whose types are above $E_2^b$. The condition $\mathbb{E}[\theta|\overline{\theta}_{-1} < \theta < \overline{\theta}_1] > \overline{E}_1$ holds if there is a sufficient mass of workers whose types are in $[\overline{E}_1, \overline{\theta}_1]$.

**Theorem 2.** *An optimal deterministic monotone HR policy exhibits either over-promotion or under-promotion but not both. If no deterministic monotone HR policy is optimal, then an optimal monotone HR policy could exhibit only over-promotion, only under-promotion, or both.*

*Proof.* Case: $p_2^o > p_1^o$. Proposition Proposition 7 deals with the case when no optimal HR policy is optimal, and Proposition Proposition 8 deals with the case when optimal HR policy is deterministic.

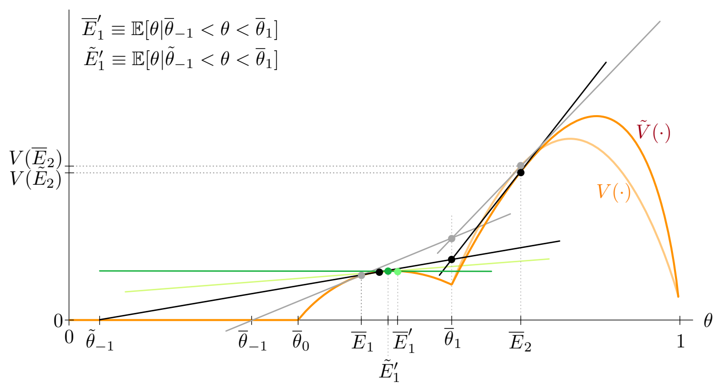
Case: $p_2^o \leq p_1^o$. From Theorem 4, the only interesting case when $p_1^o \geq p_2^o$ is if $\mathbb{E}[\theta] \in [E_1^b, E_2^b]$. In this case, the analysis of whether there is over- or under-promotion is the same as the case in which $p_2^o > p_1^o$ when the firm does not let go of any workers. ■

### A.4.3 Comparative statics on $c$

**Proposition 9.** *Suppose (25) does not hold and $E_2^b > \mathbb{E}[\theta|\theta \geq \theta_1^b]$. Then, changes in worker's productivity in job 2 that do not alter the efficient promotion threshold would not alter the nature of inefficiency with respect to the firm's promotion decision if $\overline{E}_2 < E_2^b$; if $\overline{E}_2 > E_2^b$ instead, it is possible that firm may switch from over-promotion to under-promotion.*

*Proof.* Consider a change $\tilde{c}$ such that $\tilde{c}$ is strictly increasing and strictly convex, $\tilde{c}|_{[0,\overline{\theta}_1]} = c|_{[0,\overline{\theta}_1]}$, $\tilde{c}(1) = c(1)$ and $\tilde{c}'(\overline{E}_2) > c'(\overline{E}_2)$. Let $\tilde{V} := [\max_{j\in\{1,2\}} v_j(\cdot) - \tilde{c}(\cdot)]_+$.

Figure 17: Comparative statics on $c$.




Then, $\tilde{V}'(\overline{E}_2) < V'(\overline{E}_2)$ so that the tangent to $\tilde{V}$ at $\overline{E}_2$ cuts $\overline{\theta}_2$ below the tangent to $V$ at $\overline{E}_2$ (see figure above). Given concavity of $V$ on $[\overline{\theta}_0, \overline{\theta}_1]$, $\tilde{E}_1 > \overline{E}_1$, where $\tilde{E}_1$ solves (32) (with $V$ replaced with $\tilde{V}$). In turn, $\tilde{\theta}_1$ defined via (33) (with $\overline{E}_1$ replaced with $\tilde{E}_1$) must be such that $\tilde{\theta}_{-1} \leq \overline{\theta}_{-1}$. Therefore, $\tilde{E}_1' := \mathbb{E}[\theta|\tilde{\theta}_{-1} < \theta < \overline{\theta}_1 < \overline{E}_1' := \mathbb{E}[\theta|\overline{\theta}_{-1} < \theta < \overline{\theta}_1]$. Hence, the derivative (34) must increase (see again figure). That is, under-promotion becomes more likely. ∎

### A.4.4 Comparative statics on $v_2$

Recall that $\overline{\theta}_1 \equiv \frac{s_1 - s_2}{k_2 - k_1}$ and consider an increase in $k_2$ to $\tilde{k}_2 = k_2 + \varepsilon$ for some $\varepsilon > 0$. Then, to ensure that the efficient promotion threshold remains unchanged, we need

$$\overline{\theta}_1 = \frac{s_1 - \tilde{s}_2}{\tilde{k}_2 - k_1} \Leftrightarrow \tilde{s}_2 = s_1 - \overline{\theta}_1 (k_2 - k_1) - \varepsilon \overline{\theta}_1 = s_2 - \varepsilon \overline{\theta}_1.$$

Since the efficient promotion threshold remains unchanged, $\overline{E}_1$, $\overline{E}_2$ and $\overline{\theta}$ all remain unchanged. Let $\tilde{V}$ denote the firm's value function under $(\tilde{k}_2, \tilde{s}_2)$:

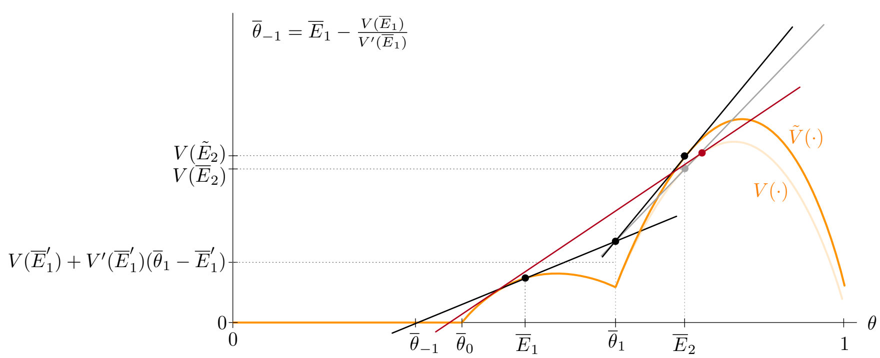
$$\tilde{V}(\theta) = \begin{cases} V(\theta) & \text{if } \theta \leq \overline{\theta}_1, \\ \tilde{s}_2 + \tilde{k}_2\theta - c(\theta) = v_2(\theta) - c(\theta) + \varepsilon\left(\theta - \overline{\theta}_1\right) & \text{if } \theta > \overline{\theta}_1. \end{cases}$$

Note that the tangent to $\tilde{V}$ at $\overline{E}_2$ is given by

$$\begin{aligned} \tilde{V}'\left(\overline{E}_2\right)\left(\theta - \overline{E}_2\right) + \tilde{V}\left(\overline{E}_2\right) &= \left(\tilde{k}_2 - c'\left(\overline{E}_2\right)\right)\left(\theta - \overline{E}_2\right) + \tilde{v}_2\left(\overline{E}_2\right) - c\left(\overline{E}_2\right) \\ &= \left(k_2 - c'\left(\overline{E}_2\right)\right)\left(\theta - \overline{E}_2\right) + \varepsilon\left(\theta - \overline{E}_2\right) \\ &\quad + v_2\left(\overline{E}_2\right) - c\left(\overline{E}_2\right) + \varepsilon\left(\overline{E}_2 - \overline{\theta}_1\right) \\ &= V'\left(\overline{E}_2\right)\left(\theta - \overline{E}_2\right) + V\left(\overline{E}_2\right) + \varepsilon\left(\theta - \overline{\theta}_1\right). \end{aligned}$$

Hence, the tangent to $\tilde{V}$ at $\overline{E}_2$ and the tangent to $V$ at $\overline{E}_2$ intersect exactly at $\overline{\theta}_1$. But then the derivative (34) does not change. This is shown in figure below.

Figure 18: Comparative statics on $v_2$.




Note further that since $\tilde{V}'|_{[\overline{\theta}_1,1]} > \tilde{V}|_{[\overline{\theta}_1,1]}$, the bi-tangency points of $\tilde{V}$ (if they exists), denoted $\tilde{E}_1^b$ and $\tilde{E}_2^b$, satisfy $\tilde{E}_1^b < E_1^b$ and $\tilde{E}_2^b > E_2^b$. Hence, if $\overline{E}_2 < E_2^b$, then $\overline{E}_2 < \tilde{E}_2^b$ so that the nature of inefficiency remains unchanged. However, if $\overline{E}_2 > E_2^b$, it is possible that $\overline{E}_2 < \tilde{E}_2^b$ so that it would result in under-promotion if $\mathbb{E}[\theta | \hat{\theta}_0 < \theta < \overline{\theta}_1] < \overline{E}_1$.

### A.4.5 Comparative statics on $F$

The following is a well-known result that tells us that if two CDFS $F$ and $G$ are ordered according to their likelihood ratios, then their conditional distributions are first-order stochastically ordered.[42]

**Fact 3.** *Let $F$ and $H$ be two continuous CDF with support* $[0,1]$. *Then,*

$$F(\cdot|a < X < b) \geq H(\cdot|a < X < b)\ \forall a,b \in [0,1] : b > a$$

*if and only if* $\frac{h(\cdot)}{f(\cdot)}$ *is increasing.*

Suppose $F$ and $H$ are two prior distributions of workers' types and that $\frac{h(\cdot)}{f(\cdot)}$ is increasing (i.e., $H$ dominates $F$ in the likelihood ratio order), then $\mathbb{E}_H[\theta|\theta_0 < \theta < \theta_1] \geq \mathbb{E}_F[\theta|\theta_0 < \theta < \theta_1]$ for any $\theta_0, \theta_1 \in [0,1]$ with $\theta_1 > \theta_0$.

**Proposition 10.** *Let $F$ and $H$ be two CDFs on $\theta$ and that $H$ dominates $F$ in the likelihood ratio order. Then, under-promotion is more likely under $F$ than under $H$.*

*Proof.* By Fact 3, $\mathbb{E}_H[\theta|\theta \geq \overline{\theta}_1] \geq \mathbb{E}_F[\theta|\theta \geq \overline{\theta}_1]$ and because $V$ is concave on $[\overline{\theta}_1, \overline{\theta}_2]$,

$$V\left(\overline{E}_2(H)\right) + V'\left(\overline{E}_2(H)\right)\left(\overline{\theta}_1 - \overline{E}_2(H)\right) > V\left(\overline{E}_2(F)\right) + V'\left(\overline{E}_2(F)\right)\left(\overline{\theta}_1 - \overline{E}_2(F)\right).$$

It must also be the case that $\mathbb{E}_H[\theta|\theta_0 < \theta < \overline{\theta}_1] \geq \mathbb{E}_H[\theta|\theta_0 < \theta < \overline{\theta}_1]$ for any $\theta_0 \in [0, \overline{\theta}_0]$. By concavity of $V$ on $[\overline{\theta}_0, \overline{\theta}_1]$,

$$\mathbb{E}_H[\theta|\overline{\theta}_{-1} < \theta < \overline{\theta}_1] \leq \mathbb{E}_F[\theta|\overline{\theta}_{-1} < \theta < \overline{\theta}_1]$$

and so by the same argument as in Proposition 9, $\mathbb{E}[\theta|\overline{\theta}_{-1} < \theta < \overline{\theta}_1] - \overline{E}_1$ must be smaller under $F$ than under $H$. ■

## A.5 Partial commitment

**Lemma 7.** *Any HR policy that solves (2) must be deterministic and monotone.*

*Proof.* Fix an HR policy, $\varphi$. The result follows if we can prove the claim: there exists measurable subsets $S, S' \subseteq \theta$ such that (i) $F(S) = F(S') = \ell > 0$; (ii) $\sup S' \leq \widehat{\theta} := \inf S$; and (iii) $\varphi(\cdot|\theta) = \delta_{j'}$ for all $\theta \in S$ and $\varphi(\cdot|\theta') = \delta_j$ for all $\theta' \in S'$ with $j > j'$; then the firm's direct payoff is weakly greater from $\widehat{\varphi}$ that equals $\varphi$ except that $\varphi(\cdot|\theta) = \delta_j$ for all $\theta \in S$ and $\varphi(\cdot|\theta') = \delta_{j'}$ for all $\theta' \in S'$. Condition (i) means that $S$ and $S'$ have the same size, (ii) means that $S'$ consists of workers whose types are below that of any worker in $S$,

[42] See, for example, Shaked and Shanthikumar (2007), Theorem 1.C.5.

and (iii) means that $\varphi$ assigns a higher position to lower types of workers who are in $S'$. The claim says that the firm is better off by assigning lower types of workers to a lower position. To prove the claim, observe that $v_j(\theta)$ has increasing differences in $(j,\theta)$ since, for any $\theta,\theta' \in \theta$ such that $\theta' > \theta$, we have

$$v_2\left(\theta'\right) - v_2\left(\theta\right) = k_2\left(\theta' - \theta\right) > k_1\left(\theta' - \theta\right) = v_1\left(\theta'\right) - v_1\left(\theta\right) > 0 = v_0\left(\theta'\right) - v_0\left(\theta\right).$$

In particular, this means that, for any $j > j'$,

$$v_j\left(\theta\right) - v_{j'}\left(\theta\right) > v_j\left(\theta'\right) - v_{j'}\left(\theta'\right).$$

This observation, together with conditions (i) and (ii), means that

$$\begin{aligned}
\int_S \left[v_j\left(\theta\right) - v_{j'}\left(\theta\right)\right] \mathrm{d}F\left(\theta\right) &\geq \int_S \left[v_j\left(\widehat{\theta}\right) - v_{j'}\left(\widehat{\theta}\right)\right] \mathrm{d}F\left(\theta\right) \\
&= \left[v_j\left(\widehat{\theta}\right) - v_{j'}\left(\widehat{\theta}\right)\right] \ell \\
&= \int_{S'} \left[v_j\left(\widehat{\theta}\right) - v_{j'}\left(\widehat{\theta}\right)\right] \mathrm{d}F\left(\theta\right) \\
&\geq \int_{S'} \left[v_j\left(\theta\right) - v_{j'}\left(\theta\right)\right] \mathrm{d}F\left(\theta\right).
\end{aligned}$$

In fact, what we have shown is

$$\int_S v_j\left(\theta\right) \mathrm{d}F\left(\theta\right) + \int_{S'} v_{j'}\left(\theta\right) \mathrm{d}F\left(\theta\right) \geq \int_S v_{j'}\left(\theta\right) \mathrm{d}F\left(\theta\right) + \int_{S'} v_j\left(\theta\right) \mathrm{d}F\left(\theta\right);$$

i.e., the payoff from swapping assignments is greater. To see why the claim implies the result, observe that if $\varphi$ is not monotone, then we can find $S,S' \subseteq \theta$ that satisfies the conditions of the claim. Now suppose that $\varphi$ is monotone and involves randomisation, then there will exist measures of workers who are sometimes assigned to a higher position and otherwise assigned to a lower position. However, the firm can improve its payoff by assigning the workers to a higher position without changing the marginal distribution via the claim. ■