Dynamic Evaluation Design*

Alex Smolin

October 28, 2020

Abstract

A principal owns a firm, hires an agent of uncertain productivity, and designs a dynamic policy for evaluating his performance. The agent observes ongoing evaluations and decides when to quit. When not quitting, the agent is paid a wage that is linear in his perceived productivity; the principal claims the residual performance. After quitting, the players secure fixed outside options. I show that equilibrium evaluation policies are Pareto efficient. In a minimally informative equilibrium, for a broad class of performance technologies, the agent's wage deterministically grows with tenure. My analysis suggests that endogenous performance evaluation plays an important role in shaping careers in organizations.

Keywords: evaluation, information design, career concerns, bandit experimentation, downward wage rigidity, up-or-out, internal labor markets

JEL Codes: C72, D82, D83, M52

*Toulouse School of Economics, 1 Esplanade de l'Université, 31080 Toulouse Cedex 06, France, alexey.v.smolin@gmail.com. This paper builds on a chapter of my Ph.D. dissertation written at Yale University that circulated as "Optimal Feedback Design" in November, 2015; I thank Dirk Bergemann, Johannes Hörner, and Larry Samuelson for support and guidance. I thank the coeditor, Michael Ostrovsky, and three anonymous referees for many productive suggestions. I am grateful to Florian Ederer, Sergei Izmalkov, Emir Kamenica, Anton Kolotilin, Chiara Margaria, Benny Moldovanu, Pauli Murto, Anne-Katrin Roesler, and Áron Tóbiás for insightful conversations. Finally, I am thankful to participants of research seminars at Berlin, Bocconi, Bonn, Cornell Johnson, HSE (Moscow), Indiana Kelley, MPI Bonn, and Yale, as well as at the 5th World Congress of the Game Theory Society and the 2016 North American Summer Meeting of the Econometric Society. I acknowledge funding from ANR under grant ANR-17-EURE-0010 (Investissements d'Avenir program).

1 Introduction

Performance evaluation is an important part of organizational life. Although much evaluation is informal, most organizations have formal evaluation policies designed to collect and distribute performance information to employees.[1] As communication and information technologies advance, many companies find it easier to provide more evaluation. As a recent example, in August 2015 General Electric (GE) announced an ongoing shift from its legacy system of annual performance reviews to more frequent conversations between managers and employees via an online application.[2] In this way, GE joined other high-profile companies such as Microsoft, Accenture, and Adobe in a move towards more frequent, exhaustive, and real-time evaluation. However, whenever adopting new evaluation policies, companies should ask: What is their effect on overall performance? Would other evaluation policies perform better? Ultimately, which evaluation policy is the best for the company?

In this paper, I develop a framework to analyze the design of evaluation policies. I consider a principal who owns a firm and hires an agent to work over time. The agent's productivity, his type, is initially uncertain to both players but affects the agent's ongoing performance via a general production technology. The agent's performance is not directly observed but can be revealed through evaluations. While at the firm, the agent's wage is linear in his expected productivity; the principal claims the residual performance. In every period, the agent evaluates his career prospects and decides whether to quit. When the agent quits, the players secure exogenous outside options. Both players are risk neutral and discount the future at the same rate.

The principal designs and adopts a dynamic evaluation policy. The policy is a sequence of statistical experiments that are informative of past performance. The experiments can vary in what and when performance is assessed. The evaluation is costless but its design should take into account the agent's incentives. On the one hand, the promise of future evaluations motivates the agent to stay at the firm and learn whether he is able to perform well. On the other hand, an evaluation may turn out negative and persuade the agent to quit.

In Section 4, I investigate equilibrium evaluation policies by developing a novel efficiency argument. First, I show that the design problem can be viewed as a dynamic persuasion problem in a bandit experimentation setting. It allows me to apply the revelation principle and characterize the set of feasible payoffs that can be possibly achieved in the relationship.

[1] According to Murphy and Cleveland (1995), between 74% and 89% of business organizations had formal performance appraisal policies by 1995.

[2] "GE's Real-Time Performance Development," *Harvard Business Review*, August 12, 2015, https://hbr.org/2015/08/ges-real-time-performance-development.

Second, I study the set of implementable payoffs—the payoffs that can be achieved by some evaluation policy and the agent's best response to it. I show that this set includes all Pareto-efficient payoffs that deliver the agent at least his safe option, and I conclude that any equilibrium evaluation policy is efficient. Under a minimally informative equilibrium policy, the agent is informed about whether he would quit if he could fully observe past performance but had a lower outside option.

In Section 5, I study the effects of optimal evaluation on the agent's career and wage dynamics. I observe several qualitative properties that hold in the minimally informative equilibrium. First, the agent's wage is a deterministic function of tenure. Second, the agent's continuation value grows with tenure so that the agent becomes more optimistic about his prospects the longer he remains at the firm. Third, for a broad class of performance technologies, the agent's wage also grows with tenure. The increase reflects the ongoing positive selection and the corresponding growth of expected productivity. The shape of the wage profile depends on the performance technology. If the technology is coarse, such that performance comes as a stream of infrequent successes, then the wage increases at the revision dates, which are spaced sparsely over the agent's career. In contrast, if the technology is detailed, such that performance can always reveal the agent's incompetence, then performance is constantly monitored, and the wage gradually grows in time.

In Section 6, I study the joint design of wage contract parameters and an evaluation policy. I show that once the principal can control both the information flow and monetary incentives, she is able to extract the full surplus from the relationship: in equilibrium, the joint surplus is maximized, and the agent is left with no rents. The principal can achieve this by offering a fixed-wage contract and, in many cases, by offering a pure-bonus contract.

I discuss the findings in Section 7. First, my analysis suggests that endogenous evaluation policies may be important in explaining economic dynamics commonly observed within firms. These include a lack of wage variation within same-tenure cohorts, downward wage rigidity, and up-or-out contracts. My results further speak in favor of retrospective evaluation and provide a rationale for rating compression and leniency bias as techniques to maintain workforce morale. Second, I highlight the important commitment role that human resource (HR) departments may play in implementing optimal evaluation policies. Third, I show that the equilibrium evaluation policy is robust to a multitude of performance leaks. Finally, I discuss how the framework can be extended to incorporate advisory feedback.

Related Literature My paper contributes to the literature on dynamic persuasion and information design built from the static models of Rayo and Segal (2010) and Kamenica and Gentzkow (2011). Orlov (2016) studies the joint design of performance evaluations and

monetary contracts when the agent exerts private effort. Renault, Solan, and Vieille (2017) and Ely (2017) study dynamic persuasion with an exogenous information flow and a myopic agent. Orlov, Skrzypacz, and Zryumov (2020) investigate a setting in which the principal lacks commitment across different periods. Importantly, my paper introduces the efficiency argument that permits the characterization of the optimal information policies by studying Pareto-efficient allocations.[3]

My framework highlights the interplay between career concerns and performance evaluations. It complements the turnover theory of Jovanovic (1979) by endogenizing the information flow. Relatedly, several papers investigate evaluation effects on private efforts in the framework of Holmström (1999). Hansen (2013) studies static incentives and focuses on partition evaluations. Hörner and Lambert (2020) study dynamic incentives with a focus on Gaussian policies.

Similar incentive effects are present in multistage contests and tournaments. Ederer (2010) compares the effectiveness of complete- and no-evaluation policies in two-stage tournaments. Halac, Kartik, and Liu (2016) study the optimal design of general multistage contests and similarly focus on the extreme evaluation policies within each period. Nevertheless, Goltsman and Mukherjee (2011) highlight that the optimal evaluation policies in tournaments are generally partially informative.

Finally, my paper contributes to the literature on dynamic contracts without transfers. Guo (2016) studies dynamic delegation when the agent is privately informed. Hörner and Guo (2015) study dynamic resource allocation when the agent's private information evolves over time. My paper highlights that information control may complement delegation and action control as a powerful management tool.

2 Model

A principal owns a firm and hires an agent. The relationship takes place in consecutive periods $t = 0, 1, 2, \ldots$ At time 0, the agent's *productivity* $\theta \in \Theta \subseteq \mathbb{R}$ is drawn according to a cumulative distribution G_0. The productivity is fixed throughout the relationship and is not directly observed by either principal or agent. The players are symmetrically informed about productivity with the prior expectation of productivity, $\mathbb{E}[\theta]$, being equal to θ_0.

Performance Productivity affects the agent's *performance* at firm $y_t \in Y \subseteq \mathbb{R}$. Conditional on productivity, performance is independently and identically distributed across

[3]This argument was later applied by Ely and Szydlowski (2020) in a setting in which the relevant state deterministically evolves over time, resulting in a predictable reversal of incentives.

periods according to a cumulative distribution F_θ. The collection of distributions $\{F_\theta\}_{\theta\in\Theta}$ defines production capabilities of the firm and is called the (performance) *technology*. I associate productivity with its expected performance: $\mathbb{E}[y_t \mid \theta] = \theta$.

It follows that performance is informative of the agent's productivity: consistently higher performance suggests higher productivity. The overall informativeness and details of the learning process are determined by technology in place. I impose no assumptions on technology in the characterization of equilibrium payoffs in Section 4. I will impose a regularity assumption in Section 5 to establish downward wage rigidity. Performance is not directly observed by either party but can be revealed through evaluations as discussed below.

Strategies The principal can publicly reveal past performance through an *evaluation policy* that she designs. The policy is costless and governs when and what performance information is available. The evaluations are objective; their outcomes cannot be manipulated by the principal. At the same time, I place no restrictions on which evaluations the principal can conduct. That is, she can conduct a complete evaluation, no evaluation, periodic reviews, grade evaluations, and so forth.

Formally, the principal chooses an evaluation policy m among all stochastic processes measurable with respect to past performance and evaluations.[4] The policy can be represented by a sequence of random messages $\{m_t\}_{t=0}^{\infty}$ that are sent to the agent:[5]

$$m_t : Y^{t-1} \times M^{t-1} \to \Delta(M). \tag{1}$$

The message space M is the same in all periods and can be freely chosen by the principal. The exact message labels are irrelevant because their meaning is determined solely by the law of m. Associate a *complete-evaluation policy* $\overline{m}$ with $m_t \equiv y_{t-1}$ and a *no-evaluation policy* $\underline{m}$ with $m_t \equiv \emptyset$. Denote the set of all possible evaluation policies of the form (1) by $\mathcal{M}$.

The concept of an evaluation policy is an extension of Kamenica and Gentzkow (2011)'s static persuasion policy and admits two possible interpretations. First, it can be viewed as a disclosure policy. In this interpretation, the principal constantly monitors performance but is bound to communicate according to the chosen policy. Second, it can be viewed as a sequence of public experiments. In this interpretation, the principal does not directly observe performance but commits to a sequential policy of public tests to inform both players of past

[4]Randomization over several evaluation policies can be represented by a single evaluation policy with a combined evaluation law.

[5]I adopt the convention that for any stochastic process x its time-t realization is denoted by subscript x_t, and the history up to time t, $\{x_s\}_{s\leq t}$, is denoted by superscript x^t. For any set X, a set X^t is the t-fold Cartesian product of X.

performance.

The evaluations may be understood as being conducted by the HR department of a firm. In this case, a realization m_t corresponds to an outcome of a particular evaluation. The evaluation policy m corresponds to the operating rules of the department and specifies which and how past performance is evaluated in any given period. I discuss the implementation details in Section 7.2.

Faced with the evaluation policy, the agent chooses whether and when to quit the firm. He decides based on past evaluations, which he correctly interprets according to Bayes' rule. Quitting is irreversible and ends the game.

Formally, the agent chooses a *quitting time* τ, which is a stopping time measurable with respect to the evaluation policy m:

$$\tau \text{ is a stopping time w.r.t. } m_0, m_1, \dots \tag{2}$$

The quitting time is a random variable. If $\tau \equiv 0$, then the agent quits at time 0 and does not generate any performance. If $\tau \equiv \infty$, then the agent stays at the firm forever, irrespective of past evaluations. Denote the set of all possible quitting times by $\mathcal{T}$.

Payoffs As long as the agent stays at the firm, the principal appropriates the performance outcomes and pays the agent a wage w_t. I assume that the wage is set according to a linear contract:

$$w_t\left(m^t\right) = w^F + \alpha \mathbb{E}\left[y_t \mid m^t\right], \tag{3}$$

with $w^F > 0$ being a fixed base wage and $\alpha \in (0, 1)$ being a bonus rate. Linear contracts are widely used in practice and capture in the simplest form the reputation effects of performance evaluations (see Carroll (2015) and the references therein). The agent wants to receive positive evaluations to be perceived as more productive because, in this case, he will be paid a higher bonus. For now, I will treat w^F and α as exogenously fixed. I study endogenous contracts in Section 6.

The timing within each period t is as follows. First, the worker receives an evaluation m_t. Then, he decides whether to stay at the firm. If he stays, then the output y_t is produced and the agent is paid wage w_t according to (3).

As soon as the agent quits the firm, he secures a total payoff of $V^A \in \mathbb{R}$, and the principal secures $V^P \in \mathbb{R}$. These payoffs are exogenous, commonly known, and fixed throughout the relationship. They can be viewed as the opportunity costs of the players.

Both players are risk neutral and discount the future with a common discount factor δ. For given evaluation policy $m \in \mathcal{M}$ and quitting strategy $\tau \in \mathcal{T}$, the normalized expected

payoffs of the players are:

$$U^P(m,\tau) = \mathbb{E}_{m,\tau}\left[(1-\delta)\sum_{t=0}^{\tau-1}\delta^t(y_t - w_t) + \delta^\tau V^P\right], \tag{4}$$

$$U^A(m,\tau) = \mathbb{E}_{m,\tau}\left[(1-\delta)\sum_{t=0}^{\tau-1}\delta^t w_t + \delta^\tau V^A\right]. \tag{5}$$

Note that evaluation policy plays two roles in the payoffs. First, it shapes the agent's wage. Second, it provides the agent with information that guides his quitting decision.

Equilibrium I study perfect Bayesian equilibria of this game. For a given evaluation policy, the agent chooses a quitting time to maximize his total expected payoff. The principal anticipates the agent's best response and designs the evaluation policy to maximize her expected payoffs.

Definition 1. An evaluation policy m^* and a quitting time τ^* constitute an equilibrium if they solve the problem:

$$\max_{m\in\mathcal{M},\tau\in\mathcal{T}} U^P(m,\tau), \tag{6}$$

$$\text{s.t.} \quad \tau \in \arg\max_{\tau\in\mathcal{T}} U^A(m,\tau). \tag{7}$$

My goal is to characterize an equilibrium evaluation policy and payoffs in this game. This characterization further allows me to study equilibrium wage dynamics. To this end, for given strategies m and τ, let *observed wage* W_t equal w_t if the agent remains at the firm, $\tau > t$, and, to complete the definition, equal to 0 otherwise. The observed wage at time t is a random variable that may take many values because the agent can possibly stay at the firm under a wide range of past evaluations. Define a wage profile as a collection of observed wages at different times $W = \{W_t\}_{t=0}^{\infty}$. The wage profile captures the dynamics of the agent's wage throughout his career at the firm.

3 Binary Example

In this section, I illustrate the workings of the setting by means of a simple example. The agent's performance is binary, low or high, $Y = \left\{y^L, y^H\right\}$. Let $y^L = 0$ and $y^H = 1$, and call y^H a "success." There are two possible types, $\Theta = \left\{\theta^L, \theta^H\right\}$, and each type is equally likely. By the definition of productivity, a type is equal to expected productivity, which in this example coincides with the probability of success, $\Pr\left(y_t = y^H \mid \theta\right) = \theta$. Assume that

only the high type is productive, $\theta^L = 0$ and $\theta^H = 1/2$. Let the payoff structure be $w^F = 0$, $\alpha = 1/2$, $V^A = 1/5$, and $V^P = 0$. Finally, let the discount factor be $\delta \simeq 0.97$.[6]

No Evaluation First, consider the case in which the principal adopts a no-evaluation policy $\underline{m}$. In this case, the same evaluation message is sent irrespective of past performance and thus is completely uninformative. The firm effectively provides no feedback. As a result, the wage is fixed at:

$$w_t \equiv w_\emptyset = w^F + \alpha\theta_0 = \frac{1}{8}.$$

Whenever staying at the firm, the agent receives a flow payoff of $w_\emptyset = 1/8$ and foregoes the opportunity flow of $V^A = 1/5$. Because $w_\emptyset < V^A$, the agent quits at time 0.

Consequently, in the absence of informative evaluations, the wage profile is nil and the firm is effectively not operating.

Complete Evaluation Now, consider the case in which the principal adopts a complete-evaluation policy $\overline{m}$. In this case, each evaluation fully reveals the agent's performance in the last period, and the agent's wage depends on past evaluations. The wage starts at $w_\emptyset = 1/8$. As long as no successes occur, the wage gradually decreases according to Bayes' rule, $w_t^L = \frac{1}{2(1+2^t)}$. If a success occurs, it indicates the the agent is of the high type; the wage jumps to $w^H \equiv 1/4$, and remains there forever. The transition to one of these two wages ensures that the wage is a martingale, the property guaranteed by Bayes' rule.

Faced with these career prospects, the agent optimally quits whenever his wage drops below a cutoff wage $\hat{w}$. The cutoff depends on the discount factor. A higher δ translates into lower $\hat{w}$, because career concerns are more important. For the considered discount factor, the agent works until $\hat{T} = 2$ and continues working if and only if success occurred in the past.

The resulting wage profile is random. Viewing the agent as a representative employee, one out of many independent draws would result in two distinct features. First, there would be cross-sectional variation in employee wages in periods before $\hat{T}$: some employees are proven to be high types, and some still attempt to achieve a success. Second, the wage of a given employee is likely to decrease during his career in the firm (Figure 1).

Equilibrium Evaluation Now, consider a partial evaluation policy that reveals whether the principal's belief falls below a certain cutoff. Given the success technology, this policy is equivalent to a revision policy, $m_T = y^{T-1}$ and $m_t = \emptyset$ for $t \neq T$. The firm provides no evaluations before or after the revision date T at which all past performance is evaluated.

[6]The exact value required for a clean demonstration is $2/1365^{1/10}$.

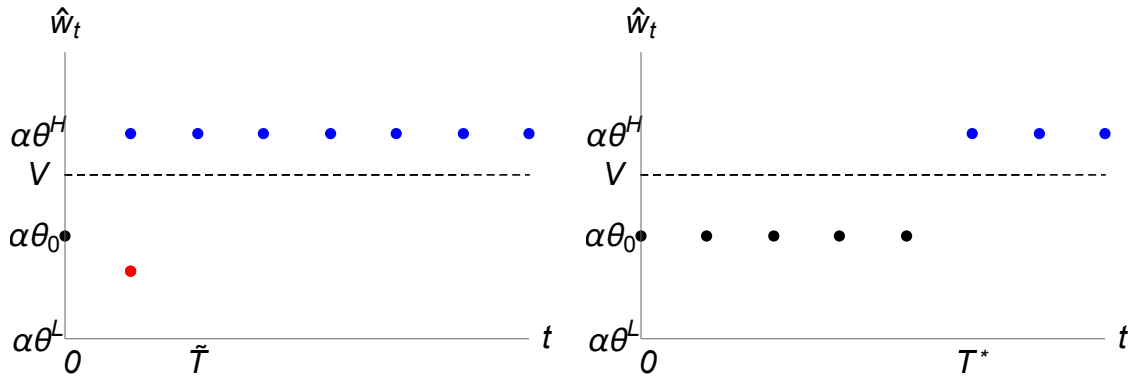

Figure 1: Wage profile under complete evaluation policy (left) and equilibrium evaluation policy (right).

Under this revision policy, the agent's wage prior to revision remains at $w_\emptyset = 1/8$. At T, the full evaluation of past performance is conducted. If a success is revealed, the agent is proven to be of a high type, and his wage jumps up to $w_T^H = 1/4$. If no success is observed, then the productivity expectation drops, as does the wage, to $w_T^L = \frac{1}{2(1+2^T)}$. After the revision time, no evaluations are conducted, so the wage remains constant.

Given these career prospects, as $w_T^H > V^A > w_T^L$, the agent's quitting problem reduces to a binary choice: to either quit at time 0 or remain until the revision time and quit only if no successes are revealed. If the revision date equals the complete evaluation time $T = \hat{T}$, then the agent prefers to remain until the revision because this strategy delivers him the same payoff as the best response to a complete evaluation policy. However, the principal can induce the agent to generate more surplus by postponing the revision until time $T > \hat{T}$. There is a limit on how late the revision can be performed because the agent may prefer to quit at time 0: the maximal revision time can be calculated to be $T^* = 10$.

In fact, the revision policy with revision time T^* is optimal for the principal. Indeed, it delivers payoffs $U^{A*} = V^A = 0.2$ and $U^{P*} \simeq 0.12$. These payoffs are Pareto efficient because the agent never quits when successful. Because the agent can guarantee his safe option by quitting at time 0, the principal cannot achieve payoffs above U^{P*}, and the result follows. The equilibrium wage profile is illustrated in Figure 1.

In what follows, I study the general setting and demonstrate that the main features of this example are general. First, the evaluation policy affects payoffs only through its effect on the quitting time and not on the wage. Second, equilibrium payoffs are Pareto efficient. Third, if productivity is binary or technology is regular, then an equilibrium wage profile is deterministically increasing in tenure. However, under general technology, an optimal evaluation policy needs to be stated in terms of the principal's beliefs and cannot be implemented

via a simple revision policy.

4 Equilibrium Analysis

Equilibrium characterization requires solving a dynamic information disclosure problem. Such problems are known to be difficult due to their inherent multidimensionality. The characterization is further complicated because the information that can be disclosed is generated gradually, and the agent is forward looking. Because of these features, I cannot use the existing techniques of Kremer, Mansour, and Perry (2014) and Ely (2017). Instead, I develop and use an efficiency approach. First, I characterize the set of Pareto-efficient payoffs (Sections 4.1 and 4.2). Second, I provide an upper bound on the principal's equilibrium payoffs. Finally, I demonstrate that the upper bound can be achieved with a particular information policy (Section 4.3).

4.1 Payoff Transformation

To characterize the set of feasible and Pareto-efficient payoffs, it is useful to observe that because both players are risk neutral, the mean-preserving spread of a wage should not affect their payoffs in any period. For example, receiving a fixed wage proportional to the expected performance at the beginning of a period should be payoff equivalent to receiving a bonus proportional to performance at the end of a period.

This intuition can be formalized. Applying the law of iterated expectations and the optional stopping theorem, the player's payoffs can be written solely in terms of performance histories and payoff parameters.

Lemma 1. (Payoff Transformation) *The players' payoffs can be written as:*

$$U^P(m,\tau) = \mathbb{E}_{m,\tau}\left[(1-\delta)\sum_{t=0}^{\tau-1}\delta^t y_t + \delta^\tau \hat{V}^P\right] \times (1-\alpha) - w^F, \tag{8}$$

$$U^A(m,\tau) = \mathbb{E}_{m,\tau}\left[(1-\delta)\sum_{t=0}^{\tau-1}\delta^t y_t + \delta^\tau \hat{V}^A\right] \times \alpha + w^F, \tag{9}$$

with $\hat{V}^P = (V^P + w^F)/(1-\alpha)$ *and* $\hat{V}^A = (V^A - w^F)/\alpha$.

Lemma 1 implies that the current setting is strategically equivalent in the sense of Thompson (1952) to the setting of persuasion in bandit experimentation.[7] In this alternative setting, the agent sequentially pulls the arm of a slot machine and decides when to stop. Pulling the

[7] Bergemann and Välimäki (2008) provide a recent overview of the bandit experimentation literature.

arm generates stochastic rewards for both players. The rewards depend on the machine's type and are not observed by the agent. Stopping delivers the players their safe options. The principal designs what reward information the agent observes to maximize her own payoffs.

The payoff representation (8) and (9) shows that the conflict of interest between the players is captured by their *safe options* $\hat{V}^P$ and $\hat{V}^A$. If these options are equal, then there is no conflict of interest. In this case, the optimal evaluation policy is to provide complete evaluation because it allows the agent to make maximally informed decisions. In contrast, if these options differ, the principal may find it optimal to coarsen the evaluations to steer agent decisions towards her interests.

Assumption 1. (Conflict of Interest) $\hat{V}^P < \hat{V}^A$.

Assumption 1 implies that there is a conflict of interest. As the principal's safe option is lower, she prefers the agent to remain longer at the firm than the agent would ideally prefer. This conflict of interest is typical in the literature on Bayesian persuasion. For example, it holds whenever the agent needs some evaluation to remain at the firm, $V^A > w^F$, and either (i) the firm appropriates a substantial share of output such that α is sufficiently small (cf. Ely and Szydlowski (2020)), or (ii) the principal's total opportunity costs, $V^P + w^F$, are sufficiently low.

4.2 Feasible Payoffs

I proceed with characterizing the set of feasible payoffs. Recall from standard game-theoretic terminology that a pair of strategies $m \in \mathcal{M}$, $\tau \in \mathcal{T}$ *delivers* payoffs $\left(u^A, u^P\right)$ if given the strategies, the payoff of the agent equals u^A and the payoff of the principal equals u^P. In turn, payoffs $\left(u^A, u^P\right)$ are *feasible* if they can be delivered by some players' strategies. The payoffs are (weakly Pareto) *efficient* if there are no strategies that deliver strictly greater payoffs to both players. Denote the set of all feasible payoffs by $\mathcal{F}$:

$$\mathcal{F} \triangleq \left\{ \left(U^A\left(m,\tau\right), U^P\left(m,\tau\right) \right) \mid m_t : Y^{t-1} \times M^{t-1} \to \Delta\left(M\right), \tau \text{ is a stopping time w.r.t. } m \right\}. \tag{10}$$

It is possible to characterize the efficient strategies by a solution to an auxiliary problem. Define the payoffs of a fictitious agent with a virtual safe option $\hat{V}^F$ as:

$$U^F\left(m,\tau,\hat{V}^F\right) = \mathbb{E}_{m,\tau}\left[(1-\delta) \sum_{t=0}^{\tau-1} \delta^t y_t + \delta^\tau \hat{V}^F \right]. \tag{11}$$

Lemma 2. (Efficient Payoffs) *The set of feasible payoffs $\mathcal{F}$ is compact and convex. Efficient payoffs are spanned by $(\overline{m}, \tau')$, where $\tau' \in \arg\max_\tau U^F\left(\overline{m}, \tau, \hat{V}^F\right)$ and $\hat{V}^F \in [\hat{V}^P, \hat{V}^A]$.*

The intuition behind the lemma is as follows. First, any feasible payoff can be obtained with a complete-evaluation policy $\overline{m}$ because this policy provides the most opportunities to respond to performance information. Second, $\mathcal{F}$ is compact as a continuous image of a compact set of quitting strategies; $\mathcal{F}$ is convex because randomizing over two quitting strategies can obtain any convex combination of their corresponding payoffs. As such, by the separating hyperplane theorem, any boundary point of $\mathcal{F}$ is achieved by a quitting strategy that maximizes a linear combination of players' payoffs. For efficient payoffs, the combination weights are positive, and the problem is equivalent to a stopping problem of a fictitious agent with a virtual safe option $\hat{V}^F \in [\hat{V}^P, \hat{V}^A]$. As the weight is shifted from the agent to the principal, $\hat{V}^F$ gradually moves from $\hat{V}^A$ to $\hat{V}^P$.

In fact, these arguments allow us to characterize the whole feasibility set $\mathcal{F}$, not only its efficient payoffs. To do so, it suffices to solve a collection of optimal stopping problems that maximize various linear combinations of players' payoffs. If these stopping problems can be solved, analytically or numerically, then $\mathcal{F}$ can be reconstructed.

In particular, this procedure allows us to depict the feasibility set of the binary example in Section 2 (Figure 2). Points A and D, as well as any payoffs on the segment AD, can be delivered by a quitting time that does not depend on evaluations. Point A is delivered by the agent never quitting, $\tau \equiv \infty$. Point D is delivered by the agent quitting at time zero, $\tau \equiv 0$. The payoffs on segment AD can be delivered by a randomization between these two quitting times.

In contrast, to obtain Pareto-efficient payoffs on the arc AC, the quitting strategy must use performance information. By Lemma 2, those payoffs are spanned by solutions to the virtual problems with outside options $\hat{V}^F \in [\hat{V}^P, \hat{V}^A]$. Point A corresponds to a virtual safe option $\hat{V}^F = \hat{V}^P$, point C corresponds to a virtual safe option $\hat{V}^F = \hat{V}^A$. A somewhat peculiar point E is delivered by a strategy that minimizes the agent's payoff: it prescribes remaining at the firm at low expected productivity and quitting the firm at high expected productivity. Point B maximizes the principal's payoff among all payoffs that deliver at least a safe payoff V^A to the agent. It plays an important role in the equilibrium characterization in the next section.

4.3 Equilibrium Payoffs

The notion of feasibility ignores players' incentives. As shown in the previous section, all feasible payoffs can be delivered by a complete evaluation policy because the quitting time

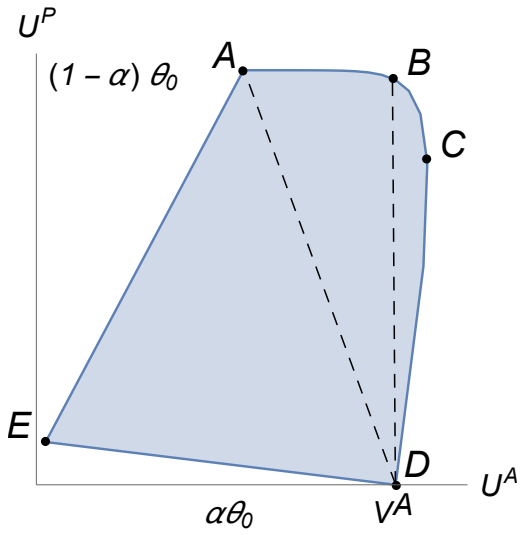

Figure 2: Feasible payoffs in the example from Section 3. $\Theta = \{0, 1/2\}$, $\theta_0 = 0.25$, $Y = \{0, 1\}$, $\Pr(y = 1 \mid \theta) = \theta$, $w^F = 0$, $\alpha = 0.5$, $V^A = 0.2$, $V^P = 0$, $\delta \simeq 0.97$. Computed numerically.

can ignore any additional information. However, under complete evaluation, the agent would act in his own interests and choose a quitting time to maximize his payoff. In Figure 2, it would correspond to point C.

To incorporate the players' incentives, I refer to mechanism-design terminology and say that an evaluation policy $m \in \mathcal{M}$ *implements* payoffs $\left(u^A, u^P\right)$ if the payoffs are delivered by that policy and some agent's best response to it. In turn, payoffs $\left(u^A, u^P\right)$ are *implementable* if they can be implemented by some evaluation policy.

In these terms, a complete-evaluation policy implements the payoffs of point C. However, it is not the only evaluation policy that implements them. Consider a policy that sends only two messages: "stay" and "quit." Let the policy mimic the agent's best response under a complete-evaluation policy; that is, to send a "quit" message only after those performance histories at which the agent himself would quit. If the agent follows the recommendations, then the joint distribution of quitting time and performance will be the same. By Lemma 1, his payoff then equals the payoffs of point C. Because it is his maximal feasible payoff, he cannot do better than follow the recommendations, and so this recommendation policy would implement payoffs C.

Definition 2. An evaluation policy is a *recommendation policy* if it places a positive probability on at most two messages: "stay" and "quit." A recommendation policy is *incentive compatible* if following the recommendations is an agent's best response.

In fact, the agent cannot do better than follow the recommendations of an arbitrary recommendation policy that mimics his best response to some evaluation policy. This fol-

lows from the standard argument of direct mechanisms of Myerson (1986). Consider an arbitrary evaluation policy m and a recommendation policy m' that mimics an agent's best response to m.[8] Because the agent always knows his actions, policy m' provides weakly less information than m. Hence, his payoff cannot be higher than that under m. Following the recommendations delivers the agent the same payoff as under m and hence is a best response to m'.

In other words, the recommendation policies provide minimal information for the agent to make his quitting decision. The principal does not need to provide any information beyond that. Note that Lemma 1 is crucial for this observation because it establishes that the net payoff effect on evaluation policy comes only through its effect on a quitting time and not on a wage.

Lemma 3. (Recommendation Principle) *All implementable payoffs can be implemented by incentive-compatible recommendation policies.*

I proceed with a characterization of implementable efficient payoffs. The agent can secure the payoff V^A by quitting at time 0. Hence, payoffs that yield the agent less than V^A cannot be implemented. In Figure 2, this means that no efficient payoffs to the west of line BD can be implemented. At the same time, point C is implementable by a complete-evaluation policy. In fact, all efficient payoffs on segment BC are also implementable.

Lemma 4. (Implementable Payoffs) *No payoffs with $u^A < V^A$ are implementable. All efficient payoffs with $u^A \geq V^A$ are implementable.*

This lemma builds on the efficient payoff characterization of Lemma 2 and the recommendation principle of Lemma 3. By the payoff characterization, the efficient payoffs can be delivered by strategies (m, τ) that maximize the payoff of an agent with a virtual safe option $\hat{V}^F \in \left[\hat{V}^P, \hat{V}^A\right]$. The agent, if faced with a recommendation policy that mimics τ, is willing to follow recommendations. The argument proceeds as follows. Recommendations to quit are incentive compatible because even the agent with a lower safe option $\hat{V}^F \leq \hat{V}^A$ is willing to follow them—it delivers to him the first-best payoff. Recommendations to stay are incentive compatible because the agent's continuation payoff weakly increases with tenure. Indeed, the agent's continuation payoff after each recommendation to stay can be viewed as a convex combination of three terms: the current wage w_t, the next-period continuation payoff if recommended to quit $\hat{V}^A$, and the next-period continuation payoff if recommended to stay. The agent's payoff must be greater than the first two terms because the agent can

[8]Such policy exists since any quitting time is measurable with respect to m and, hence, with respect to past performance.

guarantee those payoffs by remaining at the firm forever and quitting immediately, respectively. Hence, the payoff is smaller than the third term, which is equivalent to the growth of continuation value with tenure. In other words, while staying at the firm, the agent becomes increasingly optimistic about his prospects.

In equilibrium, the principal chooses an evaluation policy to maximize her payoffs. Equivalently, the principal maximizes her payoff among all implementable payoffs. It is clear from Lemma 4 that she optimally chooses an efficient payoff that either delivers V^A to the agent (point B in Figure 2) or, if implementable, delivers the first-best payoff to the principal.

Theorem 1. (Equilibrium Payoffs) *Equilibrium payoffs exist and are Pareto efficient. Either the agent is left with no rents, $u^{A*} = V^A$, or the principal obtains her first-best payoffs, $u^{P*} = \max_{(u^A,u^P)\in\mathcal{F}} u^P$.*

Proof. By Lemma 4, the principal's problem reduces to the maximization of u^P among all $\left(u^A, u^P\right) \in \mathcal{F}$ such that $u^A \geq V$. Because $\mathcal{F}$ is compact, the problem has a solution with the stated properties. □

Theorem 1 establishes that despite the conflict of interest within the firm the equilibrium outcome is Pareto efficient. However, control of performance information is a powerful tool that allows the principal to either obtain her full control payoffs or extract all rents from the agent.

5 Wage Profile

Theorem 1 implies that equilibrium payoffs are generically unique. However, several equilibrium evaluation policies could possibly deliver these payoffs but result in different wage profiles. In what follows, I concentrate on a particular equilibrium in which the principal uses an incentive-compatible recommendation policy. By Lemma 3, this equilibrium always exists and there are at least two reasons to concentrate on it. First, the recommendation policies are minimally informative; any other policy can be Blackwell garbled into a recommendation policy without affecting the players' payoffs. Providing minimal information may be desirable to avoid its misuse by the agent in ways not conceivable by the principal. Second, the recommendation policies minimize wage volatility, which may be desirable for a firm.

Definition 3. An equilibrium (m^*, τ^*) is *minimally informative* if m^* is a recommendation policy and τ^* follows its recommendations.

In what follows, by equilibrium, I mean a minimally informative equilibrium. Hence, an equilibrium wage profile is deterministic. Indeed, in any period, there is a unique history of past evaluations that results in the agent remaining at the firm, namely, a sequence of recommendations to stay. Consequently, the equilibrium agent's career takes a particularly simple form. There is a deterministic wage profile W and commonly known performance requirements to stay at the firm. If the agent's performance satisfies these requirements, he remains at the firm; otherwise, he quits and secures a safe option.

The equilibrium wage profile depends on the equilibrium recommendation policy, which, in turn, depends on technology details. However, by Theorem 1, the equilibrium policy is efficient. It allows me to establish wage profile properties that hold for a broad class of technologies.

5.1 Downward Wage Rigidity

Naïve intuition suggests that average productivity and, hence, the wage should increase with tenure because the equilibrium is efficient. Indeed, efficiency is commonly associated with ongoing positive selection that eliminates poor performers and retains good performers. Such positive selection can be implemented through a sequence of history-dependent cutoffs such that the agent is recommended to quit whenever his productivity expectation drops below the corresponding cutoff. Under such a policy, the expected productivity would increase after every history, and hence, average productivity would increase with tenure.

Definition 4. A recommendation policy is a *cutoff policy* (in expectations) if there exists a cutoff function $q_t : Y^{t-1} \to \mathbb{R}$ such that

$$m_t = \begin{cases} \text{"stay,"} & \text{if } \mathbb{E}\left[\theta \mid y^{t-1}\right] > q_{t-1}\left(y^{t-2}\right), \\ \text{"quit,"} & \text{if } \mathbb{E}\left[\theta \mid y^{t-1}\right] < q_{t-1}\left(y^{t-2}\right). \end{cases}$$

The naïve intuition overlooks the fact that, in general, efficient selection should account for the whole profile of productivity beliefs, not just productivity expectation. Roughly, a greater productivity variance increases the chances of being highly productive and, hence, increases the option value of quitting and experimentation. An agent with a lower expected productivity but higher chances of being very productive could be worth retaining, whereas an agent with higher but certain productivity could be worth terminating. As a result, an efficient policy may not be cutoff, and the resulting equilibrium wage may decrease.

Nevertheless, I show that efficient policies are cutoff in many cases. By Lemmas 2 and 3, any efficient policy maximizes a payoff of an agent with some virtual safe option $\hat{V}^F \in$

$\left[\hat{V}^P, \hat{V}^A\right]$. Such policy is Markov in the productivity beliefs, so the space of beliefs can be split into two sets: the set at which the agent stays and the set at which the agent quits. The exact characterization depends on the parameters of the problem and can be intractable. However, it can be obtained in the following cases.

If there are only two productivity types, $|\Theta| = 2$, then there is a threshold expectation q^c such that under any efficient policy, the agent stays if the productivity expectation is above the threshold and quits otherwise.[9] That is, an optimal policy is cutoff with the cutoff function being constant.

If there are more than two types, $|\Theta| > 2$, then belief is a multidimensional object, and the efficient policy can be characterized only under additional assumptions.

Definition 5. A technology is *regular* if it admits densities, $|\Theta| < \infty$, and $\forall\, \theta, \theta' \in \Theta$, $y, y' \in \text{supp} f_\theta \cap \text{supp} f_{\theta'}$, $\theta' > \theta$, $y' > y$

$$f_{\theta'}(y')\, f_\theta(y) \geq f_{\theta'}(y)\, f_\theta(y') \qquad \text{(MLRP)}.$$

Under regular technology, the conditional performance distributions satisfy the monotone likelihood property. Most technologies used in the literature satisfy this condition. The regularity assumption adds the structure necessary to analyze the multiple-type case. Banks and Sundaram (1992) use the regularity condition to establish that an optimal strategy is cutoff.

Theorem 2. (Downward Wage Rigidity) *If there are only two types* $|\Theta| = 2$ *or the technology is regular, then in any minimally informative equilibrium:*

1. An evaluation policy m is a cutoff policy;
2. The wage profile W is deterministic and weakly increasing.

The theorem establishes sufficient conditions under which the equilibrium wage exhibits downward rigidity. The conditions are plausible in that they are satisfied in most existing models of career concerns and experimentation. Interestingly, the proof presents an even stronger statement: not only does expected productivity increase, but the whole profile of productivity beliefs also shifts upwards in an MLRP sense.

5.2 Wage Profile Shape

Theorem 2 establishes that under general conditions, the equilibrium wage profile is deterministic and weakly increasing. Nevertheless, the exact shape of the wage profile depends on

[9] This standard result is also proven in the Appendix.

technology details. Calculating the wage profile in closed form is intractable outside of very specific examples. To illustrate the role of technology in determining the equilibrium, in this section, I calculate wage profiles numerically and contrast them under different performance technologies.

Coarse Performance In many industries, performance measures are coarse. A lawyer's performance is captured by the number of successful trials, a consultant's performance is measured by the outcomes of his past projects, and a drug laboratory's performance is assessed by the number of new drugs developed. In all these cases, a performance outcome within any period is limited to a few options that cannot fully reveal productivity. I illustrate such cases with the following example. Here, performance is binary, low or high, $Y = \left\{y^L, y^H\right\}$, $y^L = 0$, $y^H = 1$. There are two equally likely types, $\Theta = \left\{\theta^L, \theta^H\right\}$, and the technology is:

$f_\theta(y)$	y^L	y^H
θ^L	$1-\theta^L$	θ^L
θ^H	$1-\theta^H$	θ^H

,

with $0 < \theta^L < \theta^H < 1$, meaning that no performance realization is conclusive. Observing y^H increases expected productivity; observing y^L decreases it. $V^A > \theta_0 > 0, V^P = 0$ that ensures that the principal cannot obtain her first-best payoff.

Because the type is binary, as discussed in Section 5.1, an equilibrium evaluation policy is cutoff with a constant cutoff q. The cutoff is chosen such that the agent obtains a payoff V^A by following the recommendations. The cutoff and the corresponding wage profile and quitting rate can be calculated by Monte Carlo simulations and are presented in Figure 3.

The agent's equilibrium career can be read off these plots. It features a clear promotion ladder in which infrequent performance revisions are followed by either quitting or obtaining a permanently higher wage.

Detailed Performance In other industries, the performance measures are detailed. For example, a manager can reveal his incompetency in day-to-day interactions with his clients and employees. I illustrate such settings with the following example. Performance outcomes are rich, $Y = \mathbb{R}$. There are two equally likely types, $\Theta = \left\{\theta^L, \theta^H\right\}$. Performance is distributed according to a Gaussian distribution:

$$y_t = \theta + \varepsilon, \quad \varepsilon \sim N\left(0, \sigma^2\right).$$

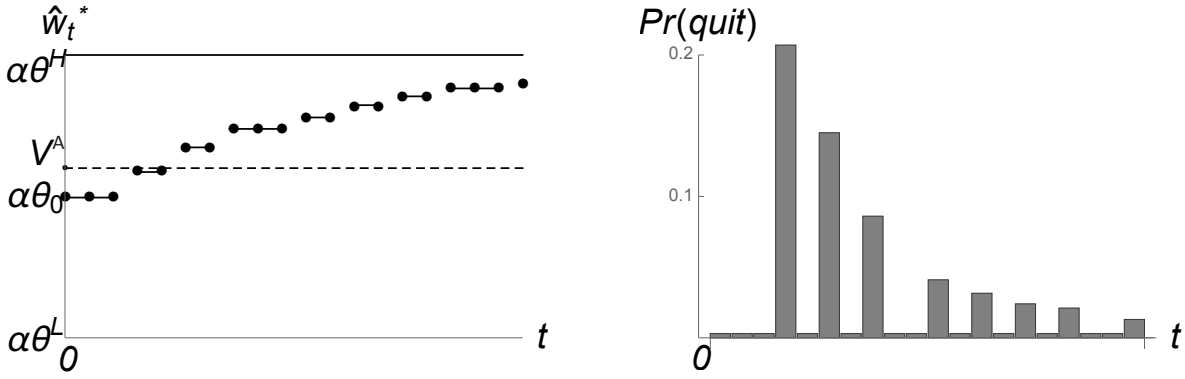

Figure 3: Wage profile (left) and quitting rate (right) in a minimally informative equilibrium. Coarse performance. $\Theta = \left\{\theta^L, \theta^H\right\}$, $Y = \{0, 1\}$, $\Pr(y = 1 \mid \theta) = \theta$, $w^F = 0$, $\alpha = 0.5$, $\theta^L = 0.2$, $\theta^H = 0.4$, $V^A = 0.6$, $V^P = 0$. Computed numerically by Monte Carlo simulation.

Expected productivity increases in performance. Moreover, the performance can be very conclusive—for any prior expectation below θ^H, the probability of interim expectations being arbitrarily close to θ^L is positive. With the payoff structure of the previous example, the equilibrium cutoff, wage profile, and quitting rate are computed numerically and are presented in Figure 4.

The quitting may occur in any period, and the wage strictly increases with tenure. Nevertheless, similar to coarse technology, the wage profile exhibits a (roughly) S shape that reflects the learning pattern and the rate of selection. At the beginning of the career, the quitting rate is low because there is little time to accumulate sufficiently negative performance information. Late in the career, the quitting rate is also low because, if retained, the agent was proven to have performed well and is likely to be a high type. Most selection occurs in the middle of the career, when information sufficient for selection is likely to be accumulated, but much productivity uncertainty remains.

6 Optimal Wage Contracts

In the previous sections, I have studied the design of optimal evaluation policies while holding the wage contract exogenously fixed. In organizations, the focus on information control can be motivated when payment schemes are more rigid and difficult to change than evaluation policies. However, it is natural to ask what contract the principal would prefer if she expected to complement it with optimal information provision. To address this question, in this section, I drop the Assumption 1 and allow the principal to freely choose the parameters of

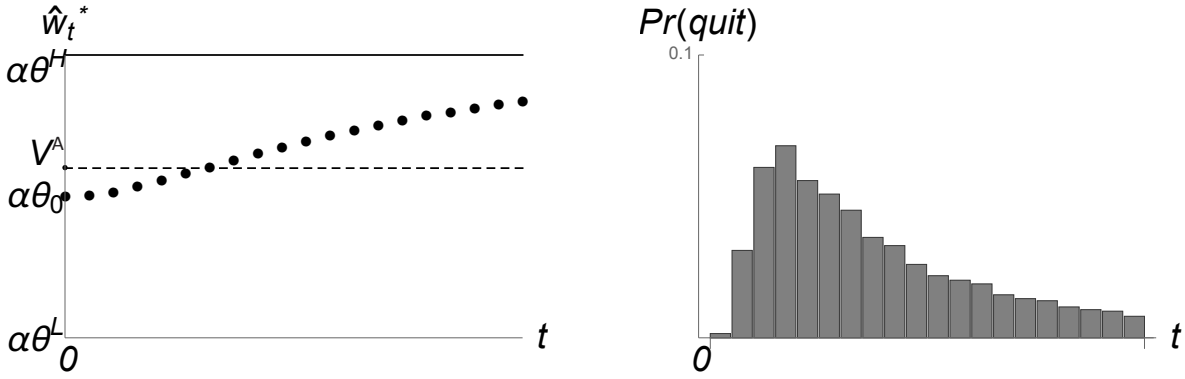

Figure 4: Wage profile (left) and quitting rate (right) in a minimally informative equilibrium. Rich performance. $\Theta = \left\{\theta^L, \theta^H\right\}$, $y_t = \theta + \varepsilon_t$, $\varepsilon_t \sim N(0, \sigma^2)$, $\sigma = 0.6$, $w^F = 0$, $\alpha = 0.5$, $\theta^L = 0$, $\theta^H = 1$, $V^A = 0.6$, $V^P = 0$. Computed numerically by Monte Carlo simulation.

the linear wage contract w^F and α.

Not very surprisingly, if the principal has both information and monetary control, then she can extract the full surplus from the relationship. To be precise, call a pair of strategies m^E, τ^E surplus-efficient if they deliver the maximal total payoff to the players. As payoffs are quasilinear in payments, wage contract details are irrelevant for efficiency. As before, m^E can be set to a complete-evaluation policy $\overline{m}$, because the quitting time can always ignore redundant information. Lemma 1 applies, and an efficient quitting time τ^E solves an optimal stopping problem with an outside option $V^E = V^A + V^P$. Call the corresponding delivered payoffs surplus efficient.

I say that a given *scheme* $\left(w^F, \alpha, m\right)$ extracts the full surplus if it implements the surplus-efficient payoffs in which the agent's payoff is equal to V^A, meaning that he obtains zero rents. As the agent can guarantee V^A by quitting at time zero, if a scheme extracts the full surplus, then this scheme is optimal for the principal among all possible schemes, with not necessarily linear contracts.

Assumption 2. (Positive Outside Options) $V^A > 0$ *and* $V^P + V^A > 0$.

The first part of Assumption 2 guarantees that the agent's would not prefer to remain at firm if offered zero compensation. The second part ensures that the principal's outside option is not too negative. It implies that the agent prefers following the surplus-efficient policy if offered a contract with $w^F = 0$ and $\alpha = 1$ to quitting at time zero.

Proposition 1. (Joint Scheme) *By jointly controlling the parameters of the wage contract and the evaluation policy, the principal can extract the full surplus. The equilibrium evaluation policy recommends a surplus-efficient quitting policy. The wage parameters can be set to either (i)* $w^F = V^A$, $\alpha = 0$, *or, under Assumption 2, (ii)* $w^F = 0$, $\alpha = \alpha^* \in \left[0, V^A/(V^A + V^P)\right]$.

The proof is constructive. As the aim of the principal is to extract the full surplus, by Lemma 3, she can focus on evaluations that recommend a surplus-efficient quitting time. The wage parameters are chosen such that the agent is willing to follow the recommendations and obtains zero rents. In the first scenario, $w^F = V^A, \alpha = 0$, the agent is effectively compensated according to his outside option. His wage is not linked to performance, and his payoff is fixed at V^A irrespective of when he quits. The incentives are trivially satisfied, and the agent obtains zero rents. In the second scenario, $w^F = 0$, $\alpha = \alpha^*$, the agent is compensated solely by a bonus. The incentive constraints are satisfied over an interval of bonuses. I show that it is possible to find a bonus that also leaves no rents to the agent.

To summarize, by using just one monetary tool, either a fixed wage or a bonus, the principal can extract the full surplus. A properly chosen fixed wage makes the agent indifferent between any quitting policy. It means that the principal can in principle provide a complete-evaluation policy, expecting the agent to quit efficiently. However, the agent's indifference may decrease the appeal of this policy in practice. In contrast, a properly chosen bonus compensation links the pay to performance and provides stronger incentives throughout the relationship. However, it should be complemented by a coarse evaluation policy: if the agent were provided complete evaluation, he would quit at his personally optimal time.

Let me highlight that in both scenarios the principal needs to provide informative evaluation to guide the agent's decisions as long as surplus-efficient decisions depend on performance. No evaluation is generically not optimal even though the principal controls the monetary incentives.

7 Discussion

7.1 Organizational Implications

Wages and Career My analysis draws attention to a role that evaluations may play in explaining internal labor market dynamics. Internal labor markets, or personnel economics, has attracted considerable attention in the economic literature.[10] Some of the robust empirical findings are (i) a well-defined career ladder within a firm (Baker, Gibbs, and Holmstrom

[10] Waldman (2013) provides an excellent overview of the related literature.

(1994a,b); Seltzer and Merrett (2000); Dohmen (2004)), (ii) nominal, and to a large extent real, downward wage rigidity (Baker et al. (1994a,b); Card and Hyslop (1997)), (iii) wage growth that cannot be attributed to human capital formation (Baker et al. (1994a,b); Medoff and Abraham (1980, 1981)), and (iv) wage compression, i.e., insufficient wage variation within worker cohorts (Akerlof and Yellen (1990); Baker et al. (1994a,b)).[11]

It has been presumed that these patterns cannot be explained by learning models of wage formation. Indeed, if a worker is provided detailed performance information and his wage is proportional to expected productivity, then, for the duration of his stay at the firm, his wage should follow a random walk. If one examined many workers within the same cohort, this would imply that over time, the cohort wage distributions present a sequence of mean-preserving spreads.

However, I show that it is possible to reconcile all these organizational patterns within a pure learning model with endogenous quitting, if a wage responds not to detailed performance but to optimally designed evaluations. In my model, the equilibrium wage is a deterministic function of tenure such that all workers from the same cohort receive the same salary. Moreover, the wage never decreases: poor evaluations are infrequent and result in an immediate resignation, whereas good evaluations lead to a permanent wage increase. Wage growth is fueled by positive ongoing selection and shapes an endogenous career ladder.

Of course, any pure learning model would be too stylized to claim universal applicability; a hybrid model that combines several features such as that of Gibbons and Waldman (1999) would be more appropriate. However, my analysis suggests that evaluation practices should be given closer attention in applications and that learning theories of wage formation have the potential to explain evidence. The learning considerations may be particularly relevant for professional firms in which employee talents have a high impact on performance and are learned gradually over time (O'Flaherty and Siow (1995)).

Evaluation Practices My analysis provides a novel perspective on evaluation practices in organizations. First, it is often observed that evaluations feature rating compression and leniency bias—too many employees are bunched at the highest grades (Murphy and Cleveland (1995)). This feature is often attributed to behavioral biases of evaluators and criticized for reducing evaluation informativeness. However, my analysis shows that these practices can in fact benefit a firm by encouraging employees to supply more effort.[12] Second, there is no consensus in the literature on whether evaluations should be retrospective, that is, account

[11] In addition, downward wage rigidity has attracted considerable attention in the development economics literature (e.g., Schmitt-Grohé and Uribe (2016); Kaur (2019)).

[12] Relatedly, Ostrovsky and Schwarz (2010) demonstrate how grade inflation can emerge as an equilibrium grading policy in school competition.

for past performance. My analysis suggests that such an accounting is generally necessary. The equilibrium evaluation policy is Markov in belief based on the entire performance history; in general, this policy cannot be implemented through a sequence of grades based solely on current performance.

The equilibrium evaluation policy can also have a psychological interpretation for maintaining workforce morale. It has been confirmed both in an experimental setting (Kuhnen and Tymula (2012)) and in the field (Kolstad (2013)) that information provision is an important incentive device that interacts with workers' intrinsic motivation. At the same time, some organizational literature has informally argued that evaluation provision should balance two opposing forces—the employee's need for evaluation and its damage to his self-esteem (Murphy and Cleveland (1995)). My model can be viewed as a formalization of these forces with the agent's morale captured by his productivity belief.[13] On one hand, evaluation is a valuable input for employees' future decisions. On the other hand, any evaluation may turn out negative and discourage an employee. The equilibrium policy can then be interpreted as the policy that optimally trades off these two competing effects.

7.2 Value of Human Resource Departments

My analysis relies on the ability of the firm to commit to its evaluation policy. In fact, several layers of commitment matter. First, I assume that evaluations are objective, so that the information is verifiable. If the evaluation were "cheap talk," then the equilibrium recommendation policy would not be credible—the principal would always recommend that the agent stay. Second, I assume that the principal commits to the evaluation rules at the beginning of the relationship and cannot change them in the future. This assumption matters because under the equilibrium policy, as long as the agent remains at a firm, his continuation value increases, and his incentive constraints become slack. In the absence of intertemporal commitment, even if the players are symmetrically informed, the principal would be tempted to change the evaluation policy to extract more continuation surplus.

For these reasons, some degree of commitment power is required to implement the equilibrium policy. In practice, this commitment can be plausibly achieved by having an HR department. Nearly all of the large firms have such departments (Gomez-Mejia, Balkin, and Cardy (2014)). They are relatively independent and administer performance appraisal of employees, controlling what information is collected and transmitted throughout a firm. HR departments provide clear instructions regarding the frequency of evaluations, the metrics that are being assessed, and the grading scales. The rules are transparent and the

[13] Fang and Moscarini (2005) study a similar tradeoff in a static setting.

performance metrics are largely objective.[14] The question of evaluation design can then be interpreted as precisely what evaluation polices should be administered by an HR department to maximize a firm's objectives.

Relatedly, my framework can be used to assess the value of having such an HR department: if the evaluations were subjective and given by biased parties, then the value of information control could be lost altogether.

7.3 Performance Leaks

Throughout the analysis, I assumed that the principal has full control over the information flow of the agent. This a typical assumption in the information design literature, and it provides an upper bound on what the principal can possibly achieve. However, it is possible that some performance information may leak; for example, the agent may have "gut feelings" about when he performed very well in the past. Such leaks can both limit the scope for information control and increase the agent's guaranteed payoff. At the extreme, if all performance information is leaked, then there is no scope for information design, and the agent obtains his first best payoff.

However, some performance leaks are benign because they may be not acted upon. To illustrate, I let the productivity space be binary, $\Theta = \left\{\theta^L, \theta^H\right\}$, and assume that the technology is *strongly regular,* meaning that it is regular, $\text{supp} f_{\theta^L} = \text{supp} f_{\theta^H} = \mathbb{R}$, and $f_{\theta^H}(y) / f_{\theta^L}(y)$ spans $[0, +\infty)$ as y spans $\mathbb{R}$. Under strongly regular technology, performance is inconclusive, meaning that the principal's belief $\mu_t \triangleq \Pr\left(\theta^H \mid y^{t-1}\right)$ never equals 0 or 1. At the same time, the belief is variable, meaning that for any $\mu_{t-1} \in (0,1)$, performance y_{t-1} can swing belief μ_t to anywhere in $(0,1)$. Call a performance leak a partitional disclosure with a cutoff κ_t if it reveals whether the principal's belief μ_t is above κ_t. A partitional disclosure reveals whether the agent is a top performer. Within each period, let the leak occur after the evaluation but before the agent's action.

Claim 1. (Performance Leaks) Let $|\Theta| = 2$, the technology be strongly regular, and the principal evaluate the agent according to the equilibrium recommendation policy. Then, there exist cutoffs $\underline{\kappa}_t < 1$ for $t = 0, 1, \ldots$ such that the leaks of partitional disclosures with cutoffs $\kappa_t \geq \underline{\kappa}_t$ do not affect players' equilibrium behavior or payoffs.

Intuitively, the leaks in Claim 1 reveal only sufficiently top performers. Such leaks make top performers even more optimistic and keep them at the firm. Less-than-top performers do become less optimistic but not sufficiently so to find it worthwhile to quit the firm. Hence,

[14] Even when seeking input from employee's supervisor, appraisal systems typically present questions in a way to minimize response bias (relative scales, numerous categories, etc.).

the leaks do not interfere with equilibrium evaluations.

More generally, first, note that whenever the agent is recommended to quit, he may obtain full access to his performance history: the principal recommends quitting only at beliefs at which the agent himself would prefer to quit. Second, as discussed in Section 4.3, as long as the agent is recommended to stay, his continuation value increases with tenure. Hence, the incentives to quit are being progressively relaxed. Consequently, if the leaks do not drive the agent's continuation value below his outside option, he is willing to follow the recommendations. This clearly occurs if the leaks are sufficiently uninformative. However, even conclusive leaks may not interfere with equilibrium evaluations as long as they do not bring very negative news, as shown in Claim 1.

7.4 Evaluation and Advice

Throughout the paper, I focused on *evaluation*—which can be regarded as feedback informative about the agent's fit at the firm. The corresponding allocation of effort is vertical—whether to stay at the firm. I showed that this feedback should be coarse to avoid discouraging the agent too quickly and prevent him from quitting prematurely. Evaluation is to not be confused with another kind of feedback—*advice*—that guides the agent's actions within a firm and promotes his professional growth (Gomez-Mejia et al. (2014)). For example, a manager may guide an employee to choose the best fitting project or otherwise coach him throughout his career. The corresponding allocation of effort may be regarded as being horizontal. It is plausible that in practice, the agent's and the firm's interests are aligned on this horizontal dimension.

In some cases, advice and evaluation can be separated. For example, the efficient advice may be driven by firm-specific goals and not depend on the agent's fit within the firm. In those cases, the current analysis can be readily applied. The firm should optimally provide exhaustive advice but coarse evaluation: the former would maximize the efficiency of the agent's stay at the firm, while the latter would shield against premature quitting. However, it is possible to envision settings in which these two kinds of feedback cannot be disentangled; for example, the manager's advice to leave a trending project may inevitably be interpreted as a signal of a poor overall fit within the company. If such entanglement is severe, then additional analysis needs to be performed to understand an optimal information policy. However, in light of the current findings, I expect the policy to be coarse and distort the agent's actions towards the firm's interests.

8 Conclusion

Performance evaluation is an important part of organizational life. Although much evaluation is informal, most firms have formal evaluation policies designed to collect and distribute performance information to employees. In this work, I showed that an equilibrium design of these policies can explain many features of internal labor markets observed in practice. Information control allows the organization to extract rents from its employees while maintaining operational efficiency.

For clarity and conciseness, I abstracted away from many realistic features of internal labor markets. Incorporating these features would make the model more applicable in practice and constitutes a plausible venue for future work. The efficiency approach developed in this paper may facilitate analysis of these richer settings. In the meantime, the proposed theory of endogenous evaluation could be viewed as complimenting existing theories of internal labor markets.

References

Akerlof, G. A. and J. L. Yellen (1990): "The Fair Wage-Effort Hypothesis and Unemployment," *Quarterly Journal of Economics*, 105, 255–283.

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

——— (1994b): "The Wage Policy of a Firm," *Quarterly Journal of Economics*, 109, 921–955.

Banks, J. S. and R. K. Sundaram (1992): "Denumerable-Armed Bandits," *Econometrica*, 1071–1096.

Bergemann, D. and J. Välimäki (2008): "Bandit Problems," in *New Palgrave Dictionary of Economics*, ed. by S. N. Durlauf and L. E. Blume, Palgrave Macmillan.

Berry, D. A. and B. Fristedt (1985): "Bandit Problems: Sequential Allocation of Experiments," *London: Chapman and Hall*, 5, 7–7.

Card, D. and D. Hyslop (1997): "Does Inflation "Grease the Wheels of the Labor Market"?" in *Reducing Inflation: Motivation and Strategy*, University of Chicago Press, 71–122.

Carroll, G. (2015): "Robustness and Linear Contracts," *American Economic Review*, 105, 536–63.

Dohmen, T. J. (2004): "Performance, Seniority, and Wages: Formal Salary Systems and Individual Earnings Profiles," *Labour Economics*, 11, 741–763.

Ederer, F. (2010): "Feedback and Motivation in Dynamic Tournaments," *Journal of Economics & Management Strategy*, 19, 733–769.

Ely, J. C. (2017): "Beeps," *American Economic Review*, 107, 31–53.

Ely, J. C. and M. Szydlowski (2020): "Moving the Goalposts," *Journal of Political Economy*, 128, 468–506.

Fang, H. and G. Moscarini (2005): "Morale Hazard," *Journal of Monetary Economics*, 52, 749–777.

Gibbons, R. and M. Waldman (1999): "A Theory of Wage and Promotion Dynamics Inside Firms," *Quarterly Journal of Economics*, 114, 1321–1358.

Gittins, J. C. and D. M. Jones (1974): "A Dynamic Allocation Index for the Sequential Design of Experiments," in *Progress in Statistics*, ed. by I. Vincze, J. Gani, and K. Sarkadi, Amsterdam: North-Holland Pub. Co., 241–266.

Goltsman, M. and A. Mukherjee (2011): "Interim Performance Feedback in Multistage Tournaments: The Optimality of Partial Disclosure," *Journal of Labor Economics*, 29, 229–265.

Gomez-Mejia, L. R., D. B. Balkin, and R. L. Cardy (2014): *Managing Human Resources*, Pearson, 8 ed.

Guo, Y. (2016): "Dynamic Delegation of Experimentation," *American Economic Review*, 106, 1969–2008.

Halac, M., N. Kartik, and Q. Liu (2016): "Optimal Contracts for Experimentation," *Review of Economic Studies*, 83, 1040–1091.

Hansen, S. E. (2013): "Performance Feedback with Career Concerns," *Journal of Law, Economics, and Organization*, 29, 1279–1316.

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

Hörner, J. and Y. Guo (2015): "Dynamic Mechanisms without Money," *Working Paper*.

Hörner, J. and N. Lambert (2020): "Motivational Ratings," *Review of Economic Studies, Forthcoming*.

Jovanovic, B. (1979): "Job Matching and the Theory of Turnover," *Journal of Political Economy*, 87, 972–990.

Kamenica, E. and M. Gentzkow (2011): "Bayesian Persuasion," *American Economic Review*, 101, 2590–2615.

Kaur, S. (2019): "Nominal Wage Rigidity in Village Labor Markets," *American Economic Review*, 109, 3585–3616.

Kolstad, J. T. (2013): "Information and Quality when Motivation is Intrinsic: Evidence from Surgeon Report Cards," *American Economic Review*, 103, 2875–2910.

Kremer, I., Y. Mansour, and M. Perry (2014): "Implementing the 'Wisdom of the Crowd'," *Journal of Political Economy*, 122, 988–1012.

Kuhnen, C. M. and A. Tymula (2012): "Feedback, Self-Esteem, and Performance in Organizations," *Management Science*, 58, 94–113.

Medoff, J. L. and K. G. Abraham (1980): "Experience, Performance, and Earnings," *Quarterly Journal of Economics*, 95, 703–736.

——— (1981): "Are Those Paid More Really More Productive? The Case of Experience," *Journal of Human resources*, 186–216.

Murphy, K. R. and J. Cleveland (1995): *Understanding Performance Appraisal: Social, Organizational, and Goal-Based Perspectives*, Sage.

Myerson, R. B. (1986): "Multistage Games with Communication," *Econometrica*, 54, 323–358.

O'Flaherty, B. and A. Siow (1995): "Up-or-out Rules in the Market for Lawyers," *Journal of Labor Economics*, 13, 709–735.

Orlov, D. (2016): "Optimal Design of Internal Disclosure," *Working Paper*.

Orlov, D., A. Skrzypacz, and P. Zryumov (2020): "Persuading the Principal to Wait," *Journal of Political Economy*, 128, 2542–2578.

Ostrovsky, M. and M. Schwarz (2010): "Information Disclosure and Unraveling in Matching Markets," *American Economic Journal: Microeconomics*, 2, 34–63.

Rayo, L. and I. Segal (2010): "Optimal Information Disclosure," *Journal of Political Economy*, 118, 949–987.

Renault, J., E. Solan, and N. Vieille (2017): "Optimal Dynamic Information Provision," *Games and Economic Behavior*, 104, 329–349.

Ross, S. (1983): *Introduction to Stochastic Dynamic Programming*, New York: Academic Press.

Schmitt-Grohé, S. and M. Uribe (2016): "Downward Nominal Wage Rigidity, Currency Pegs, and Involuntary Unemployment," *Journal of Political Economy*, 124, 1466–1514.

Seltzer, A. and D. Merrett (2000): "Personnel Policies at the Union Bank of Australia: Evidence from the 1888–1900 Entry Cohorts," *Journal of Labor Economics*, 18, 573–613.

Thompson, F. B. (1952): "Equivalence of Games in Extensive Form," Technical report rm-759, RAND Corporation, Washington, D.C.

Waldman, M. (2013): "Theory and Evidence in Internal Labor Markets," in *The Handbook of Organizational Economics*, ed. by R. S. Gibbons and J. Roberts, Princeton University Press.

9 Appendix

Proof of Lemma 1 Take any evaluation policy m and consider a stochastic process $X = \{X_s\}_{s\geq 1}$ with:

$$X_s \triangleq (1-\delta)\sum_{t=0}^{s-1}\delta^t\left(y_t - \mathbb{E}\left[\theta \mid m^t\right]\right).$$

The process X is a bounded martingale with respect to the filtration generated by m because:

$$\begin{aligned}\mathbb{E}\left[X_{s+1} \mid m^s\right] &= X_s + (1-\delta)\,\delta^t\mathbb{E}\left[y_s - \mathbb{E}\left[\theta \mid m^s\right] \mid m^s\right]\\ &= X_s + (1-\delta)\,\delta^t\left(\mathbb{E}\left[y_s \mid m^s\right] - \mathbb{E}\left[\theta \mid m^s\right]\right)\\ &= X_s + (1-\delta)\,\delta^t\left(\mathbb{E}\left[\theta \mid m^s\right] - \mathbb{E}\left[\theta \mid m^s\right]\right) = X_s.\end{aligned}$$

The second line follows from the law of iterated expectations and the third line follows from the definition of productivity. It follows from the optional stopping theorem that for any stopping time τ measurable with respect to m, $\mathbb{E}\left[X_\tau\right] = \mathbb{E}\left[X_0\right] = 0$. Equivalently, by the definition of wage (3), for any evaluation policy m and stopping time τ:

$$U^P(m,\tau) = \mathbb{E}_{m,\tau}[(1-\delta)\sum_{t=0}^{\tau-1}\delta^t\left(y_t - w_t\right)] + \delta^\tau V^P = \mathbb{E}_{m,\tau}[(1-\delta)\sum_{t=0}^{\tau-1}\delta^t((1-\alpha)\,y_t - w^F) + \delta^\tau V^P]$$

$$= \mathbb{E}_{m,\tau}[(1-\alpha)\,(1-\delta)\sum_{t=0}^{\tau-1}\delta^t y_t + \delta^\tau(V^P + w^F)] - w^F = (1-\alpha)\,\mathbb{E}_{m,\tau}[(1-\delta)\sum_{t=0}^{\tau-1}\delta^t y_t + \delta^\tau\hat{V}^P] - w^F.$$

The agent's payoff is derived analogously. □

Proof of Lemma 2. Any feasible payoff can be obtained with a complete evaluation policy $\overline{m}$. Indeed, any filtration generated by m^t is coarser than a filtration generated by y^{t-1}. Hence, the set of feasible joint distributions of y and τ is maximized by setting $m^t \equiv y^{t-1}$. Consequently, if payoffs $\left(u^A, u^P\right)$ can be generated by some (m,τ), then they can be generated by $(\overline{m},\tau')$, where $\overline{m}$ is a complete evaluation policy.

Hence, $\mathcal{F}$ is a set of payoffs that can be generated by some quitting time τ under complete evaluation. The quitting time can equivalently be identified with a quitting rule $\sigma : Y^{t-1} \to \Delta\left(\text{"stay"}, \text{"quit"}\right)$; the players' expected payoffs $\left(u^A(\sigma), u^P(\sigma)\right)$ are linear as functions of σ. Endowed with the sup norm, the set of quitting rules is a compact and convex vector space. Moreover, by the feasibility argument presented above, $\mathcal{F}$ is a subset of $\hat{\mathcal{F}}$—the set of payoffs feasible if θ were revealed at time 0. As $\mathbb{E}\left[\theta_0\right]$ exists, $\hat{\mathcal{F}}$, and thus $\mathcal{F}$ is bounded. Hence, $\mathcal{F}$ is convex and compact as a bounded linear image of a convex and compact set.

Denote the boundary of $\mathcal{F}$ by $\partial\mathcal{F}$, $\partial\mathcal{F} \triangleq \mathcal{F}\backslash\text{int}\left(\mathcal{F}\right)$. By the separating hyperplane

theorem, for payoffs $\left(u^A, u^P\right) \in \partial\mathcal{F}$, there exist $\lambda^A, \lambda^P \in \mathbb{R}$ not both equal to zero such that:

$$\left(u^A, u^P\right) \in \arg \max_{(u^{A\prime}, u^{P\prime}) \in \mathcal{F}} \lambda^A u^{A\prime} + \lambda^P u^{P\prime}. \tag{12}$$

Conversely, for any $\lambda^A, \lambda^P \in \mathbb{R}$ not both equal to zero, a solution to (12) belongs to $\partial\mathcal{F}$. By Lemma 1, (12) is equivalent to maximizing:

$$\mathbb{E}_{\overline{m},\tau}\left[(1-\delta)\sum_{t=0}^{\tau}\delta^t\left(\lambda^A\alpha + \lambda^P(1-\alpha)\right)y_t + \delta^\tau\left(\lambda^A\alpha\hat{V}^A + \lambda^P(1-\alpha)\hat{V}^P\right)\right] + w^F\left(\lambda^A - \lambda^P\right),$$

over all quitting times τ. Efficient payoffs correspond to $\lambda^A, \lambda^P \geq 0$, not both equal to zero. In this case, $\lambda^A\alpha + \lambda^P(1-\alpha) > 0$ and maximizing the combination is equivalent to maximizing a payoff of an agent with a virtual safe option $\hat{V}^F = \frac{\lambda^A\alpha}{\lambda^A\alpha+\lambda^P(1-\alpha)}\hat{V}^A + \frac{\lambda^A(1-\alpha)}{\lambda^A\alpha+\lambda^P(1-\alpha)}\hat{V}^P$. By Assumption 1, $\hat{V}^A > \hat{V}^P$; thus, as λ^A, λ^P span all possible values, $\hat{V}^F$ spans $\left[\hat{V}^P, \hat{V}^A\right]$. □

Proof of Lemma 4. Consider any efficient payoffs $\left(u^A, u^P\right)$. If $u^A < V^A$, then the payoffs cannot be implemented because the agent can secure a payoff V^A by quitting at time 0. If $u^A \geq V^A$, then by Lemma 2 the payoffs can be delivered through strategies $(\overline{m}, \tau)$ that maximize the payoff of an agent with a safe option $\hat{V}^F \in \left[\hat{V}^P, \hat{V}^A\right]$. Consider the recommendation policy m' that mimics τ. I show that m' is incentive compatible.

Recommendations to quit are incentive compatible. If the agent is recommended to quit at time t, his wage at the firm remains constant at some level w_t^Q hereafter. Because the strategy maximizes the payoff of an agent with a safe option $\hat{V}^F$, $w_t^Q \leq \hat{V}^F$; otherwise, remaining at firm forever onward would increase the payoff. Since $\hat{V}^F \leq \hat{V}^A$, the agent with a safe option $\hat{V}^A$ also agrees to quit.

Recommendations to stay are also incentive compatible. A recommendation at time 0 is incentive compatible because it gives the agent a payoff of at least u^A and $u^A \geq V^A$. Incentive compatibility at time $t+1$ follows from incentive compatibility at time t. Indeed, denote by $\tilde{U}_t^A$ the agent's optimal continuation payoff conditional on being recommended to stay at time t. Because the recommendation to stay is incentive compatible at time t and the recommendations to quit are incentive compatible everywhere, $\tilde{U}_t^A$ is a convex combination of w_t, $\tilde{U}_{t+1}^A$, and V^A:

$$\tilde{U}_t^A = (1-\delta)w_t + \delta\Pr\left(m_{t+1} = \text{"stay"} | m^t \equiv \text{"stay"}\right)\tilde{U}_{t+1}^A + \delta\Pr\left(m_{t+1} = \text{"quit"} | m^t \equiv \text{"stay"}\right)V^A.$$

The agent obtains a continuation payoff V^A if he quits immediately so $\tilde{U}_t^A \geq V^A$. The agent obtains a continuation payoff w_t if he never quits so $\tilde{U}_t^A \geq w_t$. Thus, for the equality to hold,

it must be that $\tilde{U}^A_{t+1} \geq \tilde{U}^A_t$. By forward induction, all recommendations to stay are incentive compatible. This completes the proof. □

Lemma 5. *Fix $V \in \mathbb{R}$, $\Theta = \left\{\theta^L, \theta^H\right\}$, and an arbitrary technology. Consider the following maximization problem:*

$$\max_{\tau \in \mathcal{T}} \mathbb{E}_{\overline{m},\tau}\left[(1-\delta)\sum_{t=0}^{\tau-1}\delta^t \mathbb{E}\left[\theta \mid \overline{m}^t\right] + \delta^\tau V\right].$$

Then, (i) if $V > \theta^H$, then $\tau \equiv 0$ is uniquely optimal; (ii) if $V < \theta^L$, then $\tau \equiv \infty$ is uniquely optimal; and (iii) if $V \in \left[\theta^L, \theta^H\right]$, then any optimal strategy τ is equivalent in terms of quitting distributions to a strategy that follows the recommendations of a cutoff recommendation policy with some fixed cutoff $q \in \left[\theta^L, \theta^H\right]$.

Proof. This is a standard bandit problem of an agent choosing between a risky arm with payoff $\mathbb{E}\left[\theta \mid \overline{m}^t\right]$ and a safe arm with payoff V. This is a Markov decision problem with an expected productivity $\hat{\theta}_t \triangleq \mathbb{E}\left[\theta \mid y^{t-1}\right]$ being a sufficient statistic for beliefs. An optimal quitting strategy is characterized by Gittins and Jones (1974) as an index policy. The risky arm is assigned an index $\xi\left(\hat{\theta}_t\right)$. The agent quits the arm as soon as the index drops below V; he can arbitrarily randomize at $\xi\left(\hat{\theta}_t\right) = V$.

The fact that $\xi\left(\hat{\theta}_t\right)$ is increasing in $\hat{\theta}_t$ is a direct consequence of Lemma 3.2 of Banks and Sundaram (1992) that states that the index is increasing along any belief ray that passes through the "worst" prior. For completeness, I provide a direct argument here. By Gittins and Jones (1974), the index can be calculated as:

$$\begin{aligned}
\xi\left(\hat{\theta}_t\right) &\triangleq \sup_\tau \frac{\mathbb{E}\left[\sum_{t=0}^\tau \delta^t \mathbb{E}\left[\theta \mid \overline{m}^t\right] \mid \hat{\theta}_t\right]}{\mathbb{E}\left[\sum_{t=0}^\tau \delta^t \mid \hat{\theta}_t\right]} \\
&= \sup_\tau \frac{\frac{\hat{\theta}_t - \theta^L}{\theta^H - \theta^L}\mathbb{E}\left[\sum_{t=0}^\tau \delta^t y_t \mid \theta = \theta^H\right] + \frac{\theta^H - \hat{\theta}_t}{\theta^H - \theta^L}\mathbb{E}\left[\sum_{t=0}^\tau \delta^t y_t \mid \theta = \theta^L\right]}{\frac{\hat{\theta}_t - \theta^L}{\theta^H - \theta^L}\mathbb{E}\left[\sum_{t=0}^\tau \delta^t \mid \theta = \theta^H\right] + \frac{\theta^H - \hat{\theta}_t}{\theta^H - \theta^L}\mathbb{E}\left[\sum_{t=0}^\tau \delta^t \mid \theta = \theta^L\right]} \\
&= \sup_\tau \frac{\left(\hat{\theta}_t - \theta^L\right)\theta^H \mathbb{E}\left[1 - \delta^{\tau+1} \mid \theta = \theta^H\right] + \left(\theta^H - \hat{\theta}_t\right)\theta^L \mathbb{E}\left[1 - \delta^{\tau+1} \mid \theta = \theta^L\right]}{\left(\hat{\theta}_t - \theta^L\right)\mathbb{E}\left[1 - \delta^{\tau+1} \mid \theta = \theta^H\right] + \left(\theta^H - \hat{\theta}_t\right)\mathbb{E}\left[1 - \delta^{\tau+1} \mid \theta = \theta^L\right]} \\
&= \sup_\tau \theta^L + \frac{\left(\hat{\theta}_t - \theta^L\right)\left(\theta^H - \theta^L\right)\mathbb{E}\left[1 - \delta^{\tau+1} \mid \theta = \theta^H\right]}{\left(\hat{\theta}_t - \theta^L\right)\mathbb{E}\left[1 - \delta^{\tau+1} \mid \theta = \theta^H\right] + \left(\theta^H - \hat{\theta}_t\right)\mathbb{E}\left[1 - \delta^{\tau+1} \mid \theta = \theta^L\right]} \\
&= \theta^L + \frac{\left(\hat{\theta}_t - \theta^L\right)\left(\theta^H - \theta^L\right)}{\left(\hat{\theta}_t - \theta^L\right) + \left(\theta^H - \hat{\theta}_t\right)\kappa},
\end{aligned}$$

where $\kappa \triangleq \inf_\tau \frac{\mathbb{E}\left[1-\delta^{\tau+1}|\theta=\theta^L\right]}{\mathbb{E}\left[1-\delta^{\tau+1}|\theta=\theta^H\right]} > 0$, a constant fully determined by F_θ, and the second and

third lines follow from the optional stopping theorem and the law of iterated expectations. Thus, $\xi\left(\hat{\theta}_t\right)$ strictly increases with $\hat{\theta}_t$. □

Proof of Theorem 2. I start backwards and first show that if the evaluation policy is cutoff, then the wage profile is weakly increasing. By the law of iterated expectations:

$$\mathbb{E}\left[\theta \mid m^t\right] = \mathbb{E}\left[\mathbb{E}\left[\theta \mid y^{t-1}\right] \mid m^t\right] = \mathbb{E}\left[\hat{\theta}_t \mid m^t\right].$$

If the policy always recommends to stay at the boundary, when $\hat{\theta}_t = q_{t-1}\left(y^{t-2}\right)$, then

$$\mathbb{E}\left[\hat{\theta}_t \mid m^t = (\text{"stay"}, \ldots, \text{"stay"})\right] = \mathbb{E}\left[\hat{\theta}_t \mid \hat{\theta}_1 \geq q_0, \ldots, \hat{\theta}_{t-1} \geq q_{t-2}\left(y^{t-2}\right), \hat{\theta}_t \geq q_{t-1}\left(y^{t-1}\right)\right] \geq$$
$$\mathbb{E}\left[\hat{\theta}_t \mid \hat{\theta}_1 \geq q_0, \ldots, \hat{\theta}_{t-1} \geq q_{t-2}\left(y^{t-2}\right)\right] = \mathbb{E}\left[\hat{\theta}_t \mid m^{t-1} = (\text{"stay"}, \ldots, \text{"stay"})\right],$$

where the inequality follows by the definition of a conditional expectation and the implied positive selection. The case of randomization at the boundary is analogous. As the wage is proportional to expected ability, the wage profile is weakly increasing.

I proceed with establishing that under the conditions of the theorem, the evaluation policy is indeed cutoff. If the type is binary, $|\Theta| = 2$, then by Lemma 5, an optimal policy is cutoff ($\tau \equiv 0$ and $\tau \equiv \infty$ correspond to cutoffs greater than θ^H and smaller than θ^L, respectively).

If the type is not binary, $|\Theta| > 2$, but the performance technology is regular, then I can appeal to the results of Ross (1983) and Banks and Sundaram (1992). For completeness, I present the proof here.

To proceed, we need some additional definitions. Associate the k-highest type with an index k and let $K \triangleq |\Theta|$. For any $g_1, g_2 \in \Delta(\Theta)$, say that g_2 discretely (stochastically) dominates g_1, $g_2 \succeq_d g_1$, if for all $l \in \{1, \ldots, K\}$, $\sum_{k=1}^{l} g_{2k} \geq \sum_{k=1}^{l} g_{1k}$. Similarly, for any f_1, f_2 that are probability densities over Y, say that f_2 continuously (stochastically) dominates f_1, $f_2 \succeq_c f_1$, if for any increasing function h, $\int h(y) f_{g_2}(y)\, dy \geq \int h(y) f_{g_1}(y)\, dy$. Denote by $\beta(g, y) \in \Delta(\Theta)$ the posterior distribution calculated by Bayes' rule given a prior distribution g and a one-period performance outcome $y \in Y$. With a small abuse of notation, define $f_g(y) \triangleq \sum_{\theta \in \Theta} g(\theta) f_\theta(y)$, and denote by $\hat{Y} \triangleq \cup_\theta \text{supp}\left(f_\theta(y)\right)$.

I provide some intuitive and standard consequences of regularity of technology driven my the MLRP property without a proof (see Ross (1983) and Banks and Sundaram (1992) for details).

Claim 2. For any $g \in \Delta(\Theta)$ and $y, y' \in \hat{Y}$ such that $y' > y$, $\beta(g, y') \succeq_d \beta(g, y)$.

Claim 3. For any $g, g' \in \Delta(\Theta)$ such that $g' \succeq_d g$, $f_{g'} \succeq_c f_g$.

We are ready to proceed to the main part of the proof. By Lemma 2 and Theorem 1, any optimal recommendation policy solves an optimal stopping problem of a fictitious agent. This is a Markov decision problem with the agent's beliefs $g_t \in \Delta(\Theta)$ being a state variable. Denote the agent's continuation value as a function of beliefs $\mathcal{V}(g)$:

$$\mathcal{V}(g) = \max\left[\hat{V}^F, \mathbb{E}_g[\theta] + \delta \int \mathcal{V}(\beta(g,y)) f_g(y)\,\mathrm{d}y\right].$$

By Claim 2, any beliefs following the same history $g_t = \beta(g_{t-1}, y), g'_t = \beta(g_{t-1}, y')$ can be ranked according to $\succeq_d$. To establish that the policy is cutoff, it then suffices to show that in any period t if the agent stays at belief g_{t1}, he does so for all beliefs $g_{t2} \succeq g_{t1}$. As quitting results in an immediate payoff $\hat{V}^F$, the policy is cutoff if for any $g_0 \in \Delta(\Theta)$ and $y_1, y_2 \in \hat{Y}$ such that $y_2 > y_1$, $\mathcal{V}(\beta(g_0, y_2)) \geq \mathcal{V}(\beta(g_0, y_1))$. This can be shown by an induction proof. Denote $g_2 = \beta(g_0, y_2)$ and $g_1 = \beta(g_0, y_1)$. By Claim 2, $g_2 \succeq_d g_1$. Consider the stopping problem in which the horizon is truncated to T periods, $T = 0, 1, \ldots$; this corresponds to altering the discount sequence to $\left\{1, \delta, \delta^2, \ldots, \delta^{T-1}, 0, 0, \ldots\right\}$. Denote by $\mathcal{V}_T(g)$ the value function of this problem at time 0. If $T = 0$, $\mathcal{V}_0(g_2) \geq \mathcal{V}_0(g_1)$ for any g_1, g_2 as both sides of the inequality are nil. As an induction hypothesis, assume that $\mathcal{V}_T(g_2) \geq \mathcal{V}_T(g_1)$ for all g_1, g_2 such that $g_2 \succeq_d g_1$. We show that $\mathcal{V}_{T+1}(g_2) \geq \mathcal{V}_{T+1}(g_1)$. If the optimal strategy at belief g_1 is to quit immediately, then $\mathcal{V}_{T+1}(g_1) = \hat{V}^F$ and the inequality is trivially satisfied. Otherwise,

$$\mathcal{V}_{T+1}(g_2) - \mathcal{V}_{T+1}(g_1) \geq \mathbb{E}_{g_2}[\theta] - \mathbb{E}_{g_1}[\theta] + \delta \int \mathcal{V}_T(\beta(g_2,y)) f_{g_2}(y) - \mathcal{V}_T(\beta(g_1,y)) f_{g_1}(y)\,\mathrm{d}y.$$

The first term is nonnegative because $g_2 \succeq_d g_1$ implies $\mathbb{E}_{g_2}[\theta] \geq \mathbb{E}_{g_1}[\theta]$. To establish the nonnegativity of the integral, rearrange it as:

$$\int \mathcal{V}_T(\beta(g_2,y))(f_{g_2}(y) - f_{g_1}(y))\,\mathrm{d}y + \int (\mathcal{V}_T(\beta(g_2,y)) - \mathcal{V}_T(\beta(g_1,y))) f_{g_1}(y)\,\mathrm{d}y.$$

By Claim 2 and the induction hypothesis, $\mathcal{V}_T(\beta(g_2,y))$ is increasing in y. By Claim 3, $f_{g_2} \succeq f_{g_1}$, so, by the definition of continuous domination, the first term is nonnegative. For the second term, note that $\beta(g_i, y) = \beta(\beta(g_0, y_i), y) = \beta(\beta(g_0, y), y_i)$ for $i = 1, 2$. Hence, by Claim 2, $\beta(g_2, y) \succeq_d \beta(g_1, y)$, so by the induction hypothesis $\mathcal{V}_T(\beta(g_2,y)) \geq \mathcal{V}_T(\beta(g_1,y))$, and the second term is also nonnegative. Consequently, $\mathcal{V}_{T+1}(g_2) \geq \mathcal{V}_{T+1}(g_1)$.

The induction argument establishes that for any finite T and $g_2 \succeq_d g_1$, $\mathcal{V}_T(g_2) \geq \mathcal{V}_T(g_1)$. At the same time, by Theorem 2.5.1 of Berry and Fristedt (1985), $\mathcal{V}_T \to \mathcal{V}$ as $T \to \infty$. Thus, $\mathcal{V}(g_2) \geq \mathcal{V}(g_1)$. The result follows.

□

Proof of Proposition 1. The case in which $w^F = V^A$, $\alpha_0 = 0$ is presented in the main text. Consider the case in which $w^F = 0$. By the proof of Lemma 4, if the agent's payoff from following recommendations is at least V^A and $V^E = V^A + V^P \leq \hat{V}^A$, then his incentives are satisfied. Under Assumption 2, the latter condition can be written as:

$$\alpha \leq \frac{V^A}{V^A + V^P}.$$

It is left to show that there exists such an α that under the surplus-efficient quitting policy τ^E delivers the payoff V^A to the agent. By Lemma 1, the agent's payoff can be written as:

$$U^A\left(\overline{m}, \tau^E\right) = \mathbb{E}_{\overline{m},\tau^E}\left[(1-\delta)\sum_{t=0}^{\tau-1}\delta^t y_t + \delta^\tau \hat{V}^A\right] \times \alpha = V^A + \mathbb{E}_{\overline{m},\tau^E}\left[(1-\delta)\sum_{t=0}^{\tau-1}\delta^t\left(\alpha y_t - V^A\right)\right].$$

At $\alpha = 0$, the payoff is:

$$V^A - \mathbb{E}_{\overline{m},\tau^E}\left[(1-\delta)\sum_{t=0}^{\tau-1}\delta^t V^A\right] \leq V^A,$$

where the inequality holds because $V^A > 0$. At $\alpha = V^A/(V^A + V^P)$, the payoff is:

$$V^A + \frac{V^A}{V^A + V^P}\mathbb{E}_{\overline{m},\tau^E}\left[(1-\delta)\sum_{t=0}^{\tau-1}\delta^t\left(y_t - \left(V^A + V^P\right)\right)\right] \geq V^A,$$

where the inequality holds because of Assumption 2 and the fact that a surplus-efficient policy delivers a total payoff that is at least as large as the total outside option:

$$\mathbb{E}_{\overline{m},\tau^E}\left[(1-\delta)\sum_{t=0}^{\tau-1}\delta^t\left(y_t - \left(V^A + V^P\right)\right)\right] + V^A + V^P \geq V^A + V^P.$$

Furthermore, $U^A\left(\overline{m}, \tau^E\right)$ is linear in α. Hence, by the intermediate value theorem, there exists $\alpha^* \in \left[0, V^A/(V^A + V^P)\right]$ such that the agent's payoff is equal to V^A.

□

Proof of Claim 1. By Theorem 2, the equilibrium evaluation policy is cutoff and the wage profile $\{W_t\}_{t=0}^{\infty}$ is deterministic and weakly increasing. Because the technology is strongly regular, the principal's beliefs μ_t have full support over $(0,1)$ at all t, and hence, the quitting probability is strictly positive in every period. Thus, there is strictly positive selection in every period, and W_t increases strictly in t. Because the principal's belief is a sufficient statistic for the recommendation policy, it is a sufficient statistic for the agent's continuation

value. Define by $\hat{U}_t^A(\mu_t)$ the agent's continuation value at time t if the principal's belief is μ_t:

$$\hat{U}_t^A(\mu_t) = \mathbb{E}_\tau\left[(1-\delta)\sum_{t'=t}^{\tau-1}\delta^{t'}\left(W_{t'}-V^A\right)\right]+V^A.$$

Denote the continuation distribution of quitting times by $f(\tau \mid \mu_t)$. It is continuously and strictly decreasing in μ_t at all subsequent periods: every continuation path of the principal's beliefs continuously and strictly increases in μ_t; as the technology is strongly regular, it means that there is an additional strictly positive measure of paths that stay above the optimal cutoffs. It follows that $\hat{U}_t^A(\mu_t)$ is continuous in μ_t over $[0,1]$. Moreover, the maximal value of $\hat{U}_t^A(\mu_t)$ is achieved at $\mu_t = 1$, because at this belief the agent never quits and ensures the maximal possible wage stream (which is preferred to securing an option V^A by incentive compatibility).

Proceed by constructing the thresholds $\underline{\kappa}_t$. Start with the first period. Consider the agent's continuation value if he is recommended to stay and follows the recommendations:

$$\tilde{U}_1^A = \mathbb{E}\left[\hat{U}_1^A(\mu_1) \mid m^1 \equiv \text{"stay"}\right].$$

As the wage profile is strictly increasing, $\tilde{U}_1^A > \tilde{U}_0^A \geq V^A$. As $\hat{U}_1^A(\mu_1)$ obtains its maximal value at $\mu_1 = 1$, is continuous, and $\tilde{U}_1^A > V^A$, it follows that $\hat{U}_1^A(\mu_1) > V^A$ for all $\mu_1 > x_1$ for some $x_1 < 1$. Moreover, as the technology is strongly regular, μ_1 is distributed continuously over $(0,1)$ so the probability of μ_1 being greater than x_1 is positive for all $x_1 < 1$ and continuously decreases to 0 as x_1 increases to 1. Hence, there exists a threshold $\underline{\kappa}_1 < 1$ such that for any $\kappa_1 \geq \underline{\kappa}_1$:

$$\mathbb{E}\left[\hat{U}_1^A(\mu_1) \mid \mu_1 > \kappa_1, m^1 \equiv \text{"stay"}\right] > \mathbb{E}\left[\hat{U}_1^A(\mu_1) \mid \mu_1 < \kappa_1, m^1 \equiv \text{"stay"}\right] > V^A.$$

Take this threshold as the partition threshold for period 1. The inequalities ensure that the agent is willing to follow the recommendation to stay. Recommendations to quit remain incentive compatible because the quitting recommendation occurs in the region of belief in which the agent himself would prefer to quit.

Given a sequence of thresholds $\underline{\kappa}_{t'}$ for $t' < t$, we now construct a threshold $\underline{\kappa}_t$. Consider a collection of continuation values at period $t-1$, $\mathbb{E}\left[\hat{U}_{t-1}^A(\mu_{t-1}) \mid s^{t-1}, m^{t-1} \equiv \text{"stay"}\right]$, where s^{t-1} are past realizations of performance leaks. By incentive compatibility, all these values are strictly greater than V^A. By the same argument as in the proof of Lemma 4, continuation values conditional on recommendations to stay increase and do so strictly because of the

strictly positive quitting probability:

$$\mathbb{E}\left[\hat{U}_t^A\left(\mu_t\right) \mid s^{t-1}, m^t \equiv \text{"stay"}\right] > \mathbb{E}\left[\hat{U}_{t-1}^A\left(\mu_{t-1}\right) \mid s^{t-1}, m^{t-1} \equiv \text{"stay"}\right].$$

Hence, we can find 2^{t-1} thresholds $\kappa_t\left(s^{t-1}\right)$ such that the recommendation to stay whenever $\mu_t \geq \kappa_t\left(s^{t-1}\right)$ is incentive compatible after leak history s^{t-1}. Setting $\underline{\kappa}_t = \max_{s^{t-1}} \kappa_t\left(s^{t-1}\right)$ creates an incentive-compatible policy.

□