Learning-by-Employing: The Value of Commitment under Uncertainty[*]

Braz Camargo

Sao Paulo School of Economics-FGV

Elena Pastorino

Federal Reserve Bank of Minneapolis
and University of Minnesota

ABSTRACT _____

We analyze commitment to employment in an environment in which an infinitely lived firm faces a sequence of finitely lived workers who differ in their ability to produce output. A worker's ability is initially unknown to both the worker and the firm. A worker's effort affects the information on ability conveyed by his performance. We characterize equilibria and show that they display commitment to employment only when effort has a persistent but delayed impact on output. In this case, by providing insurance against early termination, commitment to employment encourages workers to exert effort, thus improving the firm's ability to identify workers' talent. The incentive value of commitment to retention helps explain the use of probationary appointments in environments in which there is uncertainty about individual ability.

Keywords: Career concerns; Learning; Retention; Commitment
JEL Classification: C73, D21, D83, J41

1 Introduction

It is widely recognized that talented individuals can be identified only through careful selection. As a result, firms usually employ a range of methods to evaluate job candidates. Standard practices include the review of resumes, the evaluation of references, various forms of testing, and interviewing. As part of their hiring process, many firms also rely on probationary appointments—temporary contracts that grant employment for a prespecified period of time—in order to determine whether new workers are suited to their jobs.[1]

The use of probationary appointments is common in management consulting, the legal profession, academia, and government bureaucracies.[2] In all of these instances, a worker's output can critically depend on his skill, but often the qualities that distinguish a successful individual are only revealed over time. Why then should an employer commit to retain a worker of uncertain ability for a certain period of time rather than decide on employment as information on performance is acquired? Intuitively, if performance on the job provides information about ability, and thus is a signal of future productivity, then the flexibility to replace workers whose performance is unsatisfactory should be valuable to a firm. In this paper, we show that when the ability of new hires is uncertain, an employer might nevertheless benefit from committing not to dismiss workers early in their careers.

The reason why commitment to employment can be beneficial is twofold. First, the quality of the information that a worker's performance provides about ability can be affected by a worker's behavior. For instance, whether a researcher is successful at a project depends not only on his talent but also on the effort spent by the researcher on the project. Likewise, whether a restructuring project is successful in addressing the needs of a client firm depends both on the ability

---

[1]Probationary periods are also understood as the stage at the beginning of an employment relationship during which an employer has greater discretion to dismiss workers; see Loh (1994), for example. This type of employment arrangement is common in unionized industries. In this paper, a probationary period is a period of time, ranging from a few months to several years, during which a firm is committed to employing a worker: the firm can dismiss the worker only at the end of this period.

[2]The literature on labor markets in professional service industries provides evidence that probationary appointments, usually coupled with so-called up-or-out rules, are often used and entail an explicit commitment to retain a newly hired worker for a prespecified period of time. See, for instance, Carmichael (1988), Gibbons (1998), Wilkins and Gulati (1998), Rebitzer and Taylor (2007), and Ghosh and Waldman (2010). We discuss the evidence on the use of probationary appointments in more detail in Section 6.

of the consultant assigned to the project and on his effort. Second, it may take time before effort affects the informational content of performance. For example, the completion of a research project normally requires a few years. Hence, it may take time for effort to have a perceivable impact on a researcher's output, and thus on the information about the researcher's ability that can be acquired based on his output. Analogously, even if progress on a consultancy project might be measurable on a month-to-month basis, the final outcome is usually the best indicator of the ability of a consultant. In such circumstances, the prospect of an early dismissal may discourage a worker from exerting effort, thus reducing the informativeness of his performance. In this case, offering a probationary appointment may be valuable to a firm if insurance against early failure encourages workers to produce informative signals about their ability.

Formally, we consider a labor market in which an infinitely lived firm faces a constant inflow of finitely lived workers. At any date, the firm can employ at most one worker. We model probationary appointments as a short-term commitment to employment on the part of the firm and assume that incentive pay is not feasible. We do not consider incentive pay for two reasons. First, as mentioned above, the use of probationary appointments is common in academia and bureaucracies, where the use of incentive pay is quite limited. So, a framework where incentive pay is not possible makes for a natural benchmark. Second, abstracting from incentive pay allows us to analyze the trade-offs involved in the use of probationary appointments in a more transparent way.

We assume that workers differ in their ability, either high or low, to produce output and that a worker's ability is initially unknown to both the worker and the firm. The performance of an employed worker also depends on his choice of effort. More precisely, effort increases the probability of good performance only if the worker is of high ability. Hence, when a worker exerts effort, good performance is a more precise signal of high ability. Note that our environment differs from a standard moral hazard setting in that the firm benefits from effort in two ways. As in a standard moral hazard setting, effort increases output. Unlike in a standard moral hazard setting, effort also makes performance more informative about ability, which allows the firm to better sort the workers it employs.

Each worker in the market has an outside option whose value increases with the worker's

reputation, which we model as the firm's belief that the worker is of high ability.[3] Then, workers of higher reputation are perceived to be more productive, but can be employed only at a higher wage. Since effort affects a worker's reputation, and thus wage, a worker's concern for his future career influences his choice of effort. Indeed, a worker who has yet to prove his ability is rewarded with a wage increase after good performance, as good performance increases reputation. Thus, a worker who has yet to prove his ability has an incentive to exert effort. The worker's value to the firm also increases with the worker's reputation. Hence, the firm faces an opportunity cost by retaining a worker whose initial performance is poor: such a worker is less likely to be of high ability than a worker who is new to the market.

The firm benefits from offering probation to a new worker when commitment to employment strengthens the worker's incentive to exert effort. Commitment to employment, however, is costly because it prevents the firm from dismissing the worker if he performs poorly. In our analysis we identify circumstances under which the use of probation can be beneficial to a firm by contrasting two cases that are distinguished by the effect of effort on output. In the first case, our benchmark, the impact of effort on output is independent and identical over time. We refer to this case as the *IID case*. In the second case, effort has a delayed impact on output. We refer to the second case as the *non-IID case*. The non-IID case captures situations in which it takes time for effort to affect the informativeness of performance.

We begin by showing that the firm does not benefit from offering probation in the IID case. In this case, a new worker's concern for his future career is sufficient to induce him to exert effort. In particular, commitment to employment does not provide any additional incentive for effort. Moreover, by offering probation, the firm is forced to employ a worker whose initial performance is poor, and such a worker is less attractive to the firm than a new worker.

The firm can benefit from offering probation to new hires in the non-IID case, though. Commitment to employment is valuable in the non-IID case because now the incentive problem under consideration is compounded by the time separation of costs and benefits typical of investment problems. Specifically, when effort affects mostly future output, the worker can gain from invest-

---

[3]Note that our use of the word "reputation" differs from its use in the literature on repeated games, where an individual's reputation refers to the belief about his behavior in the game.

ing in his reputation only if he is guaranteed to participate in its return, which can occur only if employment lasts until the impact of effort on output materializes. Although the decision to retain a newly hired worker after poor performance is ex ante optimal for the firm, this decision is not ex post optimal, however, because a worker whose initial performance is poor is less profitable to the firm than a new worker. Hence, when it offers probation, the firm overcomes its incentive to dismiss underperforming workers, thereby inducing new hires to exert effort and generate more precise information about their ability. In other words, the firm can benefit from probation in the non-IID case, since commitment to employment solves a time-inconsistency problem. Importantly, we show that in this case, the use of probation can be justified by informational considerations alone. Given that effort makes performance more informative about ability, the firm can gain from commitment to employment simply because this commitment improves the firm's ability to identify workers' talent.

The rest of the paper is organized as follows. We discuss the related literature in the remainder of this section. We introduce the model in Section 2, derive some auxiliary results in Section 3, analyze the IID case in Section 4, and examine the non-IID case in Section 5. We present new evidence on the use of probationary appointments and discuss some of our modeling choices in Section 6. We conclude in Section 7. Omitted proofs and details are in the appendices.

Related Literature

In our model, a worker's concern for his future career influences his choice of effort. The idea that career concerns can induce workers to exert effort even in the absence of explicit incentives for performance was first formalized by Holmström (1999); see also Scharfstein and Stein (1990). A few papers have studied the interplay between explicit incentive contracts and career concerns. For instance, Gibbons and Murphy (1992) analyze the optimal combination of incentive pay and career concerns incentives.[4] Mukherjee (2008) investigates the substitutability between incentives from career concerns and incentives from implicit bonus contracts when a firm can disclose information about a worker's productivity to prospective employers. Our work differs from the existing literature on career concerns (and, more generally, from the literature on dynamic moral hazard) in

---

[4]Andersson (2002) extends the analysis of Gibbons and Murphy to the case in which contracts are unobservable.

that it considers a framework in which the key decision for the firm is whether to retain a worker.[5]

The retention problem the firm faces in our setting is an example of a multi-armed bandit, that is, the sequential sampling problem of a decision maker choosing between alternatives with uncertain rewards. In our case, the alternatives correspond to the different workers the firm can employ.[6] Jovanovic (1979, 1984) is the first application of the multi-armed bandit framework to the analysis of employer learning in labor markets. More recent papers are Harris and Weiss (1984), Felli and Harris (1996), Moscarini (2005), and Eeckhout and Weng (2010).

Differently from a standard bandit problem, in our environment rewards are endogenous: a worker's choice of effort depends on the firm's employment decisions. Manso (2011) also analyzes a contracting environment in which a firm faces a bandit problem with endogenous rewards. In his setting, the problem of the firm is to motivate an agent to innovate, that is, to select an action with unknown payoffs that could be superior to an action with known payoffs. Our environment differs from Manso's environment in that the uncertainty is about an agent's ability rather than about the payoffs from an agent's actions.

The fact that rewards are endogenous is crucial for our analysis. The firm can never gain from offering commitment to employment if the incentive problem is absent, that is, if a worker's choice of effort is not affected by the firm's behavior. Indeed, in the absence of incentive problems, the firm faces a standard multi-armed bandit. In this case, since rewards are exogenous, any strategy that involves commitment to employment, that is, commitment to the use of a given arm, can be replicated by a strategy that does not involve commitment. So, commitment to employment is not beneficial.

Bull and Tedeschi (1989) and Wang and Weiss (1998) provide an alternative explanation for the use of probationary appointments. They show that when workers are privately informed about their ability, probation can be used as a mechanism to induce workers to self-select into jobs according to their skill.[7] More precisely, Bull and Tedeschi model probation as a period during which a firm

---

[5]Banks and Sundaram (1998) study the problem of agent retention by a long-lived principal when there is both moral hazard and adverse selection, contracting is not possible, and agents live for two periods.

[6]See Berry and Fristedt (1985) for an exposition of the theory of multi-armed bandits. In the non-IID case, the firm faces a so-called experimentation problem with signal dependence. See Datta, Mirman, and Schlee (2002) for an analysis of such a problem.

[7]See Guasch and Weiss (1981, 1982) for labor market models of adverse selection.

commits to monitoring a worker's choice of effort. By adjusting the length of the probationary period, firms can discourage low-ability workers, for whom effort is more costly, from accepting employment offers. Wang and Weiss model probation as a period during which a firm tests newly hired workers. By letting wages and retention decisions depend on test results, firms can induce low-ability workers, who are less likely to pass the selection test, to reject employment. However, in many of the settings in which probation is used, such as academia, it is likely that workers do not know their ability initially and only learn about it over time. A natural benchmark in this case is an environment in which a firm and a worker are initially symmetrically uninformed about the worker's ability, the case we consider.

Our paper belongs to the literature on time-inconsistency and internal labor market practices. Kahn and Huberman (1988) study the problem of workers who make a nonverifiable investment in firm-specific human capital. Although rewarding such an investment is ex ante optimal for a firm, the decision to do so is not ex post optimal. Anticipating this, workers underinvest in human capital.[8] Kahn and Huberman show that up-or-out contracts specifying that a worker is fired if not promoted can solve this double moral hazard problem on the part of workers and firms. Waldman (1990) extends Kahn and Huberman's analysis to the case of general human capital.[9] Prendergast (1993) shows how promotion to different jobs or tasks can replace the use of up-or-out contracts as a mechanism to induce the acquisition of firm-specific human capital.

Milgrom and Roberts (1988) study a time-inconsistency problem that arises when workers can engage in inefficient influence activities. Ex post firms have an incentive to base the promotion of a worker on his evaluation. This, however, creates an incentive for individuals to expend effort in manipulating the evaluation process. Milgrom and Robert's analysis shows that firms can sometimes benefit from committing to not always promoting the most qualified workers, so as to weaken the incentive to engage in influence activities. Waldman (2003) analyzes a time-inconsistency problem that arises from the dual role of promotions: they serve both to reward performance and to

---

[8]Our paper is also related to the literature on the so-called hold-up problem in that we analyze contractual remedies to underinvestment in relation-specific capital, here information about ability.

[9]Carmichael (1988) and O'Flaherty and Siow (1992) provide alternative explanations for the use of up-or-out rules. Carmichael shows how the institution of tenure in academia can induce incumbent professors to hire researchers who are potentially more talented than themselves. O'Flaherty and Siow analyze a model of on-the-job screening and firm growth, and show that the optimal retention decision is an up-or-out rule.

efficiently assign workers to tasks. Since at the time of a promotion decision, only the assignment motive matters for a firm, the use of promotions as an incentive device is undermined. Waldman shows that the practice of favoring internal candidates for promotion can be interpreted as a solution to this time-inconsistency problem. Finally, Ghosh and Waldman (2010) study the choice between standard promotion practices and up-or-out contracts when promotions serve the dual role just described. They show that up-or-out contracts are superior to standard promotion practices when the level of a worker's firm-specific human capital is low.

Our paper differs from the previous literature on time-inconsistency and internal labor market practices in that we consider how probationary appointments can be used to induce workers to generate nonverifiable information about their ability. Our goal informs a number of modeling choices. First, all of the above-mentioned papers consider two-period settings, so that retention amounts to permanent retention. We, instead, assume that workers live for at least three periods, so that the notion of short-term commitment to employment is meaningful. Second, unlike the papers that study how promotion practices affect incentives (Prendergast (1993), Waldman (2003), and Ghosh and Waldman (2010)), we assume there is a single task and many workers who are heterogeneous in their ability to perform this task. Thus, the firm faces a bandit problem. Finally, the moral hazard problem we study is nonstandard. First, in our setting, effort affects not only output but also the informativeness of performance.[10] Second, in the non-IID case, a worker's current output only depends on his previous choice of effort.[11]

2 Environment

We consider a labor market with one firm and a countably infinite number of workers. Time is discrete and indexed by $t \geq 1$.

---

[10]In Ghosh and Waldman (2010), a worker's effort affects the mean of the posterior belief about his ability, but not its variance. Thus, unlike in our framework, effort does not make performance more informative about ability.

[11]A few papers have considered the problem of repeated moral hazard with effort persistence. Jarque (2010) provides conditions under which this problem is observationally equivalent to a problem without persistence. Mukoyama and Şahin (2005) study a two-period problem and show that it can be optimal for a principal to perfectly insure an agent in the first period when effort is persistent.

Workers. Workers enter the market sequentially, one in each period. They have a concave and strictly increasing utility function $v : \mathbb{R}_+ \to \mathbb{R}$, live for $T \geq 3$ periods once they enter the market, and discount future utility at rate $\delta \in (0, 1)$.[12] Each worker is either of high (H) or low (L) *ability*. A worker's type, that is, his ability, is unknown to both the worker and the firm. The probability that a worker entering the labor market is of high ability is $\phi_0 \in (0, 1)$. We refer to the firm's belief that a worker is of high ability as the worker's *reputation*.

In each period of employment, a worker can either exert effort ($\overline{e}$), incurring a cost $c > 0$, or no effort ($\underline{e}$), and can produce either high ($\overline{y}$) or low ($\underline{y}$) output. A worker's effort is unobservable and affects his output. We consider two cases. In the first case, our benchmark, a worker's output in a period depends on his current choice of effort. In the second case, a worker's output in a period depends on his choice of effort in the previous period. We refer to the first case as the IID case and to the second case as the non-IID case. Formally, let e and e_- denote a worker's choice of effort in the current and previous periods, respectively. Moreover, let $\Pr\{y|\tau, e, e_-\}$ be the probability, as a function of e and e_-, that a worker of type $\tau \in \{L, H\}$ produces $y \in \{\underline{y}, \overline{y}\}$. Then,

$$\Pr\{\overline{y}|\tau, e, e_-\} = \begin{cases} \alpha + \eta(e, e_-) & \text{if} \quad \tau = H \\ 0 & \text{if} \quad \tau = L \end{cases}, \tag{1}$$

where $\alpha \in (0, 1)$ and $\eta(e, e_-)$ is the increase in the probability that a high-ability worker produces high output when his current choice of effort is $e \in \{\overline{e}, \underline{e}\}$ and his previous choice of effort was $e_- \in \{\overline{e}, \underline{e}\}$. In the IID case, $\eta(e, e_-) \equiv \eta(e)$ with $\eta(\overline{e}) > \eta(\underline{e}) = 0$. In the non-IID case, $\eta(e, e_-) \equiv \eta(e_-)$ with $\eta(\overline{e}) > \eta(\underline{e}) = 0$ and $e_- = \underline{e}$ for age 1 workers; in Appendix B, we show that assuming $e_- = \underline{e}$ for age 1 workers is not crucial for our results. Thus, in the non-IID case, only effort in the previous period affects the probability of high output.[13] For ease of notation, in what follows we let $\eta = \eta(\overline{e})$.

Notice that in both the IID and non-IID cases, a low type worker cannot produce high output. Hence, a worker whose performance is good, that is, a worker who produces high output, reveals that he is of high ability. In Section 6, we show that our results extend to the case in which the

---

[12]The restriction that $T \geq 3$ is to avoid the uninteresting case in which the decision to retain a worker for one additional period after his first period of employment amounts to permanent retention.

[13]The results we obtain in the non-IID case clearly also hold when $\eta(\overline{e}, e_-) - \eta(\underline{e}, e_-)$ is positive but small.

probability that a low type worker produces high output is greater than zero but sufficiently small, so that good performance is a strong signal of high ability.

Observe also from (1) that in both cases, effort increases the probability of good performance only if the worker is of high ability. So, effort increases the likelihood that a high-ability worker reveals his type. In this precise sense, effort makes a worker's performance more informative about his ability. The informational role of effort is central to our analysis.

Workers in the market have an outside option paying a (non-negative) wage w_R that depends on their reputation. A worker who collects his outside option can no longer be hired by the firm. This feature captures the fact that in practice it may not always be profitable or possible for a firm to rehire a worker who has previously separated from it. We comment on this point in Section 6. We assume that $w_R(\phi) = w_R(\phi_0)$ for all $\phi \leq \phi_0$. So, as long as a worker fails to produce high output, his outside option is constant at $\underline{w} := w_R(\phi_0)$, but it increases to $\overline{w} := w_R(1)$ the first time he produces high output. The restriction that a worker's outside option cannot decrease below $\underline{w}$ captures situations in which a worker can seek employment in an alternative labor market where a different type of ability is valued. We discuss our modeling of the labor market in Section 6.

The firm.   The firm is infinitely lived and risk neutral. For ease of notation, we assume that the firm discounts future payoffs at the same rate as the workers. Our results extend to the case in which the firm is at least as patient as the workers. The firm can employ at most one worker in each period, and its flow payoff when it does not employ a worker is $\Pi < \underline{y} - \underline{w}$.[14] So, the firm prefers employing a worker it knows is of the low type rather than not employing any worker. We normalize flow payoffs to the firm by $(1 - \delta)$.

Besides an *incumbent*, a worker employed by the firm in the previous period, the only other worker the firm can employ in a given period is the available age 1 worker. Following the career concerns literature, we assume that wages in a period cannot be conditioned on that period's output. More precisely, at the beginning of each period, the firm can offer a worker to pay him a (non-negative) wage w^0 at the end of the period if he accepts employment. The firm can also commit to make minimum one-period (non-negative) wage offers to a worker of age $k \leq T - 1$ for the

---

[14]We can extend our analysis to the case in which the firm has a finite number of vacancies.

next $q \in \{1, \ldots, T - k\}$ periods, where these minimum offers can depend on the worker's age. Hence, an offer to a worker is a list $(w^0, (q, \{w^s\}_{s=1}^q))$ consisting of a one-period wage offer w^0, the number q of subsequent periods in which the firm is committed to make one-period wage offers to the worker, and the schedule $\{w^s\}_{s=1}^q$ of future minimum one-period wage offers, where w^s is the minimum one-period wage offer s periods into the future.[15] Notice that if the firm offers $(w^0, (q, \{w^s\}_{s=1}^q))$ with $q \geq 1$ to a worker of age k, then it must propose $(w^0, 0)$ with $w^0 \geq w^s$ to him when he is of age $k + s$.[16] We say the firm offers *probation* to an age 1 worker when it offers him $(w^0, (q, \{w^s\}_{s=1}^q))$ with $q \geq 1$.

Let $y(\phi, \xi) = \phi \xi \overline{y} + (1 - \phi \xi) \underline{y}$ be the expected output of a worker of reputation ϕ when the probability that he produces high output is ξ if he is of the high type. The following restriction is a maintained assumption:

(A1) $y(1, \alpha) - \overline{w} > y(\phi_0, \alpha + \eta)$.

Recall that in the IID case, η is the increase in the probability of high output in a given period when a worker exerts effort in that period, whereas in the non-IID case, η is the increase in the probability of high output in a given period if the worker exerted effort in the previous period.

Note that the firm cannot hire a worker known to be of the high type for less than $\overline{w}$ and the reputation of an age $k \geq 2$ worker who has never produced high output is smaller than ϕ_0. Thus, (A1) implies that the firm's flow payoff from employing a worker who has revealed himself to be of high ability is greater than the firm's flow payoff from employing a worker who has not revealed himself to be of high ability.[17] Observe that (A1) is satisfied if high-ability workers are sufficiently scarce, that is, if ϕ_0 is sufficiently small.

Assumption (A1) also implies that the firm prefers a worker of the high type to an age 1 worker,

---

[15] As we show in the next section, in equilibrium the firm never makes an offer that it expects to be rejected. Thus, commitment to future wage offers amounts to commitment to employment: if the firm makes an offer with $q \geq 1$ to a worker, then the worker remains employed at least until the end of the commitment period.

[16] Thus, the firm cannot offer to extend the length of the commitment period once an offer with commitment is accepted. This assumption is without loss of generality, since any equilibrium in which the firm extends the commitment period for a worker is outcome equivalent to an equilibrium in which the firm offers a longer commitment period to the worker in the first place and does not extend this commitment afterward.

[17] As it turns out, age 1 workers may be willing to work for a zero wage, in which case (A1) is also necessary for the firm to prefer workers who have proved themselves to be of high ability to workers who have not.

even if the high type worker exerts low effort. This assumption plays an important role in our analysis because it allows us to focus on the problem of interest, which is whether probationary appointments can help the firm identify high-ability workers.

Timing. The sequence of events in a period is as follows. If the firm has no incumbent, then it either collects its outside option or makes an offer to the available age 1 worker. If the firm has an incumbent to which it is committed to make a one-period wage offer, then it makes him such an offer. If the firm has an incumbent but is not committed to make him an offer, then it can collect its outside option, make an offer to the incumbent, or make an offer to the available age 1 worker. The worker who receives an offer decides whether to accept it or not. In case the worker accepts the offer, he chooses how much effort to exert, output is realized, and the firm pays him the wage that it promised at the beginning of the period. A worker collects his outside option if he either does not receive an offer or does receive an offer and rejects it. Likewise, the firm collects its outside option if its offer is rejected.

Equilibrium. Let $\Sigma_w(t)$ be the set of behavior strategies for a worker who enters the market in period t and Σ_f be the set of behavior strategies for the firm. We assume that workers do not observe the history of play before they enter the market. Thus, we let $\Sigma_w(t) \equiv \Sigma_w$. Note that since workers do not observe the history of play before they enter the market, any deviation by the firm can at most affect the behavior of the worker it currently employs. Therefore, even though the firm is infinitely lived, it cannot develop a reputation for a particular behavior.

A strategy profile for workers is a map $\sigma_w : \mathbb{N} \to \Sigma_w$, where $\sigma_w(t)$ is the behavior strategy of the worker who enters the market in period t. A strategy profile (σ_w, σ_f) is worker symmetric if $\sigma_w(t)$ is independent of t. We restrict attention to worker-symmetric perfect Bayesian equilibria. In what follows, we use the expression "in equilibrium" as a shorthand for "in every equilibrium".

3 Preliminaries

This section consists of two parts. First, we discuss the option value of employment. Then, we present some results that are useful for the analysis that follows.

3.1 The Option Value of Employment

An important feature of our environment is that a worker of age $T - 1$ or less who has not revealed himself to be of high ability is willing to work for less than his outside option. To see why, consider such a worker, and let $k \leq T - 1$ be his age and $\pi \leq \phi_0$ be his (private) belief that he is of the high type. Since an option for the worker is to accept employment, exert no effort, and collect his outside option when of age $k + 1$, the worker obtains a payoff of at least

$$v(w) + \pi\alpha\delta(1 - \delta)^{-1}(1 - \delta^{T-k})v(\overline{w}) + (1 - \pi\alpha)\delta(1 - \delta)^{-1}(1 - \delta^{T-k})v(\underline{w}) \qquad (2)$$

when he accepts a one-period wage of w. Indeed, if the worker exerts no effort, then he produces high output with probability $\pi\alpha$ and low output with the remaining probability. Moreover, if the worker collects his outside option when of age $k + 1$, then his payoff is $(1 - \delta)^{-1}(1 - \delta^{T-k})v(\overline{w})$ if he produces high output when of age k and $(1 - \delta)^{-1}(1 - \delta^{T-k})v(\underline{w})$ otherwise. Now observe that if $w = \underline{w}$, then (2) is greater than $(1 - \delta)^{-1}(1 - \delta^{T-k+1})v(\underline{w})$, which is the worker's payoff from taking his outside option when of age k. Hence, the worker is willing to accept (an offer with) a one-period wage smaller than his outside option. Intuitively, by accepting employment at the firm, the worker has the opportunity to prove that he is of high ability and thus to increase his future compensation.

It turns out that in equilibrium, the workers who are willing to sacrifice the most in order to work for the firm are age 1 workers. Intuitively, among the workers who have not revealed themselves to be of high ability, age 1 workers are the most likely to do so. Age 1 workers are also those who benefit from revealing that they are of high ability for the greatest number of periods. Thus, an age 1 worker not only holds more promise than an age $k \geq 2$ worker who has only produced low output, and so has a lower reputation, but also is less expensive to employ.

3.2 Auxiliary Results

Here, we present some results that are useful for the analysis that follows. First, note that a straight-forward consequence of the assumption that $\Pi < \underline{y} - \underline{w}$ is that in equilibrium the firm makes an offer in every period and never makes an offer that it anticipates will be rejected with positive probability. See Appendix A for a proof of this result.

We now prove that a worker has no incentive to exert effort if there is no uncertainty about his ability. Hence, whenever the firm makes an offer to an incumbent known to be of the high type, the one-period wage it offers is the lowest possible.

Lemma 1. *Suppose the firm has an incumbent it knows is of the high type, and let w' be the smallest one-period wage the firm can offer him if it is committed to doing so. The following holds in equilibrium: (i) the firm never offers the incumbent a one-period wage greater than $\max\{\overline{w}, w'\}$; (ii) if the firm makes an offer to the incumbent, then it never commits to future minimum one-period wage offers greater than $\overline{w}$; (iii) the incumbent never exerts effort.*

The proof of Lemma 1 is in Appendix A. A sketch of the proof is as follows. Consider a worker known to be of the high type who is in his last period of employment. As in a standard career concerns model, the only incentive for him to exert effort is the variation of his future payoff in his output. This variation, however, is zero, because the worker's reputation does not change with his output if his ability is known. Thus, exerting no effort is uniquely optimal for him. This, in turn, implies that the firm has no incentive to offer a one-period wage greater than $\max\{\overline{w}, w'\}$ to the worker. The desired result now follows from a backward induction argument.

The next result we establish follows from the fact that workers use symmetric strategies. Let $V(h|\sigma)$ denote the firm's lifetime payoff after a history h when the strategy profile under play is σ. Now suppose there exist histories h and h' for the firm after which the firm makes an offer to the available age 1 worker with $V(h'|\sigma) > V(h|\sigma)$. Consider then the deviation for the firm where it behaves after h as if h' had happened. Since workers follow symmetric strategies and do not observe the history of play before they enter the market, this deviation increases the firm's payoff after h by $V(h'|\sigma) - V(h|\sigma)$. So, σ is not an equilibrium. This argument proves that in equilibrium the firm's (expected present discounted) continuation payoff from hiring an age 1

worker is independent of calendar time. Lemma 2 summarizes this discussion.

Lemma 2. *Let σ be an equilibrium. Then, $V(h|\sigma) = V(h'|\sigma)$ for any two histories h and h' for the firm after which the firm makes an offer to the available age 1 worker.*

We conclude this section by characterizing the firm's retention decision. A consequence of assumption (A1) is that the highest flow payoff the firm obtains is when it employs a high-ability worker at wage $\overline{w}$. This fact in turn implies that the firm always retains an incumbent it knows is of high ability; see Appendix A for a proof. For convenience, in the remainder of the paper we say the firm "offers w" to a worker whenever it makes an offer in which the one-period wage is w.

Lemma 3. *Suppose the firm has an incumbent it knows is of high ability, and let w' be the smallest one-period wage the firm can offer him if it is committed to doing so. In equilibrium, the firm offers $\max\{\overline{w}, w'\}$ to this worker. Moreover, on the equilibrium path, a worker known to be of high ability always accepts such an offer.*

4 The IID Case

In this section, we investigate the role of commitment to employment when, conditional on a worker's type, the impact of effort on output is identical and independent over time. Our main result in the IID case is that the firm does not benefit from commitment to employment.

In order to establish our main result, consider an age 1 worker who is not retained if his performance is poor, that is, if he produces low output. Lemma 1 implies that if the worker produces high output, then his wage in every subsequent period is $\overline{w}$. Hence, $\phi_0\eta\delta(1-\delta)^{-1}(1-\delta^{T-1})[v(\overline{w})-v(\underline{w})]$ is the worker's expected lifetime payoff gain from exerting effort: $\phi_0\eta$ is the increase in the probability that the worker produces high output if he exerts effort, and $\delta(1-\delta)^{-1}(1-\delta^{T-1})[v(\overline{w})-v(\underline{w})]$ is the increase in the worker's lifetime payoff if he produces high output. Thus, when

$$\phi_0\eta\delta(1-\delta)^{-1}(1-\delta^{T-1})[v(\overline{w}) - v(\underline{w})] \geq c, \tag{3}$$

an age 1 worker who is not retained after low output has an incentive to exert effort. We can now state the main result of this section.

Proposition 1. *Suppose that* (3) *holds. There exists an equilibrium* σ^* *in which the firm dismisses age 1 workers after low output and these workers exert effort. Moreover, the firm's payoff in any equilibrium in which it retains an age 1 worker after low output is strictly smaller than its payoff under* σ^*. *Suppose now that* (3) *does not hold. The firm dismisses age 1 workers after low output in every equilibrium.*

An immediate consequence of Proposition 1 is that commitment to employment is not beneficial to the firm. The proof of Proposition 1 is in Appendix A. Here, we provide a sketch of the proof. Suppose first that (3) holds, so that an age 1 worker who is dismissed after low output has an incentive to exert effort. In order to show that there exists an equilibrium σ^* in which the firm dismisses age 1 workers after low output, we just need to verify that doing so is optimal for the firm. The optimality of dismissing age 1 workers after poor performance follows from the fact that an age $k \geq 2$ worker who has not revealed himself to be of high ability is less attractive to the firm than an age 1 worker who exerts effort. Indeed, such an age $k \geq 2$ worker is worse than an age 1 worker in flow payoff terms, since his expected output is lower and, as discussed in Subsection 3.1, it is costlier to induce his participation. Moreover, such an age $k \geq 2$ worker is worse than an age 1 worker in continuation payoff terms for the following two reasons: the likelihood that such an age $k \geq 2$ worker reveals himself to be of high ability is lower, and in case he proves himself to be of high ability, he can work at the firm for a shorter period of time. This argument also implies that the firm's payoff in any equilibrium in which it retains an age 1 worker after low output is strictly smaller than its payoff under σ^*.

Consider now the case in which (3) does not hold, so that the threat of dismissal after poor performance is not sufficient to induce an age 1 worker to exert effort. The question then is whether retention after low output can provide an age 1 worker with the incentive to exert effort. As we discuss below, it turns out that the answer is no. More generally, no worker exerts effort when (3) is not satisfied. Thus, commitment to employment does not improve incentives. Consequently, it is optimal for the firm to dismiss age 1 workers after they perform poorly.

The argument for why workers do not exert effort when (3) is not satisfied is as follows. First notice by Lemma 1 that we only need to consider workers who have not produced high output. Next

16

note that a worker of age T who has not revealed himself to be of high ability has no incentive to exert effort. Consider then a worker of age $k \in \{1, \ldots, T-1\}$ who has not revealed himself to be of high ability, and let $\pi \leq \phi_0$ be his belief that he is of high ability. Recall that $\pi < \phi_0$ if $k \geq 2$. The worker's incentive-compatibility constraint for effort exertion is

$$-c + \pi(\alpha + \eta)\delta R(\overline{y}, \overline{e}|\pi, k) + [1 - \pi(\alpha + \eta)]\delta R(\underline{y}, \overline{e}|\pi, k)$$
$$\geq \pi\alpha\delta R(\overline{y}, \underline{e}|\pi, k) + (1 - \pi\alpha)\delta R(\underline{y}, \underline{e}|\pi, k), \qquad (4)$$

where $R(y, e|\pi, k)$ is his continuation payoff if he chooses e and produces y, which depends on π and k. Since the worker reveals himself to be of high ability if he produces high output, Lemma 1 implies that $R(\overline{y}, \underline{e}|\pi, k) = R(\overline{y}, \overline{e}|\pi, k)$. Then we can rewrite (4) as

$$\pi\eta\delta\underbrace{[R(\overline{y}, \overline{e}|\pi, k) - R(\underline{y}, \overline{e}|\pi, k)]}_{\Delta_1} + (1 - \pi\alpha)\delta\underbrace{[R(\underline{y}, \overline{e}|\pi, k) - R(\underline{y}, \underline{e}|\pi, k)]}_{\Delta_2} \geq c. \qquad (5)$$

As it turns out, Δ_1 is bounded above by $(1 - \delta)^{-1}(1 - \delta^{T-k})[v(\overline{w}) - v(\underline{w})]$. Indeed, since $R(\underline{y}, \overline{e}|\pi, k) \geq (1 - \delta)^{-1}(1 - \delta^{T-k})v(\underline{w})$, given that the worker can always collect his outside option, we immediately obtain the above bound on Δ_1 if the worker's wage after he reveals himself to be of high ability is always $\overline{w}$. Now observe that once the worker reveals himself to be of high ability, the firm pays him a wage w greater than $\overline{w}$ in some period t only if it committed to doing so before the worker produced high output. But if so, then w is also the wage the firm pays the worker in period t if he has not revealed himself to be of high ability by then. In the proof of Proposition 1, we show that this fact implies that the above upper bound on Δ_1 holds even if the firm pays the worker wages greater than $\overline{w}$ after he produces high output.

Next, notice that the worker's private belief about his ability after he produces low output is higher when he does not exert effort than when he does. Since an option for the worker in case he exerts no effort and produces low output is to behave from the next period on as if he had exerted effort (and produced low output), it follows that $R(\underline{y}, \overline{e}|\pi, k) \leq R(\underline{y}, \underline{e}|\pi, k)$. Intuitively, the higher the probability the worker is of high ability, the higher the probability that he produces high output, which leads to higher wages. Therefore, Δ_2 is at most zero. We can then conclude that a necessary condition for (5) is

$$\pi\eta(1 - \delta)^{-1}(1 - \delta^{T-k})[v(\overline{w}) - v(\underline{w})] \geq c,$$

which is not satisfied when (3) does not hold. Consequently, if an age 1 worker has no incentive to exert effort when he is dismissed after low output, then no other worker has an incentive to exert effort either. This step completes the proof of Proposition 1.

Proposition 1 admits the following corollary.

Corollary 1. *There exists no equilibrium in which age 1 workers are retained after low output if either* (3) *does not hold or* (3) *holds with equality.*

Indeed, we know by Proposition 1 that the above result is true if (3) is not satisfied. Suppose now that (3) holds with equality and let $\phi_1 = (1 - \alpha)\phi_0/(1 - \phi_0\alpha)$. Observe that ϕ_1 is the highest reputation possible for an age 2 worker who produced low output when of age 1. By assumption, we have that

$$\phi_1\eta\delta(1 - \delta^{T-2})[v(\overline{w}) - v(\underline{w})] < (1 - \delta)c. \tag{6}$$

But then, a straightforward modification of the argument preceding Corollary 1 shows that if (6) holds, then no worker of age $k \geq 2$ who has only produced low output has an incentive to exert effort. In this case, it is also never optimal for the firm to retain an age 1 worker after poor performance.

A natural conjecture is that under (3), all equilibria in the IID case are payoff-equivalent to the equilibrium σ^* of Proposition 1. As it turns out, this is not the case. In Appendix A, we show that when (3) holds, an equilibrium may exist in which age 1 workers do not exert effort and the firm retains them after poor performance. The payoff to the firm in this second equilibrium is smaller than its payoff under σ^*. A consequence of the multiplicity just described is that the firm would actually benefit from committing to dismiss rather than to retain age 1 workers after poor performance. Indeed, commitment to dismissal eliminates this multiplicity and, by Proposition 1, selects the most favorable outcome for the firm.

5 The Non-IID Case

We now study the case in which effort has an impact only on future output, so that an age 1 worker has no incentive to exert effort if dismissed after low output. We show that in this case, there exist

circumstances under which the firm offers probation to age 1 workers.

We divide the analysis of the non-IID case into three parts. In the first part, we introduce an assumption that is necessary for an age 1 worker to have an incentive to exert effort if retained after low output. Otherwise, commitment to employment does not improve incentives for effort and, thus, cannot be beneficial for the firm. We also derive sufficient conditions for an age 1 worker to exert effort if the firm offers him one period of probation, that is, if the firm makes him an offer of $(w^0, (q, \{w^s\}_{s=1}^q))$ with $q = 1$.

In the second part, we derive conditions under which the firm retains an age 1 worker who produces low output only if it commits to do so. In other words, we derive conditions under which it is ex ante optimal for the firm to retain an age 1 worker who produces low output, but it is not ex post optimal for the firm to do so. Hence, a time-inconsistency problem arises implying that commitment to employment is potentially beneficial to the firm. We also show that the conditions we derive imply that the firm does not benefit from either offering more than one period of probation to age 1 workers or offering commitment to employment to any other worker.

In the third part, we determine conditions under which the firm effectively benefits from offering commitment to employment. We also find conditions under which the gain from commitment is greater than just the increase in output, so as to make explicit the informational role of effort when commitment to employment is valuable. For simplicity, in what follows we let $\gamma = \alpha + \eta$. Also note that in what follows, assumptions (A2) and (A3) correspond, respectively, to assumptions (A3) and (A4) in Camargo and Pastorino (2013).

5.1 The Scope for Commitment

In what follows, we take the following condition as given:

(A2) $\phi_0(1 - \alpha)(\gamma - \alpha)(1 - \delta)^{-1}\delta^2(1 - \delta^{T-2})[v(\overline{w}) - v(\underline{w})] > c$.

The interpretation of (A3) is straightforward. When an age 1 worker exerts effort, $\phi_0(1-\alpha)(\gamma-\alpha)$ is the increase in the probability that he produces high output when of age 2 after producing low output when of age 1, whereas $\delta^2(1 - \delta)^{-1}(1 - \delta^{T-2})[v(\overline{w}) - v(\underline{w})]$ is the worker's lifetime payoff gain in this case. Thus, (A3) implies that an age 1 worker who is retained after low output, but is

dismissed if he produces low output one more time, has an incentive to exert effort. In Appendix A, we show that if (A2) does not hold, then, except for the knife-edged case in which the left side of (A2) equals its right side, no worker exerts effort in equilibrium, so that there is no scope for commitment to employment.

Now let $\phi_k = (1-\alpha)^k \phi_0 / [(1-\alpha)^k \phi_0 + 1 - \phi_0]$, with $k \geq 1$, be the highest reputation possible for a worker of age $k+1$ who has never produced high output. By construction, $\phi_k \gamma$ is an upper bound on the probability that such a worker produces high output. Notice that ϕ_1 is the reputation of an age 2 worker who produced low output when of age 1. In order to simplify the exposition, we also take the following condition as given:

(A3) $v(\underline{w}) + \phi_1 \alpha \delta (1-\delta)^{-1} (1-\delta^{T-2})[v(\underline{w}) - v(\overline{w})] \leq v(0)$.

We analyze the case in which (A3) does not hold in Appendix B. Assumption (A3) implies that the option value of employment to an age 2 worker who failed to reveal himself to be of high ability is large enough that he always accepts employment at the firm. Assumption (A3) also implies that an age 1 worker accepts any offer by the firm. Since (A3) implies that an age 1 worker accepts any offer of the form $(w^0, (1, w^1))$, the following result holds. See Appendix A for omitted details and a proof of Lemma 4.

Lemma 4. *An age 1 worker who is retained after low output but is dismissed when of age 3 if he has not produced high output by then, accepts employment and exerts effort when the firm offers him one period of probation.*

Lemma 4 is important to show that the firm cannot gain from either offering more than one period of probation to age 1 workers or offering commitment to employment to any other worker.

5.2 Commitment Is Necessary for Incentives

We know that an age 1 worker exerts effort only if he is retained after poor performance. We now derive conditions under which the firm faces a time-inconsistency problem in that it can induce an age 1 worker to exert effort only if it commits to employing him after low output. In order to do so, let $\Delta_y = \overline{y} - \underline{y}$, $\Delta = [y(1, \alpha) - \overline{w}] - y(\phi_0, \alpha)$ be the firm's payoff gain from employing a

high-ability worker instead of an age 1 worker, and $\Upsilon = \Delta/\Delta_y$ be the same payoff gain in units of output.

Proposition 2. *Suppose that $\phi_2\gamma \leq \phi_0\alpha$ and that*

$$\phi_1\gamma < \phi_0\alpha \left[1 + \min\left\{\frac{\eta + \Upsilon + \phi_0\alpha\eta(T-2)}{1 + \phi_0\alpha + [\Upsilon + \phi_0\alpha(1-\eta)](T-2)}, \frac{\alpha(1-\alpha)(1-\phi_0)}{1 - \phi_0\alpha}\right\}\right]. \quad (7)$$

There exists $\delta^* \in (0,1)$ such that if $\delta > \delta^*$, then in equilibrium the firm retains a worker of age $k \geq 2$ who has never produced high output only if it is committed to doing so.

The restriction that $\phi_2\gamma \leq \phi_0\alpha$ limits the gain to the firm from employing a worker of age 3 or more who has not revealed himself to be of the high type, so that the firm employs such a worker only if committed to doing so. The restriction that $\phi_1\gamma$ cannot be much greater than $\phi_0\alpha$ is also intuitive. Indeed, if $\phi_1\gamma - \phi_0\alpha$ is sufficiently large, then despite his lower reputation and the fact that he lives for one less period, an age 2 worker who exerted effort when of age 1 and produced low output is more attractive to the firm than an age 1 worker regardless of the discount factor. In this case, the firm retains the age 2 worker even if not committed to doing so, and no issue of time-inconsistency emerges. In other words, commitment to employment is not necessary to induce an age 1 worker to exert effort when $\phi_1\gamma - \phi_0\alpha$ is large enough.

It follows from Proposition 2 that $\phi_1\gamma \leq \phi_0\alpha$ is a simpler sufficient condition for the firm to have no incentive to retain an age 2 worker who produced low output when of age 1. However, as we show below, $\phi_1\gamma > \phi_0\alpha$ is necessary for commitment to employment to be beneficial from an informational point of view. Now observe that if $\phi_1\gamma > \phi_0\alpha$, then the firm's flow payoff when it employs an age 2 worker who exerted effort and produced low output when of age 1 is greater than the firm's flow payoff when it employs an age 1 worker. Hence, when $\phi_1\gamma - \phi_0\alpha$ is positive, but not large enough to eliminate the time-inconsistency problem altogether, commitment is necessary for the firm to retain a worker of age 2 who produced low output when of age 1 only when the firm is sufficiently patient. This explains the discount factor restriction in Proposition 2.

Finally, note that since $\Delta \geq \phi_0\eta\Delta_y$ by (A1), the right side of (7) is strictly decreasing in T. This result is intuitive. When $\phi_1\gamma > \phi_0\alpha$, an increase in T makes an age 2 worker who exerted effort when of age 1 and produced low output more attractive to the firm compared with an age

1 worker. Thus, an increase in T mitigates the firm's time-inconsistency problem, in which case commitment to employment is less likely to be necessary to provide incentives.

An immediate consequence of Proposition 2 is that the firm never offers commitment to employment to a worker of age $k \geq 2$. A further implication is that it is not optimal for the firm to offer more than one period of probation to age 1 workers; see Appendix A for details.[18]

Corollary 2. *In equilibrium, the firm offers at most one period of probation to age 1 workers.*

5.3 Commitment Is Beneficial

Proposition 2 provides sufficient conditions for the firm to face a time-inconsistency problem, so that it can potentially benefit from commitment to employment. Here, we first derive a condition under which the firm actually benefits from offering probation to age 1 workers. We then study when offering probation to age 1 workers is beneficial just for the information about ability that the firm can acquire during probation. We take the conditions of Proposition 2 as given in what follows.

In order to determine when the firm benefits from offering probation to age 1 workers, we begin by deriving the firm's equilibrium payoff V_{nc} when it cannot offer commitment to employment. In this case, Proposition 2 implies that the firm always dismisses an age 1 worker who produces low output, so that age 1 workers do not exert effort. Since (A3) implies that an age 1 worker accepts a one-period wage of zero, V_{nc} then satisfies

$$V_{nc} = (1 - \delta)y(\phi_0, \alpha) + \phi_0\alpha \left\{\delta(1 - \delta^{T-1})[y(1, \alpha) - \overline{w}] + \delta^T V_{nc}\right\} + (1 - \phi_0\alpha)\delta V_{nc}.$$

To see why, note that in equilibrium the firm retains an age 1 worker who produces high output, which happens with probability $\phi_0\alpha$, and replaces him with a new age 1 worker otherwise.

Now observe from Corollary 2 that if the firm is not constrained in the offers it can make, then it offers at most one period of probation to age 1 workers. Moreover, by Lemma 4 and Proposition

---

[18]The result that there is no scope for more than one period of probation depends in part on the assumption that there is a one-period delay in the impact of effort on output. This result also depends on the condition that $\phi_2\gamma \leq \phi_0\alpha$, which limits the gain to the firm from employing a worker of age 3 or more. Our analysis can allow for more than one period of probation as an equilibrium outcome if, for instance, it takes more than one period for effort to affect output.

2, an age 1 worker accepts any offer with one period of probation and exerts effort. Hence, the highest payoff V_c the firm obtains when it offers probation to age 1 workers is when it always offers $(w^0, (1, w^1))$ with $w^0 = w^1 = 0$ to such workers. It is easy to see that V_c satisfies

$$V_c = (1 - \delta)y(\phi_0, \alpha) + \phi_0\alpha \left\{ \delta(1 - \delta)[y(1, \gamma) - \overline{w}] + \delta^2(1 - \delta^{T-2})[y(1, \alpha) - \overline{w}] + \delta^T V_c \right\}$$
$$+ (1 - \phi_0\alpha) \left\{ \delta(1 - \delta)y(\phi_1, \gamma) + \phi_1\gamma \left\{ \delta^2(1 - \delta^{T-2})[y(1, \alpha) - \overline{w}] + \delta^T V_c \right\} + (1 - \phi_1\gamma)\delta^2 V_c \right\}.$$

To interpret V_c, note that now the firm retains an age 1 worker after low output, which happens with probability $1 - \phi_0\alpha$. In this event, the firm retains the worker only if he produces high output when of age 2, which happens with probability $\phi_1\gamma$. Observe that if the worker is of high ability, then he produces high output with probability γ when of age 2 if he exerted effort when of age 1.

Therefore, the firm benefits from commitment to employment if, and only if, $V_c > V_{nc}$. The cost of offering probation to an age 1 worker is that it prevents the firm from replacing the worker if he performs poorly. The gain from probation is that when an age 1 worker exerts effort, his output increases and his performance becomes more informative about his ability.

In order to determine when commitment to employment is beneficial, and also to understand the distinct roles of effort in our environment, let ρ_O be the lifetime value to the firm of the extra output it obtains when age 1 workers exert effort. Notice that ρ_O satisfies

$$\rho_O = \phi_0\alpha \left\{ \delta(1 - \delta)[y(1, \gamma) - y(1, \alpha)] + \delta^T \rho_O \right\}$$
$$+ (1 - \phi_0\alpha) \left\{ \delta(1 - \delta)[y(\phi_1, \gamma) - y(\phi_1, \alpha)] + \phi_1\gamma\delta^T \rho_O + (1 - \phi_1\gamma)\delta^2 \rho_O \right\}.$$

To see why, note that when an age 1 worker exerts effort, the (expected) increase in output in the following period is either $y(1, \gamma) - y(1, \alpha)$ or $y(\phi_1, \gamma) - y(\phi_1, \alpha)$, depending on whether the worker produces high or low output when of age 1. Moreover, if an age 1 worker produces high output, then the firm retains the worker (and so only hires a new age 1 worker T periods later). On the other hand, if an age 1 worker produces low output, the firm retains the worker if he produces high output when of age 2 and replaces him otherwise. Solving the above recursions for V_c, V_{nc}, and ρ_O, it follows that

$$V_c - V_{nc} = \rho_O + \rho_I,$$

where

$$\rho_I = \left\{ \frac{1-\delta}{1-\delta+\phi_0\alpha\delta(1-\delta^{T-1})} - \frac{(1-\delta)[1+\delta(1-\phi_0\alpha)(1+\psi)]}{1-\delta^2+\phi_0[\alpha+\gamma(1-\alpha)]\delta^2(1-\delta^{T-2})} \right\} \Delta$$

and $\psi = \phi_0(1-\phi_0)\alpha^2/(1-\phi_0\alpha)\Upsilon$. So, the gain from commitment can be decomposed into two parts: the output gain, ρ_O, and the informational gain net of the commitment cost, ρ_I.

Given that the focus of our analysis is on how commitment to employment can allow the firm to better sort workers, the case of interest is the one in which ρ_I is greater than zero, so that the benefit of probation is greater than the extra output it generates. In particular, when ρ_I is greater than zero, even if the gain in output due to probation were lost, commitment to employment would still be valuable to the firm. We then say that commitment to employment has *informational value* when $\rho_I > 0$. Straightforward algebra shows that $\rho_I > 0$ is equivalent to

$$\phi_1\gamma > \phi_0\alpha(1+\psi)\frac{1-\delta^{T-1}}{1-\delta^{T-2}} + \psi\frac{(1-\delta)}{\delta(1-\delta^{T-2})}. \tag{8}$$

A necessary condition for (8) is that $\phi_1\gamma > \phi_0\alpha$. This restriction is intuitive: when $\phi_1\gamma \leq \phi_0\alpha$, the performance of an age 2 worker who failed to reveal himself to be of high ability is less informative than the performance of an age 1 worker. In this case, the firm can benefit from commitment to employment only because of the additional output it obtains from effort. Straightforward algebra shows that the right side of (8) is strictly decreasing in T. This is also intuitive. The longer a worker's lifetime is, the more the firm benefits from identifying high-ability workers. Thus, the informational value of probation increases with T, making it easier for (8) to be satisfied. Notice that when δ is high, which is the case of interest, (8) becomes

$$\phi_1\gamma > \phi_0\alpha(1+\psi)\frac{T-1}{T-2}.$$

We can now state the main result of this section, which follows from the discussion so far.

Proposition 3. *Suppose that $\phi_2\gamma \leq \phi_0\alpha$ and that*

$$\phi_0\alpha(1+\psi)\frac{T-1}{T-2} < \phi_1\gamma$$
$$< \phi_0\alpha\left[1+\min\left\{\frac{\eta+\Upsilon+\phi_0\alpha\eta(T-2)}{1+\phi_0\alpha+[\Upsilon+\phi_0\alpha(1-\eta)](T-2)}, \frac{\alpha(1-\alpha)(1-\phi_0)}{1-\phi_0\alpha}\right\}\right]. \tag{9}$$

There exists $\delta^{**} \in (0,1)$ such that if $\delta > \delta^{**}$, then: (i) the firm always offers probation to age 1 workers in equilibrium; (ii) commitment to employment has informational value.

Recall that $\phi_2 \gamma \leq \phi_0 \alpha$ implies that the firm prefers an age 1 worker to an age $k \geq 3$ worker who has only produced low output. Condition (9) has also a clear interpretation. On the one hand, as already discussed, $\phi_1 \gamma$ needs to be high enough for the informational gain from commitment to employment to be positive. On the other hand, we know from Proposition 2 that $\phi_1 \gamma$ cannot be too high, otherwise there is no time-inconsistency problem and probation is not necessary to induce age 1 workers to exert effort. Note from (9) that the required bounds on $\phi_1 \gamma$ depend on a worker's lifetime. Specifically, by the discussions following Proposition 2 and condition (8), both the left and right sides of (9) are decreasing in T. Thus, one expects that (9) holds for intermediate values of T. Intuitively, T cannot be so low that commitment to employment has no informational value and T cannot be so large that the firm does not face a time-inconsistency problem. Indeed, the conditions of Proposition 3 are satisfied for intermediate values of T. For instance, for T between 10 and 30, all the model's restrictions are easily satisfied for values of ϕ_0 and α smaller than 1/2.

To summarize our results in this section, we have established that when effort only affects future output, an age 1 worker can benefit from exerting effort only if he is retained after low output. Thus, when the firm cannot credibly promise to retain an age 1 worker after low output, the use of probationary appointments can provide such workers with the appropriate incentives to exert effort. Crucially, since effort makes performance more informative about ability, the firm can benefit from offering commitment to employment purely from an informational point of view.

6 Discussion

In this section we first provide evidence on the use of probationary appointments and then discuss some of our modeling choices.

6.1 Evidence on Probationary Appointments

A number of papers in the literature on internal labor markets provide evidence on the use of probationary appointments in professional service industries. For instance, in discussing employment practices in academia, Carmichael (1988) states that "incumbents are reviewed after a specified period, and those who do not meet with the approval of their senior colleagues are involuntarily

released" (page 468). Similarly, in the case of law firms, Wilkins and Gulati (1998) state: "Law firms traditionally have only two categories of workers: partners and associates. Associates are hired on the express understanding that at the end of a fixed probationary period, some of them will be promoted and the rest will be asked to leave" (page 1591). See Gibbons (1998), Rebitzer and Taylor (2007), and Ghosh and Waldman (2010) for similar evidence.

A natural question is whether the commitment not to dismiss a worker before the end of a probationary period, the commitment we focus on, is a common feature of probationary contracts used in practice. To address this question, we surveyed a private and a public research university in the United States about their employment practices for tenure-track hires at the faculty level. The questionnaire we submitted is in Appendix C. In both universities, the employment contract offered to tenure-track assistant professors consists of a probationary period of predetermined length and specifies an intermediate and a final review of an assistant professor's record. The first review is to decide on the renewal of the probationary appointment, while the second review is to decide on the award of promotion to associate professor and permanent tenure. An explicitly stated concern in both universities is that the probationary period should be sufficiently long, since it is believed that a candidate's later research output is more informative. The circumstances under which an assistant professor on tenure track can be terminated before either review period are very rare and similar to the circumstances under which a tenured professor can be dismissed.[19]

Overall, the evidence just discussed suggests that the type of commitment we analyze is common in professional service industries. This evidence also supports our assumption that effort has a delayed impact on output, at least in academia.

6.2 Modeling Choices

Labor Market. As in a standard career concerns model, we focus on a labor market in which ability is valuable but scarce and is revealed over time through performance. In our setting, a worker obtains a higher wage once he is revealed to be of high ability. However, unlike standard models of career concerns, we consider a market in which either: (*i*) firms possess enough monop-

---

[19]Some of these instances correspond to termination by just cause, or situations in which a professor falls short of satisfying his job requirements, or situations in which the professor otherwise compromises students' education.

sony power to be able to extract more surplus from a match with a high-ability worker than from a match with a worker of unknown ability; or (*ii*) high-ability workers accumulate firm-specific human capital. Either way, a positive rent accrues to the firm when an employed worker is revealed to be of high ability. With perfect competition and no accumulation of firm-specific human capital for high-ability workers, a firm cannot strictly benefit from identifying workers of high ability. In this case, commitment to employment is not valuable.

Nondecreasing Outside Option. As discussed in Section 2, the assumption that a worker's outside option cannot decrease below $\underline{w}$ captures situations in which a worker can seek employment in an alternative labor market where his ability (in the market we consider) has no value. The model can accommodate the case in which $w_R(\phi)$, the workers' outside option, falls below $\underline{w}$ when $\phi < \phi_0$ as long as this decrease is not too large. Indeed, when $w_R(0)$ is sufficiently close to $w_R(\phi_0)$, the worker who is willing to accept the smallest wage is still the age 1 worker, and this is sufficient for the analysis to proceed as above. Allowing for a decreasing outside option would complicate the analysis without bringing any additional insight.

No Recall. In principle, a worker who is still in the market after being dismissed by the firm could prove himself to be of high ability, in which case the firm could consider re-hiring him. Our assumption of no recall rules out this possibility. This assumption is consistent with the possibility that the firm may not be willing or able to rehire a worker. For instance, this is the case if the firm employs a younger worker known to be of high ability. The assumption of no recall is only important in the non-IID case because it implies that an age 1 worker has no incentive to exert effort unless he is retained after low output. When recall is possible, the results in the non-IID case continue to hold as long as being dismissed after poor performance reduces a worker's return from exerting effort to the point of eliminating the incentive for effort altogether. Considering a more general environment with recall would provide no additional insights.

Good Performance Reveals High Ability. We assume that only high type workers can produce high output. This assumption is consistent with common features of the environments we consider,

namely, that workers can repeatedly fail regardless of their ability, especially early in their careers, and that success is highly rewarded. Hence, we think the assumption that success reveals ability is reasonable.

Here, we also argue that our results still hold if the probability that a low-ability worker produces high output is positive but small.[20] For simplicity, assume that this probability, which we denote by α_L, does not depend on a low type worker's current (in the IID case) or previous (in the non-IID case) choice of effort. We can adapt our argument to the case in which a low type worker's effort has an impact on the probability that he produces high output.

To see that our results still hold when a low-ability worker produces high output, observe first that there exists $\phi^* \in (0, 1)$, which depends on T, such that Lemma 1 applies to any worker with reputation in the interval $[\phi^*, 1]$. Moreover, by increasing ϕ^* if necessary, we can take ϕ^* to be such that $y(\phi^*, \alpha) - \overline{w} > y(\phi_0, \alpha + \eta)$, in which case a straightforward modification of the proof of Lemma 3 shows that it is optimal for the firm to retain an incumbent with reputation in the interval $[\phi^*, 1]$. Now observe that if α_L is sufficiently small, then high output still provides a strong signal that a worker is of high ability. More precisely, there exists $\alpha'_L \in (0, 1)$, which also depends on T, such that if $\alpha_L \leq \alpha'_L$, then once a worker produces high output, his reputation stays in the interval $[\phi^*, 1]$ until he exits the market. Thus, as long as $\alpha_L \in (0, \alpha'_L)$, producing high output plays the same role as revealing oneself to be of the high type. Moreover, by reducing α'_L if necessary, we can ensure that in the non-IID case, an age 1 worker can benefit from effort only if he is retained after low output. Proceeding as in Sections 4 and 5, we obtain the same results.

Incentive Pay.   We assume that wages in a period cannot be made contingent on the period's output. Therefore, commitment to employment is the only instrument the firm can use to strengthen the incentives for effort provided by career concerns. This allows us to analyze the trade-offs involved in the use of probation in a transparent way. Moreover, as pointed out in the introduction,

---

[20]It may be possible that even in the IID case, the firm can benefit from commitment to employment if high output is not sufficiently informative about ability. However, if effort matters little for output, then commitment to employment has no value regardless of the probability that a low-ability worker produces high output. Indeed, the limiting case in which effort has no impact on output corresponds to a standard multi-armed bandit problem. As we discussed in the related literature, in this case there is no scope for commitment to employment. Hence, by continuity, there is no scope for commitment either when the impact of effort on output is positive but small.

probationary appointments are common in academia and bureaucracies, where the use of incentive pay is limited. Nevertheless, probationary appointments are also used in environments in which performance pay is observed, such as consulting firms. Then, a natural question is whether the main insights of the paper would change in the presence of performance pay. We argue that this is not the case.

Consider first the IID case. Suppose that (3) holds, and so career concerns are strong enough to induce an age 1 worker who is dismissed after low output to exert effort. In this case, the best that the firm can do is to dismiss age 1 workers after poor performance and use output-contingent pay to induce a worker who reveals himself to be of the high type to exert effort. Thus, the use of probation is still not beneficial to the firm. Suppose now that (3) does not hold. We know from the proof of Proposition 1 that the incentive to exert effort for a worker who has only produced low output is maximized when he is dismissed after poor performance. So, it is still optimal for the firm to dismiss age 1 workers after low output (and use output-contingent contracts to induce high ability workers to exert effort). The only difference from the case in which (3) holds is that now the firm can reward high output from age 1 workers in order to induce them to exert effort. Thus, the use of probation is still not beneficial for the firm in the IID case.

Consider now the non-IID case. We know that when performance pay is not possible, the firm can induce an age 1 worker to exert effort only by retaining him after poor performance. With performance pay, the firm can induce age 1 workers to exert effort even if they are dismissed after low output. It can do so by retaining an age 1 worker after good performance and rewarding him for high output when he is of age 2. Nevertheless, the firm can still use probationary appointments as an incentive device. Whether the firm employs them or not depends on how restricted it is in the use of performance pay. In particular, in situations in which an age 1 worker does not have a strong career concerns motive to exert effort and the use of performance pay is limited, the firm may benefit by combining performance pay with commitment to retention.

7 Conclusion

n this paper we provide a rationale for the use of short-term commitment to employment in labor markets in which a worker's ability is uncertain. We prove that a firm can benefit from committing to employing workers of unknown ability if this commitment encourages workers to exert effort, thus making their performance more informative about their ability. Specifically, we show that firms do not gain from commitment to employment in the standard case in which the impact of effort on output is independent and identical over time. However, probation can be valuable when the impact of effort on output is delayed. In this case, commitment to employment solves a time-inconsistency problem. Even though it is ex ante optimal for a firm to retain a new hire after bad performance, without commitment to employment a firm cannot credibly promise to retain a worker whose initial performance is poor. In turn, this inability to credibly promise retention undermines a worker's incentives for effort. Hence, when the impact of effort on output is delayed, it can be ex ante optimal for a firm to commit to ex post inefficient outcomes. Importantly, since effort increases the informativeness of performance, we show that the use of probation can be justified solely on the basis of the informational role of effort.

In practice, the type of commitment we focus on and up-or-out contracts are frequently linked. In our environment, the fact that high output reveals high ability implies that the firm's retention decision follows an up-or-out rule. More generally, as discussed in the previous section, the firm's retention decision has an up-or-out form as long as high output is a strong enough signal of high ability. Thus, our analysis provides a simple theory in which probationary appointments are coupled with up-or-out rules. The literature on internal labor market practices has suggested other reasons for why probationary appointments combined with up-or-out contracts solve a time-inconsistency problem: for instance, the need to provide incentives for human capital acquisition (Kahn and Huberman (1988), Waldman (1990), and Ghosh and Waldman (2010)) or incentives for the selection of new workers of high ability by currently employed workers (Carmichael (1988)). A more general theory of the joint use of probationary appointments and up-or-out rules constitutes an interesting topic for future research.

References

[1] Andersson, F. (2002): "Career Concerns, Contracts, and Effort Distortions," *Journal of Labor Economics*, 20(1), 42-58.

[2] Banks, J. S., and R. K. Sundaram (1998): "Optimal Retention in Agency Problems," *Journal of Economic Theory*, 82(2), 293-323.

[3] Berry, D. A., and B. Fristedt (1985): *Bandit Problems: Sequential Allocation of Experiments*, London: Chapman and Hall.

[4] Bull, C., and P. Tedeschi (1989): "Optimal Probation for New Hires," *Journal of Institutional and Theoretical Economics*, 145(4), 627-642.

[5] Camargo, B., and E. Pastorino (2013): "Learning-by-Employing: The Value of Commitment under Uncertainty," Research Department Staff Report 475, Federal Reserve Bank of Minneapolis.

[6] Carmichael, H. L. (1988): "Incentives in Academics: Why Is There Tenure?," *Journal of Political Economy*, 96(3), 453-472.

[7] Datta, M., L. J. Mirman, and E. E. Schlee (2002): "Optimal Experimentation in Signal-Dependent Decision Problems," *International Economic Review*, 43(2), 577-607.

[8] Eeckhout, J., and X. Weng (2010): "Assortative Learning," mimeo.

[9] Felli, L., and C. Harris (1996): "Learning, Wage Dynamics, and Firm-Specific Human Capital," *Journal of Political Economy*, 104(4), 838-868.

[10] Ghosh, S., and M. Waldman (2010): "Standard Promotion Practices versus Up-or-Out Contracts," *RAND Journal of Economics*, 41(2), 301-325.

[11] Gibbons, R. (1998): "Incentives in Organizations," *Journal of Economic Perspectives*, 12(4), 115-132.

[12] Gibbons, R., and K. J. Murphy (1992): "Optimal Incentive Contracts in the Presence of Career Concerns: Theory and Evidence," *Journal of Political Economy*, 100(3), 468-505.

[13] Guasch, J. L., and A. Weiss (1981): "Self-Selection in the Labor Market," *American Economic Review*, 71(3), 275-284.

[14] Guasch, J. L., and A. Weiss (1982): "An Equilibrium Analysis of Wage-Productivity Gaps," *Review of Economic Studies*, 49(4), 485-497.

[15] Harris, M., and Y. Weiss (1984): "Job Matching with Finite Horizon and Risk Aversion," *Journal of Political Economy*, 92(4), 758-779.

[16] Holmström, B. (1999): "Managerial Incentives: A Dynamic Perspective," *Review of Economic Studies*, 66(1), 169-182.

[17] Jarque, A. (2010): "Repeated Moral Hazard with Effort Persistence," *Journal of Economic Theory*, 145(6), 2412-2423.

[18] Jovanovic, B. (1979): "Job Matching and the Theory of Turnover," *Journal of Political Economy*, 87(5), 972-990.

[19] Jovanovic, B. (1984): "Matching, Turnover, and Unemployment," *Journal of Political Economy*, 92(1), 108-122.

[20] Kahn, C., and G. Huberman (1988): "Two-Sided Uncertainty and 'Up-or-Out' Contracts," *Journal of Labor Economics*, 6(4), 423-444.

[21] Loh, E. S. (1994): "The Determinants of Employment Probation Lengths," *Industrial Relations*, 33(3), 386-406.

[22] Manso, G. (2011): "Motivating Innovation," *Journal of Finance*, 66(5), 1823-1860.

[23] Milgrom, P., and J. Roberts (1988): "An Economic Approach to Influence Activities in Organizations," *American Journal of Sociology*, 94(Supplement), S154-S179.

[24] Moscarini, G. (2005): "Job Matching and the Wage Distribution," *Econometrica*, 73(2), 481-516.

[25] Mukherjee, A. (2008): "Sustaining Implicit Contracts When Agents Have Career Concerns: The Role of Information Disclosure," *RAND Journal of Economics*, 39(2), 469-490.

[26] Mukoyama, T., and A. Şahin (2005): "Repeated Moral Hazard with Persistence," *Economic Theory*, 25(4), 831-854.

[27] O'Flaherty, B., and A. Siow (1992): "On the Job Screening, Up or Out Rules, and Firm Growth," *Canadian Journal of Economics*, 25(2), 346-368.

[28] Prendergast, C. (1993): "The Role of Promotion in Inducing Specific Human Capital Acquisition," *Quarterly Journal of Economics*, 108(2), 523-534.

[29] Rebitzer, J. B., and L. J. Taylor (2007): "When Knowledge Is an Asset: Explaining the Organizational Structure of Large Law Firms," *Journal of Labor Economics*, 25(2), 201-229.

[30] Scharfstein, D. S., and J. C. Stein (1990): "Herd Behavior and Investment," *American Economic Review*, 80(3), 465-479.

[31] Waldman, M. (1990): "Up-or-Out Contracts: A Signaling Perspective," *Journal of Labor Economics*, 8(2), 230-250.

[32] Waldman, M. (2003): "Ex Ante versus Ex Post Optimal Promotion Rules: The Case of Internal Promotion," *Economic Inquiry*, 41(1), 27-41.

[33] Wang, R., and A. Weiss (1998): "Probation, Layoffs, and Wage-Tenure Profiles: A Sorting Explanation," *Labour Economics*, 5(3), 359-383.

[34] Wilkins, D. B., and G. M. Gulati (1998): "Reconceiving the Tournament of Lawyers: Tracking, Seeding, and Information Control in the Internal Labor Markets of Elite Law Firms," *Virginia Law Review*, 84(8), 1581-1681.

Appendix A: Omitted Proofs

Claim 1. *In equilibrium, the firm makes an offer in every equilibrium and never makes an offer that it anticipates will be rejected with positive probability.*

Proof. Suppose there is a period t in which the firm either does not make an offer or makes an offer that is rejected. The firm's lifetime payoff in period t is then $(1-\delta)\Pi + \delta V$, where V is the firm's expected continuation payoff. There are two possibilities in period t: (*i*) the firm is not committed to making an offer to its incumbent or it has no incumbent; (*ii*) the firm is committed to making an offer to its incumbent. Suppose (*i*) holds and consider the following deviation for the firm: offer $(\underline{w}, 0)$ to the age 1 worker in period t and behave from period $t+1$ on as if it did not deviate in t. Since an age 1 worker always accepts a one-period wage of $\underline{w}$, the firm's payoff from this deviation is at least $(1-\delta)[\underline{y} - \underline{w}] + \delta V$, which is greater than $(1-\delta)\Pi + \delta V$ by assumption. Thus, the firm has a profitable deviation. Consider now case (*ii*). By (*i*), we only need to consider the case in which the incumbent is known to be of the high type. Since $y(1, \alpha) - \overline{w} > \underline{y} - \underline{w}$, and a worker known to be of high ability accepts any wage greater than $\overline{w}$, an argument similar to the one used in case (*i*) shows that the firm has a profitable deviation. $\square$

Proof of Lemma 1: Let $k \in \{2, \ldots, T\}$ be the incumbent's age and $\ell \in \{0, \ldots, T-k\}$ be the maximum number of future periods the firm employs the incumbent if it makes him an offer that he accepts. The proof is by induction in ℓ. Note that if $\ell = 0$, then: (*i*) the firm never offers the incumbent a one-period wage greater than $\max\{\overline{w}, w'\}$; (*ii*) if the firm makes the incumbent an offer, then it never commits to future minimum one-period wage offers greater than $\overline{w}$ (trivially satisfied); (*iii*) the incumbent does not exert effort if employed. Suppose then, by induction, that there exists $\ell' \in \{0, \ldots, T-k\}$ such that (*i*) to (*iii*) hold if $\ell \leq \ell'$ and let $\ell = \ell' + 1$. We claim that (*iii*) is true. Indeed, the induction hypothesis implies that the incumbent's continuation payoff does not depend on his output. Note that when the firm is committed to making a minimum one-period wage offer $w' > \overline{w}$ to the incumbent in the future, the value of w' does not depend on the incumbent's current period output. It is now easy to see that (*ii*) must also hold, for otherwise the firm can profitably deviate by lowering the present discounted value of the future wages it offers to

the incumbent while still satisfying the incumbent's participation constraints. A similar argument shows that the firm has a profitable deviation if (*i*) does not hold. $\square$

Proof of Lemma 3: Consider first the IID case. Suppose σ is an equilibrium. By Lemma 1, an incumbent known to be of the high type never exerts effort. Moreover, such a worker rejects any offer with a one-period wage smaller than $\overline{w}$. Hence, by (A1), $V(h|\sigma) < y(1, \alpha) - \overline{w}$ if h is the initial history of the game. Lemma 2 then implies that $V(h'|\sigma) < y(1, \alpha) - \overline{w}$ for every history h' for the firm after which it hires the available age 1 worker. Since a worker known to be of the high type always accepts an offer greater than $\overline{w}$, it must be that the firm offers $\max\{\overline{w}, w'\}$ to such a worker. To finish, note that on the equilibrium path, a worker known to be of the high type always accepts an offer of $\overline{w}$. Otherwise, the firm has a profitable deviation, since $\Pi < \underline{y} - \underline{w} < y(1, \alpha) - \overline{w}$. A straightforward induction argument then implies that on the equilibrium path, a worker known to be of the high type always accepts an offer of $\overline{w}$.

Consider now the non-IID case. The present value of the output generated by a worker who has not revealed himself to be of the high type is at most $y(\phi_0, \alpha) + \delta\phi_0[y(1, \gamma) - y(1, \alpha)]$. Indeed, when a worker is of high ability, he creates an extra output of $y(1, \gamma) - y(1, \alpha)$ in the following period if he exerts effort. Now observe that

$$y(\phi_0, \alpha) + \delta\phi_0[y(1, \gamma) - y(1, \alpha)] = y(\phi_0, \alpha + \delta(\gamma - \alpha)) < y(1, \alpha) - \overline{w},$$

where the last inequality follows from (A1). The desired result now follows from the same argument as in the IID case. $\square$

Proof of Proposition 1: We divide the proof of Proposition 1 into two parts. First, we consider the case in which (3) holds. Then, we consider the case in which (3) is violated.

Case 1. We divide Case 1 into three parts. First, we establish an auxiliary result that plays an important role in the argument that follows. Then, we show that there exists an equilibrium σ^* in which the firm dismisses age 1 workers after poor performance and these workers exert effort. It is clear that the firm's payoff in any equilibrium in which it dismisses age 1 workers after low output is the same as the firm's payoff under σ^*. Finally, we show that the firm's payoff in any equilibrium in which it retains an age 1 worker after low output is smaller than its payoff in σ^*.

Step 1. Let $\widetilde{w}_1$ be such that

$$v(\widetilde{w}_1) + \phi_0(\alpha + \eta)\delta(1 - \delta)^{-1}(1 - \delta^T)[v(\overline{w}) - v(\underline{w})] - c = v(\underline{w}).$$

By construction, $\widetilde{w}_1$ is the smallest one-period wage an age 1 worker accepts if he exerts effort and is not retained after poor performance. Recall that Lemma 1 implies that if the firm does not offer probation to an age 1 worker and he produces high output, then the worker's wage in every subsequent period is $\overline{w}$. We have the following result.

Claim 2. *Suppose the firm employs an age 1 worker until he is of age $q \leq T$ and then retains him only if he has revealed himself to be of the high type. A lower bound on the firm's present discounted wage bill from employing the worker is achieved when it pays him a wage of $\widetilde{w}_1$ as long as he does not produce high output and a wage of $\overline{w}$ once he reveals himself to be of high ability.*

Proof. Let w_k, e_k, and ϕ_k be, respectively, the worker's wage, choice of effort, and reputation when of age $k \in \{1, \ldots, q\}$ if he has not revealed himself to be of high ability by then. Note that $\phi_1 = \phi_0$ and that ϕ_k is strictly decreasing in k. Now define $\widehat{\xi}(e_k)$ and $\widehat{c}(e_k)$ to be such that: *(i)* $\widehat{\xi}(e_k) = \alpha$ and $\widehat{c}(e_k) = 0$ if $e_k = \underline{e}$; *(ii)* $\widehat{\xi}(e_k) = \alpha + \eta$ and $\widehat{c}(e_k) = c$ if $e_k = \overline{e}$.

As the first step, we show that $w_k \leq \underline{w}$ for all $k \in \{1, \ldots, q\}$ is necessary for the firm to minimize the present discounted wage bill from employing the worker. Clearly, $w_1 \leq \underline{w}$, because an age 1 worker always accepts an offer with a one-period wage of $\underline{w}$. Suppose now that there exists $k \in \{1, \ldots, q-1\}$ such that $w_k \leq \underline{w}$ but $w_{k+1} > \underline{w}$, and let κ be given by

$$v(w_k + \kappa) - v(w_k) = [1 - \phi_k\widehat{\xi}(e_k)]\delta\{v(\underline{w} + \varepsilon) - v(\underline{w})\},$$

where $\varepsilon = w_{k+1} - \underline{w}$. Since v is (weakly) concave and $w_k \leq \underline{w}$, we then have that

$$v(\underline{w} + \kappa) \leq v(\underline{w}) + v(w_k + \kappa) - v(w_k)$$
$$\leq [1 - \delta + \delta\phi_k\widehat{\xi}(e_k)]v(\underline{w}) + [1 - \phi_k\widehat{\xi}(e_k)]\delta v(\underline{w} + \varepsilon) \leq v(\underline{w} + (1 - \phi_k\widehat{\xi}(e_k))\delta\varepsilon)),$$

where the last inequality also follows from the concavity of v. Hence, $\kappa \leq [1 - \phi_k\widehat{\xi}(e_k)]\delta\varepsilon$, because v is strictly increasing. Therefore, if the firm decreases w_{k+1} to $\underline{w}$ and increases w_k to $w_k + \kappa$, it reduces the present discounted wage bill from employing the worker by at least

$\delta^k \varepsilon - [1 - \phi_k \widehat{\xi}(e_k)]\delta^k \varepsilon$ while still satisfying his participation constraints (the worker has the option of not changing his behavior after the change in wages). The desired result follows by an induction argument.[21]

Suppose then that $w_k \leq \underline{w}$ for all $k \in \{1, \ldots, q\}$. We now show that a necessary condition for the firm to minimize the present discounted wage bill from employing the worker is that it pays him a wage of $\overline{w}$ once he produces high output. Indeed, let $w_{k,k+s}$ be the wage the firm pays the worker when he is of age $k + s$ if he first produces high output when of age k, and suppose that $w_{k,k+s} > \overline{w}$ for some $k \in \{1, \ldots, q\}$ and $s \in \{1, \ldots, T-k\}$. Let $\varepsilon = w_{k,k+s} - \overline{w}$, and define κ to be such that

$$v(w_k + \kappa) - v(w_k) = \phi_k \widehat{\xi}(e_k)\delta^s[v(\overline{w} + \varepsilon) - v(\overline{w})].$$

Given that $w_k \leq \underline{w} < \overline{w}$, the same argument as in the previous paragraph shows that $\kappa \leq \phi_k \widehat{\xi}(e_k)\delta^s \varepsilon$. Thus, the firm reduces the present discounted wage bill from employing the worker if it decreases $w_{k,k+s}$ to $\overline{w}$ and increases w_k to $w_k + \kappa$. This proves the desired result.

To finish the proof of the claim, suppose the firm pays the worker a wage of $\overline{w}$ once he produces high output for the first time. This implies that w_q must be such that

$$v(w_q) + \phi_q \widehat{\xi}(e_q)\delta(1 - \delta)^{-1}(1 - \delta^{T-q+1})[v(\overline{w}) - v(\underline{w})] - \widehat{c}(e_q) \geq v(\underline{w}). \tag{10}$$

Since $\phi_q \leq \phi_0$ and (A2) implies that

$$\phi_0(\alpha + \eta)\delta(1 - \delta)^{-1}(1 - \delta^T)[v(\overline{w}) - v(\underline{w})] - c \geq$$
$$\phi_0 \widehat{\xi}(e_q)\delta(1 - \delta)^{-1}(1 - \delta^T)[v(\overline{w}) - v(\underline{w})] - \widehat{c}(e_q),$$

we then have that $w_q \geq \widetilde{w}_1$. Now note that

$$v(w_q) + \phi_{q-1}\widehat{\xi}(e_q)\delta(1 - \delta)^{-1}(1 - \delta^{T-q+2})[v(\overline{w}) - v(\underline{w})] - \widehat{c}(e_q)$$
$$> v(w_q) + \phi_q \widehat{\xi}(e_q)\delta(1 - \delta)^{-1}(1 - \delta^{T-q+1})[v(\overline{w}) - v(\underline{w})] - \widehat{c}(e_q) \geq v(\underline{w}),$$

---

[21]Note that if the firm and the workers have different discount factors, then decreasing w_{k+1} to $\underline{w}$ and increasing w_k to $w_k + \kappa$ changes the present discounted wage bill from employing the worker by $\delta_f^k \varepsilon - \delta_w^k \kappa$, where δ_f is the firm's discount factor and δ_w is the workers' discount factor. Thus, the conclusion that $w_k \leq \underline{w}$ for all $k \in \{1, \ldots, q\}$ is necessary to minimize the present discounted wage bill from employing the worker holds as long as $\delta_f \geq \delta_w$. More generally, Claim 2 holds as long as $\delta_f \geq \delta_w$.

and so the worker accepts a one-period wage of w_q when he is of age $q - 1$ if he has not revealed himself to be of high ability by then. Thus, the firm can reduce the present discounted wage bill from employing the worker if $w_{q-1} > w_q$.

Suppose then that $w_{q-1} \leq w_q$, and for each $\varepsilon > 0$ let κ be such that

$$v(w_{q-1} + \kappa) - v(w_{q-1}) = [1 - \phi_{q-1}\widehat{\xi}(e_{q-1})]\delta\{v(w_q + \varepsilon) - v(w_q)\}.$$

Given that $w_{q-1} \leq w_q$, the same argument as in the first paragraph of the proof shows that $\kappa \leq [1 - \phi_{q-1}\widehat{\xi}(e_{q-1})]\delta\varepsilon$. Hence, if w_q is such that (10) holds with strict inequality, then the firm can reduce the present discounted wage bill from employing the worker by reducing w_q and increasing w_{q-1}. A straightforward induction argument then shows that a lower bound on the firm's present discounted wage bill from employing the worker is achieved if for all $k \in \{1, \ldots, q\}$, w_k satisfies

$$v(w_k) + \phi_k\widehat{\xi}(e_k)\delta(1 - \delta)^{-1}(1 - \delta^{T-k+1})[v(\overline{w}) - v(\underline{w})] - \widehat{c}(e_k) = v(\underline{w}).$$

In particular, $w_k \geq \widetilde{w}_1$ for all $k \in \{1, \ldots, q\}$. This establishes the claim. □

Step 2. We now show that there exists an equilibrium in which the firm dismisses age 1 workers after low output and these workers exert effort. Let σ^* be a strategy profile such that: (I) the firm offers $(\overline{w}, 0)$ to an incumbent it knows is of high ability if it is not committed to making him an offer; (II) the firm offers $(\widetilde{w}_1, 0)$ to the available age 1 worker if it has no incumbent or if its incumbent has always produced low output and the firm is not committed to employing him; (III) an incumbent who has revealed himself to be of the high type does not exert effort; (IV) an age 1 worker who accepts an offer of $(w, 0)$ exerts effort; (V) a worker accepts an offer if he is indifferent between taking it and collecting his outside option. Note that we have not completely specified the behavior of the equilibrium path for the firm and the workers. We do this later.

Observe that the behavior described by (I) and (III) to (V) is optimal. In order to prove that the behavior described by (II) is optimal, we need to prove that: (*A*) regardless of the worker's behavior, it is optimal for the firm to dismiss an incumbent who has never produced high output if it is not committed to employing him; and (*B*) if the firm makes an offer to the available age 1 worker, then it is optimal for the firm to offer $(\widetilde{w}_1, 0)$. The following facts are useful. First, if V^*

38

is the firm's payoff when play is given by σ^*, then

$$
V^* = (1 - \delta)[y(\phi_0, \xi) - \widetilde{w}_1] + \phi_0 \xi \left\{ \delta(1 - \delta^{T-1})[y(1, \alpha) - \overline{w}] + \delta^T V^* \right\} + (1 - \phi_0 \xi) \delta V^*
$$

$$
= (1 - \delta)[y(\phi_0, \xi) - \widetilde{w}_1] + \delta V^* + \phi_0 \xi \delta(1 - \delta^{T-1})[y(1, \alpha) - \overline{w} - V^*], \tag{11}
$$

where $\xi = \alpha + \eta$. Second, $y(\phi_0, \xi) - \widetilde{w}_1 < V^* < y(1, \alpha) - \overline{w}$ by (A1).

We begin with (A). Suppose the firm has an incumbent of age $k \geq 2$ who has never produced high output. Consider first the case in which the firm offers $(w^0, 0)$ to him. Given that (II) implies that the firm dismisses the incumbent after he produces low output, w^0 must be greater than $\widetilde{w}_1$ in order for the firm's offer to be accepted; we know from the main text that it is never optimal for the firm to make an offer that is rejected. Thus, the firm's payoff from offering $(w^0, 0)$ to the incumbent is bounded above by

$$
V_0 = (1 - \delta)[y(\phi_0, \xi) - \widetilde{w}_1] + \phi_0 \xi \left\{ \delta(1 - \delta^{T-k})[y(1, \alpha) - \overline{w}] + \delta^{T-k+1} V^* \right\} + (1 - \phi_0 \xi) \delta V^*,
$$

which is smaller than V^* since $V^* < y(1, \alpha) - \overline{w}$ and $k \geq 2$. We are done if we show that no matter how the incumbent behaves when employed, the firm's payoff from making an offer with $q \geq 1$ to him is smaller than V_0.

Suppose the firm makes an offer with $q \in \{0, \ldots, T - k\}$ to the incumbent. By (II), the firm dismisses the incumbent when he is of age $k + q + 1$ if he has not revealed himself to be of high ability by then. The same argument as in the proof of Claim 2 shows that a lower bound on the firm's present discounted wage bill from employing the worker under consideration is if it pays him a wage of $\widetilde{w}_1$ as long as he does not produce high output and a wage of $\overline{w}$ once he reveals himself to be of high ability. Since at best for the firm the incumbent exerts effort as long as he does not reveal himself to be of high ability, an upper bound on the firm's payoff is then given by

$$
V_q = \phi_0 \sum_{j=0}^{q} (1 - \xi)^j \xi \left\{ (1 - \delta^j)[\underline{y} - \widetilde{w}_1] + \delta^j(1 - \delta)[\overline{y} - \widetilde{w}_1] + \delta^{j+1}(1 - \delta^{T-k-j})[y(1, \alpha) - \overline{w}] \right.
$$

$$
\left. + \delta^{T-k+1} V^* \right\} + \left[\phi_0(1 - \xi)^{q+1} + 1 - \phi_0 \right] \left\{ (1 - \delta^{q+1})[\underline{y} - \widetilde{w}_1] + \delta^{q+1} V^* \right\}.
$$

Now observe that

$$(1 - \delta^{q+1})[\underline{y} - \widetilde{w}_1] + \delta^{q+1}V^* > (1 - \delta^{q+1})[\underline{y} - \widetilde{w}_1] + \delta^{q+1}\phi_0\xi \left\{ (1 - \delta)[\overline{y} - \widetilde{w}_1] \right.$$

$$\left. + \delta(1 - \delta^{T-k-q-1})[y(1, \alpha) - \overline{w}] + \delta^{T-k-q}V^* \right\} + \delta^{q+1}(1 - \phi_0\xi) \left\{ (1 - \delta)[\underline{y} - \widetilde{w}_1] + \delta V^* \right\}$$

$$= \phi_0\xi \left\{ (1 - \delta^{q+1})[\underline{y} - \widetilde{w}_1] + \delta^{q+1}(1 - \delta)[\overline{y} - \widetilde{w}_1] + \delta^{q+2}(1 - \delta^{T-k-q-1})[y(1, \alpha) - \overline{w}] \right.$$

$$\left. + \delta^{T-k+1}V^* \right\} + (1 - \phi_0\xi) \left\{ (1 - \delta^{q+2})[\underline{y} - \widetilde{w}_1] + \delta^{q+2}V^* \right\},$$

where the inequality follows from (11) and the fact that $V^* < y(1, \alpha) - \overline{w}$. Hence, since $\phi_0(1 - \xi)^{q+1} + 1 - \phi_0 = (1 - \xi)^{q+1} + (1 - \phi_0)[1 - (1 - \xi)^{q+1}]$, we have that

$$\left[\phi_0(1 - \xi)^{q+1} + 1 - \phi_0 \right] \left\{ (1 - \delta^{q+1})[\underline{y} - \widetilde{w}_1] + \delta^{q+1}V^* \right\} >$$

$$(1 - \phi_0)[1 - (1 - \xi)^{q+1}] \left\{ (1 - \delta^{q+2})[\underline{y} - \widetilde{w}_1] + \delta^{q+2}V^* \right\} + (1 - \xi)^{q+1}\phi_0\xi \left\{ (1 - \delta^{q+1})[\underline{y} - \widetilde{w}_1] \right.$$

$$\left. + \delta^{q+1}(1 - \delta)[\overline{y} - \widetilde{w}_1] + \delta^{q+2}(1 - \delta^{T-k-q-1})[y(1, \alpha) - \overline{w}] + \delta^{T-k+1}V^* \right\}$$

$$+ (1 - \xi)^{q+1}(1 - \phi_0\xi) \left\{ (1 - \delta^{q+2})[\underline{y} - \widetilde{w}_1] + \delta^{q+2}V^* \right\}.$$

Given that $(1 - \xi)^{q+1}(1 - \phi_0\xi) + (1 - \phi_0)[1 - (1 - \xi)^{q+1}] = 1 - \phi_0 + \phi_0(1 - \xi)^{q+2}$, we can then conclude that

$$\left[\phi_0(1 - \xi)^{q+1} + 1 - \phi_0 \right] \left\{ (1 - \delta^{q+1})[\underline{y} - \widetilde{w}_1] + \delta^{q+1}V^* \right\} >$$

$$[1 - \phi_0 + \phi_0(1 - \xi)^{q+2}] \left\{ (1 - \delta^{q+2})[\underline{y} - \widetilde{w}_1] + \delta^{q+2}V^* \right\} + (1 - \xi)^{q+1}\phi_0\xi \left\{ (1 - \delta^{q+1})[\underline{y} - \widetilde{w}_1] \right.$$

$$\left. + \delta^{q+1}(1 - \delta)[\overline{y} - \widetilde{w}_1] + \delta^{q+2}(1 - \delta^{T-k-q-1})[y(1, \alpha) - \overline{w}] + \delta^{T-k+1}V^* \right\}.$$

Thus, $V_q > V_{q+1}$, and so (A) is true.

We now establish (B). The above argument together with Claim 2 shows that if the firm makes an offer with $q \geq 1$ to an age 1 worker, then its payoff is bounded above by

$$y(\phi_0, \xi) - \widetilde{w}_1 + \phi_0\xi \left\{ \delta(1 - \delta^{T-1})[y(1, \alpha) - \overline{w}] + \delta^T V^* \right\} + (1 - \phi_0\xi)V_0,$$

which is smaller than V^*. This implies the desired result, since the firm's payoff in case it offers $(\widetilde{w}_1, 0)$ to an age 1 worker is V^* and $\widetilde{w}_1$ is the smallest one-period wage an age 1 worker accepts if he exerts effort and is dismissed after low output.

In order to finish the equilibrium construction, we need to complete the description of the behavior of the firm and the workers off the equilibrium path. By Lemma 1, we are left with

40

the task of determining: (*i*) the one-period wage offers the firm makes when it is committed to employing its incumbent and the incumbent has only produced low output; (*ii*) the effort choice of a worker of age $k \geq 2$ who has never produced high output; (*iii*) the effort choice of an age 1 worker who receives an offer with $q \geq 1$. Note that the behavior given by (II) is optimal regardless of how we specify (*i*) to (*iii*).

Consider a worker of age $k \geq 2$ who has not revealed himself to be of the high type, and let ϕ and π be, respectively, his reputation and private belief that he is of high ability. We proceed by induction in ℓ, the number of future periods the firm retains the worker if he only produces low output. Suppose first that $\ell = 0$, so that $k \geq 2$. The worker exerts effort if, and only if, $\pi\eta(1-\delta)^{-1}(1-\delta^{T-k})[v(\overline{w}) - v(\underline{w})] \geq c$. If committed to employing him, the firm offers the worker a one-period wage of $\max\{\widetilde{w}, w'\}$, where w' is the smallest one-period wage the firm can offer and $\widetilde{w}$ is the wage that makes the worker indifferent between accepting employment and taking his outside option when $\pi = \phi$. Suppose now that $\ell = 1$, so that by (II) the firm is committed to making an offer to the worker when he is of age $k+1$. The worker's choice of effort is optimal given π and his and the firm's behavior after he produces low output. As when $\ell = 0$, if committed to employing him, the firm offers the worker a one-period wage of $\max\{\widetilde{w}, w'\}$, where w' is the smallest one-period wage the firm can offer and $\widetilde{w}$ is the wage that makes the worker indifferent between accepting employment and taking his outside option when $\pi = \phi$. Moving backward, we completely determine (*i*), (*ii*), and (*iii*). This completes the proof that there exists an equilibrium in which the firm dismisses age 1 workers after low output and these workers exert effort.

Step 3. To finish Case 1, we show that the firm's payoff in any equilibrium in which it retains an age 1 worker after low output with positive probability is strictly less than V^*. Consider such an equilibrium, and let V^{**} be the firm's payoff in this equilibrium. By assumption, there exist $t \geq 1$, $q \in \{2, \ldots, T\}$, and probabilities λ_1 to λ_q, with $\lambda_1 < 1$, such that with probability λ_k the firm retains the age 1 worker it hires in period t until he is of age $k \in \{1, \ldots, q\}$ regardless of his performance. Since at best for the firm, the workers it hires exert effort as long as they do not reveal themselves to be of high ability, Lemma 2 and Claim 2 imply that $V^{**} \leq \overline{V}$, where $\overline{V}$ is

such that $\overline{V} = \sum_{k=1}^{q} \lambda_k T_k \overline{V}$ and

$$T_k \overline{V} = \phi_0 \sum_{j=0}^{k-1} (1-\xi)^j \xi \left\{ (1-\delta^j)[\underline{y} - \widetilde{w}_1] + \delta^j(1-\delta)[\overline{y} - \widetilde{w}_1] + \delta^{j+1}(1-\delta^{T-1-j})[y(1,\alpha) - \overline{w}] \right.$$
$$\left. + \delta^T \overline{V} \right\} + \left[\phi_0(1-\xi)^k + 1 - \phi_0 \right] \left\{ (1-\delta^k)[\underline{y} - \widetilde{w}_1] + \delta^k \overline{V} \right\}.$$

It is immediate to see that each map T_k is a contraction from $\mathbb{R}$ into $\mathbb{R}$, and so is the map $T = \sum_{k=1}^{q} \lambda_k T_k$. Now note, by (11), that $T_1 V^* = V^*$. Moreover, $T_k V^*$ is strictly decreasing in k by Step 2. Thus, since $\lambda_1 < 1$ by hypothesis, we have that $TV^* < T_1 V^* = V^*$, and so $T^{n+1} V^* < T^n V^*$ for all $n \geq 0$. Given that $T^n V^*$ converges to $\overline{V}$ by the contraction mapping theorem, we can then conclude that $\overline{V} < V^*$, which implies the desired result.

Case 2. Consider now the case in which (3) does not hold. We divide of proof of Case 2 into two steps. First, we establish that if not committed to doing so, the firm does not make an offer to a worker of age $k \geq 2$ who has only produced low output and does not exert effort in the current and future periods. Next, we use this fact to show that if (3) does not hold, then no worker ever exerts effort, so that in equilibrium the firm always dismisses an age 1 worker after low output.

Step 1. Suppose, by contradiction, that even though not committed to doing so, the firm makes an offer to an age $k \geq 2$ worker who has only produced low output and never exerts effort, and let $\ell \in \{1, \ldots, T-k+1\}$ be such that the worker is not employed when he is of age $k + \ell$ if he has not revealed himself to be of high ability by then. We know from the main text that it is not optimal for the firm to make an offer that is rejected.

It is immediate to see that the firm has a profitable deviation if $\ell = 1$. Suppose then, by induction, that there exists $q \in \{2, \ldots, T-k+1\}$ such that the firm has a profitable deviation if $\ell \leq q-1$ and let $\ell = q$. By the induction hypothesis, it must be that the firm commits to employing the worker for the next q periods: let $(w^0, (q, \{w^s\}_{s=1}^q))$ be the offer that it makes to the worker. The payoff to the firm is

$$V(\phi) = \phi \sum_{j=0}^{q} (1-\alpha)^j \alpha V(j) + [\phi(1-\alpha)^{q+1} + 1 - \phi]V(\emptyset),$$

where: (*i*) $\phi < \phi_0$ is the worker's reputation; (*ii*) $V(j)$ is the firm's payoff in case the worker reveals himself to be of the high type when he is of age $k + j$, with $j \in \{0, \ldots, q\}$; (*iii*) $V(\emptyset)$

is the firm's payoff in case the worker never produces high output. Since $\phi \sum_{j=0}^{q} (1-\alpha)^j \alpha = 1 - [\phi(1-\alpha)^{q+1} + 1 - \phi]$, we can rewrite $V(\phi)$ as

$$V(\phi) = \phi \sum_{j=0}^{q} (1-\alpha)^j \alpha [V(j) - V(\emptyset)] + V(\emptyset). \tag{12}$$

Now let $V < y(1,\alpha) - \overline{w}$ be the firm's payoff after it hires an age 1 worker. Since, by assumption, the worker in question does not exert effort when employed, we have that

$$V(\emptyset) = \sum_{s=0}^{q} \delta^s (1-\delta)[\underline{y} - w^s] + \delta^{q+1} V$$

and, by Lemma 1, that

$$V(j) = \sum_{s=0}^{j-1} \delta^s (1-\delta)[\underline{y} - w^s] + \delta^j (1-\delta)[\overline{y} - w^j]$$
$$+ \sum_{s=j+1}^{q} \delta^s (1-\delta)[y(1,\alpha) - \max\{w^s, \overline{w}\}] + \sum_{s=q+1}^{T-k} \delta^s (1-\delta)[y(1,\alpha) - \overline{w}] + \delta^{T-k+1} V$$

for all $j \in \{0, \ldots, q\}$. Since (A1) implies that $y(1,\alpha) - \overline{w} > \underline{y}$, it is immediate to see that $V(j) > V(\emptyset)$ for all $j \in \{0, \ldots, q\}$, so that $V(\phi)$ is strictly increasing in ϕ.

Consider then the following deviation for the firm: offer $(w^0, (q, \{w^s\}_{s=1}^{q}))$ to the available age 1 worker and then behave as if no deviation has occurred; in particular, the firm treats the age 1 worker as if he were the incumbent he is replacing. Since the age 1 worker has the option of never exerting effort, it is easy to see that he accepts the offer and then accepts a one-period wage of w^s when he is of age $s+1$ if he has only produced low output by then. The firm's payoff after this deviation is at least $V(\phi_0)$, which is the payoff it obtains if the age 1 worker never exerts effort. So, the firm has a profitable deviation when $\ell = q$. Thus, by induction, it is never optimal for the firm to make an offer to a worker of age $k \geq 2$ who has only produced low output and never exerts effort. This establishes the desired result.

Step 2. We can now prove that no worker exerts effort when (3) does not hold. By Lemma 1, we only need to consider workers of age $k = 1$ or workers of age $k \in \{2, \ldots, T-1\}$ who have never produced high output. Consider such a worker and suppose the firm employs him. The worker's

incentive-compatibility constraint for effort exertion is

$$-c + \pi(\alpha + \eta)\delta R(\overline{y}, \overline{e}|\pi, k) + [1 - \pi(\alpha + \eta)]\delta R(\underline{y}, \overline{e}|\pi, k)$$
$$\geq \pi\alpha\delta R(\overline{y}, \underline{e}|\pi, k) + (1 - \pi\alpha)\delta R(\underline{y}, \underline{e}|\pi, k), \tag{13}$$

where $\pi \leq \phi_0$ is the worker's belief that he is of high ability and $R(y, e|\pi, k)$ is the worker's continuation payoff if he chooses e and produces y, which depends on π and k. Note that π need not coincide with the worker's reputation. Given that $R(\overline{y}, \overline{e}|\pi, k) = R(\overline{y}, \underline{e}|\pi, k)$, we can rewrite (13) as

$$\pi\eta\delta\underbrace{[R(\overline{y}, \overline{e}|\pi, k) - R(\underline{y}, \overline{e}|\pi, k)]}_{\Delta_1} + \delta(1 - \pi\alpha)\underbrace{[R(\underline{y}, \overline{e}|\pi, k) - R(\underline{y}, \underline{e}|\pi, k)]}_{\Delta_2} \geq c. \tag{14}$$

Now let $1 \leq \ell \leq T - k + 1$ be the maximum number of periods (including the current one) that the worker is employed if he only produces low output. We proceed by induction in ℓ.

1) Suppose that $\ell = 1$, that is, the worker is dismissed after low output (and so $\Delta_2 = 0$). In this case, Lemma 1 implies that (14) reduces to

$$\pi\eta\delta(1 - \delta)^{-1}(1 - \delta^{T-k})[v(\overline{w}) - v(\underline{w})] \geq c,$$

which is not satisfied by assumption. Hence, the worker does not exert effort when $\ell = 1$.

2) Suppose now, by induction, that there exists $1 \leq \ell' \leq T - k + 1$ such that the worker does not exert effort if $\ell \leq \ell'$ and let $\ell = \ell' + 1$. By Step 1 and the induction hypothesis, the firm must commit to employ the worker for the next ℓ' periods. Let $\underline{\pi}(e) = [1 - \widehat{\xi}(e)]\pi/[1 - \pi\widehat{\xi}(e)]$ be the worker's updated belief about his ability after he produces low output, as a function of his effort choice. Moreover, let w_s be the lowest one-period wage the firm can offer him when he is of age $k + s + 1$, with $s \in \{0, \ldots, \ell - 2\}$. Note that it need not be the case that the worker's participation constraint is satisfied until he is of age $k + \ell - 1$ if he only produces low output. Let then $\widehat{\ell}(e)$ be such that if the worker's choice of effort (when of age k) is e, then he remains at the firm until he is of age $k + \widehat{\ell}(e) - 1$ if he only produces low output. Since $\underline{\pi}(\underline{e}) > \underline{\pi}(\overline{e})$, the induction hypothesis implies that $\widehat{\ell}(\overline{e}) \leq \widehat{\ell}(\underline{e}) = \ell$.

Following the same logic used in the derivation of (12), we have that

$$R(\underline{y}, e|\pi, k) = \underline{\pi}(e) \sum_{j=0}^{\widehat{\ell}(e)-2} (1 - \alpha)^j \alpha[R(j) - R(\emptyset, e)] + R(\emptyset, e),$$

44

where: (*i*) $R(j)$ is the worker's continuation payoff after low output if he produces high output for the first time when of age $k + j + 1$, where $j \in \{0, \ldots, \widehat{\ell}(e) - 2\}$; (*ii*) $R(\emptyset, e)$ is the worker's continuation payoff after low output if he never produces high output afterwards, which depends on his effort choice when of age k. By the induction hypothesis, the worker does not exert effort from the next period on. So, the firm has no incentive to offer the worker more than w_s when he is of age $k + s + 1$ if he has not produced high output by then. Lemma 1 then implies that

$$R(j) = \sum_{s=0}^{j} \delta^s v(w_s) + \delta^{j+1} \sum_{s=j+1}^{\ell'-1} v(\max\{\overline{w}, w_s\}) + \delta^{\ell'}(1-\delta)^{-1}(1 - \delta^{T-k-\ell'})v(\overline{w})$$

and

$$R(\emptyset, e) = \sum_{s=0}^{\widehat{\ell}(e)-2} \delta^s v(w_s) + \delta^{\widehat{\ell}(e)-1}(1-\delta)^{-1}(1 - \delta^{T-k-\widehat{\ell}(e)+1})v(\underline{w}).$$

Now observe that if $\widehat{\ell}(\overline{e}) < \ell$, then an option for the worker when he chooses $e = \underline{e}$ is to collect his outside option when he is of age $k + \widehat{\ell}(\overline{e})$ if he has not revealed himself to be of the high type by then. This implies that

$$R(\underline{y}, \underline{e}|\pi, k) \geq \underline{\pi}(\underline{e}) \sum_{j=0}^{\widehat{\ell}(\overline{e})-2} (1-\alpha)^j \alpha [R(j) - R(\emptyset, \overline{e})] + R(\emptyset, \overline{e})$$

Given that $R(j) > R(\emptyset, \overline{e})$ for all $j \leq \widehat{\ell}(\overline{e}) - 2$ and $\underline{\pi}(\overline{e}) < \underline{\pi}(\underline{e})$, we can conclude that $R(\underline{y}, \underline{e}|\pi, k) > R(\underline{y}, \overline{e}|\pi, k)$, which implies that $\Delta_2 \leq 0$.

To finish the induction argument, observe that $\Delta_1 \leq (1-\delta)^{-1}(1 - \delta_w^{T-k})[v(\overline{w}) - v(\underline{w})]$. It is immediate to see that this bound on Δ_1 holds if $w_s \leq \overline{w}$ for all $s \in \{0, \ldots, \ell' - 1\}$, given that $R(\underline{y}, \overline{e}|\pi, k) \geq (1-\delta)^{-1}(1-\delta_w^{T-k})v(\underline{w})$. Now recall that if $w_s > \overline{w}$ for some $s \in \{0, \ldots, \widehat{\ell}(\overline{e})-2\}$, then Lemma 1 implies that the worker's wage when he is of age $k + s + 1$ and has not revealed himself to be of high ability by then is the same as his wage when he produces high output before he is of age $k + s + 1$. Therefore, since the worker does not exert effort in any future period, we have that $\partial \Delta_1 / \partial w_s \leq 0$ if $w_s \geq \overline{w}$, which implies the desired result. Consequently, a necessary condition for (14) is that

$$\pi\eta\delta(1-\delta)^{-1}(1 - \delta^{T-k})[v(\overline{w}) - v(\underline{w})] \geq c,$$

45

which is not satisfied by assumption. We can then conclude, by induction, that if (3) does not hold, then no worker ever exerts effort when employed.

Since no worker exerts effort, an argument similar to the one used in Case 1 shows that the highest payoff the firm can obtain is when it dismisses age 1 workers after low output. Thus, in equilibrium, the firm always dismisses age 1 workers after poor performance. This completes the proof of Proposition 1. □

Claim 3. *Suppose that* (3) *holds with strict inequality. Let* $\phi_1(e) = [1 - \xi(e)]\phi_0/[1 - \phi_0\xi(e)]$, *with* $\xi(\underline{e}) = \alpha$ *and* $\xi(\overline{e}) = \alpha + \eta$, *and let* $\phi_2 = (1 - \alpha)^2\phi_0/[(1 - \alpha)^2\phi_0 + 1 - \phi_0]$. *Assume that*

$$v(\underline{w}) - \phi_1(\overline{e})\alpha(1 - \delta)^{-1}(1 - \delta^{T-2})[v(\overline{w}) - v(\underline{w})] < v(0); \tag{15}$$

$$\phi_1(\underline{e})\eta\delta(1 - \delta)^{-1}(1 - \delta^{T-2})[v(\overline{w}) - v(\underline{w})] > c; \tag{16}$$

$$\max\{\phi_2, \phi_1(\overline{e})\}\eta\delta(1 - \delta)^{-1}(1 - \delta^{T-2})[v(\overline{w}) - v(\underline{w})] < c; \tag{17}$$

$$\phi_1(\underline{e})(\alpha + \eta) > \phi_0\alpha(T - 1)/(T - 2); \tag{18}$$

$$\phi_0\eta\delta[v(\overline{w}) - v(0)] < c. \tag{19}$$

There exists an equilibrium in which the firm retains age 1 worker after low output.

Proof. Note that $\phi_1(e)$ is the reputation of an age 2 worker who chose e and produced low output when of age 1. Note also that ϕ_2 is the highest reputation possible for a worker of age $k \geq 3$ who has never produced high output. Condition (15), which can only be satisfied if v is bounded below, implies that both an age 1 worker and an age 2 worker who produced low output when of age 1 accept a one-period wage of zero. Condition (16) implies that an age 2 worker who exerted no effort and produced low output when of age 1 exerts effort if he is dismissed after low output. The same argument as in the proof of Proposition 1 shows that if (17) is satisfied, then both a worker of age $k \geq 3$ who has only produced low output and a worker of age 2 who exerted effort and produced low output when of age 1 do not exert effort when employed.

Consider then the strategy profile σ^{**} where: (I) the firm offers $(\overline{w}, 0)$ to an incumbent it knows is of high ability if it is not committed to making him an offer; (II) the firm offers $(w^0, 0)$, with $w^0 = 0$, to the available age 1 worker if it has no incumbent or if its incumbent is of age 3 or more, has always produced low output, and the firm is not committed to employing him; (III) the

firm offers $(w^0, 0)$, with $w^0 = 0$, to an age 2 incumbent who failed to produce high output if it is not committed to making him an offer; (IV) the firm offers $(\max\{\overline{w}, w\}, 0)$ to an incumbent if it is committed to offering him a one-period wage of at least w; (V) an age 1 worker, an age $k \geq 2$ worker who has revealed himself to be of high ability, an age $k \geq 3$ worker who has only produced low output, and an age 2 worker who exerted effort and produced low output when of age 1 do not exert effort; (VI) the effort choice of an age 2 worker who exerted no effort and produced low output when of age 1 is sequentially rational given the firm's offer and (I) to (V); (VII) a worker accepts an offer if indifferent between taking it or not. Notice that according to (VI), an age 2 worker who exerted no effort and produced low output when of age 1 exerts effort if offered $(w, 0)$ with $w \geq 0$. We claim that σ^{**} is an equilibrium if δ is sufficiently close to one.

It is easy to see that the firm is indifferent between offering $(w^0, 0)$ with $w^0 = 0$ and $(w^0, (1, w^1))$ with $w^0 = w^1 = 0$ to an age 1 worker. Thus, (II) is incentive-compatible if it is not optimal for the firm to make an offer with $q \geq 2$ to an age 1 worker. This, however, is a consequence of (V).[22]

Now observe, by Lemma 1 and (17), that we only need to show that it is never optimal for an age 1 worker to exert effort in order to establish that (V) is incentive-compatible. By assumption, the incentive-compatibility constraint for effort exertion for an age 1 worker is the same whether the firm offers him $(w^0, 0)$ with $w^0 \geq 0$ or $(w^0, (1, w^1))$ with $w^0 \geq 0$ and $w^1 = 0$, and is given by

$$-(1 - \delta)c + \phi_0(\alpha + \eta)\delta(1 - \delta^{T-1})v(\overline{w})$$
$$+[1 - \phi_0(\alpha + \eta)]\delta\left\{(1 - \delta)v(0) + \phi_1(\overline{e})\alpha\delta(1 - \delta^{T-2})v(\overline{w}) + [1 - \phi_1(\overline{e})\alpha]\delta(1 - \delta^{T-2})v(\underline{w})\right\}$$
$$\geq \phi_0\alpha\delta(1 - \delta^{T-1})v(\overline{w}) + (1 - \phi_0\alpha)\delta\left\{(1 - \delta)v(0)\right.$$
$$\left. + \phi_1(\underline{e})(\alpha + \eta)\delta(1 - \delta^{T-2})v(\overline{w}) + [1 - \phi_1(\underline{e})(\alpha + \eta)]\delta(1 - \delta^{T-2})v(\underline{w})\right\}. \tag{20}$$

Straightforward algebra shows that (20) reduces to

$$\phi_0\eta\delta[v(\overline{w}) - v(0)] \geq c,$$

---

[22]Observe that σ^{**} remains an equilibrium if we replace (II) with: (II') the firm offers $(w^0, (1, w^1))$ with $w^0 = w^1 = 0$ to the available age 1 worker if it has no incumbent or if its incumbent is of age 3 or more, has only produced low output, and the firm is not committed to employing him. However, the use of probation by the firm in this equilibrium is irrelevant in the sense that the firm would retain an age 1 worker who produces low output even if it were not committed to doing so.

which is not satisfied by (19). Now observe that the same argument as in the proof of Claim 4 below shows that if (20) is not satisfied, then an age 1 worker has no incentive to exert effort regardless of the firm's offer. Thus, (V) is indeed incentive-compatible.

We are done if we show that (III) is incentive-compatible. By (V), the best alternative to the firm is if it offers $(w^0, 0)$ with $w^0 = 0$ to an age 1 worker. Let V be the firm's payoff if it offers $(w^0, 0)$ with $w^0 = 0$ to an age 1 worker and V' be the firm's payoff if it offers $(w^0, 0)$ with $w^0 = 0$ to its (age 2) incumbent. We need to show that $V' > V$. Straightforward algebra shows that

$$V = y(1, \alpha) - \overline{w} - \frac{(1 - \delta)\Delta + \delta(1 - \delta)(1 - \phi_0\alpha)\Delta'}{1 - \delta^2 + \phi_0[\alpha + (1 - \alpha)(\alpha + \eta)]\delta^2(1 - \delta^{T-2})},$$

where $\Delta = y(1, \alpha) - \overline{w} - y(\phi_0, \alpha)$ and $\Delta' = y(1, \alpha) - \overline{w} - y(\phi_1(\underline{e}), \alpha + \eta)$. Now observe that

$$V' = (1 - \delta)y(\phi_1(\underline{e}), \alpha + \eta) + \phi_1(\underline{e})(\alpha + \eta)\delta(1 - \delta^{T-2})[y(1, \alpha) - \overline{w} - V] + \delta V.$$

Hence, $V' > V$ if, and only if,

$$\left\{1 + \phi_1(\underline{e})(\alpha + \eta)\frac{\delta(1 - \delta^{T-2})}{1 - \delta}\right\} \frac{(1 - \delta)\Delta + \delta(1 - \delta)(1 - \phi_0\alpha)\Delta'}{1 - \delta^2 + \phi_0[\alpha + (1 - \alpha)(\alpha + \eta)]\delta^2(1 - \delta^{T-2})} > \Delta',$$

which reduces to

$$[1 + \phi_1(\underline{e})(\alpha + \eta)(T - 2)]\Delta > [1 + \phi_0\alpha + \phi_0\alpha(T - 2)]\Delta'.$$

when $\delta \approx 1$. Since $\Delta > \Delta'$, as $\phi_1(\underline{e})(\alpha + \eta) > \phi_0\alpha$ by (18), the last inequality is satisfied if $\phi_1(\underline{e})(\alpha + \eta)(T - 2) > \phi_0\alpha(T - 1)$, which is true by (18). This completes the proof. $\qquad\square$

Claim 4. *In equilibrium, no worker exerts effort if*

$$\phi_0(1 - \alpha)(\gamma - \alpha)(1 - \delta)^{-1}\delta^2(1 - \delta^{T-2})[v(\overline{w}) - v(\underline{w})] < c.$$

Proof. Suppose the condition in the statement holds. By Lemma 1, we only need to consider a worker of age $k = 1$ or a worker of age $k \in \{2, \ldots, T - 2\}$ who has never produced high output; in the non-IID case, a worker of age $T - 1$ or more has no incentive to exert effort. Suppose the firm employs such a worker. The worker's incentive-compatibility constraint for effort is

$$-c + \pi\xi\delta R(\overline{y}, \overline{e}) + (1 - \pi\xi)\delta R(\underline{y}, \overline{e}) \geq \pi\xi\delta R(\overline{y}, \underline{e}) + (1 - \pi\xi)\delta R(\underline{y}, \underline{e}),$$

where $\pi \leq \phi_0$ is the worker's private belief that he is of high ability, ξ is the probability that he produces high output, and $R(y, e)$ is his continuation payoff if he chooses e and produces y. We omit the dependence of $R(y, e)$ on π and k for ease of notation. Notice that, unlike the IID case, the probability ξ is independent of the worker's current choice of effort. Since $R(\overline{y}, \overline{e}) = R(\overline{y}, \underline{e})$, we can rewrite the last inequality as

$$\delta(1 - \pi\xi) \left[R(\underline{y}, \overline{e}) - R(\underline{y}, \underline{e}) \right] \geq c. \tag{21}$$

Let $1 \leq \ell \leq T - k + 1$ be the maximum number of periods (including the current one) that the worker is employed if he never produces high output, where k is the worker's age. Notice that $R(\underline{y}, \overline{e}) = R(\underline{y}, \underline{e}) = (1 - \delta)^{-1}(1 - \delta^{T-k})v(\underline{w})$ if $\ell = 1$, in which case (21) is violated. Suppose then, by induction, that there exists $1 \leq \ell' \leq T - k - 1$ such that the worker does not exert effort if $\ell \leq \ell'$ and let $\ell = \ell' + 1$. Now let: (i) $\pi' = (1 - \xi)\pi/(1 - \pi\xi)$ be the worker's belief that he is of high ability after he produces low output; (ii) w_{k+1} be the wage offer to the worker when he is of age $k + 1$; (iii) R_+ be the worker's lifetime payoff when he is of age $k + 2$ if he produces low output when of age k and high output when of age $k + 1$; and (iv) $R_-(e)$ be the worker's lifetime payoff when he is of age $k + 2$ if he produces low output when of ages k and $k + 1$ and chooses e when of age k (by the induction hypothesis, the worker does not exert effort when of age $k + 1$ or older). Then,

$$R(\underline{y}, e) = v(w_{k+1}) + \delta\pi'\xi'(e)R_+ + [1 - \pi'\xi'(e)]\delta R_-(e),$$

where $\xi'(\overline{e}) = \gamma$ and $\xi'(\underline{e}) = \alpha$. The same argument as in the proof of Case 2 of Proposition 1 shows that $R_-(\overline{e}) \leq R_-(\underline{e})$.[23] This implies that a necessary condition for (21) is that

$$(1 - \pi\xi)\pi'(\gamma - \alpha)\delta^2(R_+ - R_-(\underline{e})) = \pi(1 - \xi)(\gamma - \alpha)\delta^2(R_+ - R_-(\underline{e})) \geq c.$$

To finish, notice, also by the same argument as in the proof of Case 2 of Proposition 1, that

$$R_+ - R_-(\underline{e}) \leq (1 - \delta)^{-1}(1 - \delta^{T-k-1})[v(\overline{w}) - v(\underline{w})].$$

Hence, (21) cannot be satisfied by assumption. Thus, by induction, the worker never exerts effort when employed if (A2) is not satisfied. $\qquad\square$

---

[23]By the induction hypothesis, the worker does not exert effort when of age $k+1$ or older. Thus, the only difference between the case in which the worker exerts effort when of age k and the case in which the worker does not exert effort when of age k is his private belief about his ability when he is of age $k + 2$.

Claim 5. *Assumption (A3) implies that an age 1 worker and age 2 worker who produce low output when of age 1 accept any offer by the firm.*

Proof. Consider an age 2 worker who produced low output when of age 1. A lower bound on his payoff if he is employed is

$$v(0) + \phi_1\alpha\delta(1-\delta)^{-1}(1-\delta^{T-2})[v(\overline{w}) - v(\underline{w})] + \delta(1-\delta)^{-1}(1-\delta^{T-2})v(\underline{w}). \qquad (22)$$

The worker obtains this payoff when he accepts a one-period wage of zero and collects his outside option in the next period regardless of his output. By (A4), (22) is greater than

$$v(\underline{w}) + \delta(1-\delta)^{-1}(1-\delta^{T-2}),$$

which is the payoff the worker obtains if he is not employed. Hence, the worker accepts any offer by the firm. To finish, note that (22) is also a lower bound on the payoff to an age 1 worker who accepts a one-period wage of zero. $\qquad \square$

Proof of Lemma 4: We know that an age 1 worker accepts any offer by the firm. Consider then an age 1 worker who is employed by the firm and suppose that: (*i*) the firm offers him $(w^0, 0)$ with $w^0 = 0$ in the next period if he produces low output, but dismisses him when he is of age 3 if he has not produced high output by then; (*ii*) his flow payoff is $v(\overline{w})$ in every period once he produces high output for the first time. Since the worker has no incentive to exert effort after he produces low output, his incentive-compatibility constraint for effort exertion is

$$-c + \phi_0\alpha\delta(1-\delta)^{-1}(1-\delta^{T-1})v(\overline{w})$$
$$+(1-\phi_0\alpha)\delta\left[v(0) + \phi_1\gamma\delta(1-\delta)^{-1}(1-\delta^{T-2})v(\overline{w}) + (1-\phi_1\gamma)\delta(1-\delta)^{-1}(1-\delta^{T-2})v(\underline{w})\right]$$
$$\geq \phi_0\alpha\delta(1-\delta)^{-1}(1-\delta^{T-1})v(\overline{w})$$
$$+(1-\phi_0\alpha)\delta\left[v(0) + \phi_1\alpha\delta(1-\delta)^{-1}(1-\delta^{T-2})v(\overline{w}) + (1-\phi_1\alpha)\delta(1-\delta)^{-1}(1-\delta^{T-2})v(\underline{w})\right],$$

which is implied by (A2). This proves the desired result. $\qquad \square$

Proof of Proposition 2: We divide the proof of Proposition 2 into two parts. We first show that if $\phi_2\gamma \leq \phi_0\alpha$, then no matter δ the firm retains an age $k \geq 3$ worker who has only produced low

output only if the firm is committed to doing so. We then show that if (7) holds, then there exists δ^* such that $\delta \geq \delta^*$ implies that the firm retains a worker of age 2 who produced low output when of age 1 only if the firm is committed to doing so.

Step 1. Since this is useful in Appendix B, we proceed in Step 1 without assuming (A2). For this, let $\widetilde{w}_1^N$ be such that

$$v(\widetilde{w}_1^N) = \max \left\{ v(\underline{w}) - \phi_0 \alpha \delta (1-\delta)^{-1} (1 - \delta^{T-1}) [v(\overline{w}) - v(\underline{w})], v(0) \right\}.$$

We know from the discussion in Subsection 3.1 that an age 1 worker accepts a one-period wage of $\widetilde{w}_1^N$. Moreover, $\widetilde{w}_1^N$ is the smallest one-period wage an age 1 worker accepts if he is not retained after poor performance. Notice that $\widetilde{w}_1^N = 0$ if (A2) holds.

Suppose the firm retains a worker of age $k \geq 3$ who has never produced high output despite not being committed to doing so. Consider first the case in which the firm dismisses the worker after low output, so that he does not exert effort. Since an upper bound on the probability the worker produces high output is $\phi_{k-1}\gamma$ and $\phi_{k-1}\gamma \leq \phi_2\gamma \leq \phi_0\alpha$, a straightforward modification of the argument in the previous paragraph shows that the worker does not accept a one-period wage smaller than $\widetilde{w}_1^N$. Thus, an upper bound on the firm's payoff, which is obtained if the worker exerted effort in the previous period, is

$$\overline{V}_0^k = (1-\delta)[y(\phi_{k-1}, \gamma) - \widetilde{w}_1^N] + \phi_{k-1}\gamma\delta \left\{ (1 - \delta^{T-k})[y(1,\alpha) - \overline{w}] + \delta^{T-k}V \right\} + (1 - \phi_{k-1}\gamma)\delta V$$

$$= (1-\delta)[y(\phi_{k-1}, \gamma) - \widetilde{w}_1^N] + \delta V + \phi_{k-1}\gamma\delta(1 - \delta^{T-k})[y(1,\alpha) - \overline{w} - V], \qquad (23)$$

where V is the firm's payoff after it hires an age 1 worker. Given that $\phi_{k-1}\gamma \leq \phi_2\gamma \leq \phi\alpha$ and $V < y(1,\alpha) - \overline{w}$ by the proof of Lemma 3, we have that

$$\overline{V}_0^k < (1-\delta)[y(\phi_0, \alpha) - \widetilde{w}_1^N] + \delta V + \phi_0\alpha\delta(1 - \delta^{T-1})[y(1,\alpha) - \overline{w} - V].$$

However, since an age 1 worker accepts any offer by the firm, the payoff on the right side of the above inequality is a lower bound on the firm's payoff if it offers $(w^0, 0)$ with $w^0 = 0$ to an age 1 worker and dismisses him after he produces low output. So, the firm has a profitable deviation.

Consider now the case in which there exists $\ell \in \{1, \ldots, T - k\}$ such that the firm employs the worker for ℓ more periods and retains him afterward only if he has revealed himself to be of

high ability. We know that if the worker has only produced low output before he is of age $k + j$, with $j \in \{0, \ldots, \ell\}$, then an upper bound on the probability that he produces high output when of age $k + j$ is $\phi_2 \gamma \leq \phi_0 \alpha$. An argument similar to the one used to establish Claim 2 in the proof of Proposition 1 then shows that: (*i*) a lower bound on the firm's present discounted wage bill from employing the worker is if it pays him a wage of $\widetilde{w}_1^N$ as long as he does not produce high output and a wage of $\overline{w}$ once he reveals himself to be of the high type. Now observe that: (*ii*) at best for the firm, the worker exerted effort in the previous period and exerts effort as long as he does not produce high output and is of age $k + \ell - 1$ or less. Recall that a worker who has only produced low output has no incentive to exert effort if he is dismissed after poor performance. Let $\overline{V}_\ell^k$ be the firm's payoff when the worker behaves as in (*ii*) and wage payments are as in (*i*). We claim that $\overline{V}_\ell^k < V$ for all $k \geq 3$ and $\ell \in \{1, \ldots, T - k\}$, so that the firm can profitably deviate by employing an age 1 worker.

In order to prove the claim in the previous paragraph, first note that

$$
\begin{aligned}
\overline{V}_1^k &= (1 - \delta)[y(\phi_{k-1}, \gamma) - \widetilde{w}_1^N] + \phi_{k-1}\gamma\delta \left\{ (1 - \delta)[y(1, \gamma) - \overline{w}] + \delta(1 - \delta^{T-k-1})[y(1, \alpha) - \overline{w}] \right. \\
&\quad \left. + \delta^{T-k}V \right\} + (1 - \phi_{k-1}\gamma)\delta \left\{ (1 - \delta)[y(\phi_k^*, \gamma) - \widetilde{w}_1^N] \right. \\
&\quad \left. + \phi_k^*\gamma \left\{ \delta(1 - \delta^{T-k-1})[y(1, \alpha) - \overline{w}] + \delta^{T-k}V \right\} + (1 - \phi_k^*\gamma)\delta V \right\},
\end{aligned}
$$

where $\phi_k^* = (1 - \gamma)\phi_{k-1}/(1 - \phi_{k-1}\gamma)$. Hence,

$$
\begin{aligned}
\overline{V}_1^k &= (1 - \delta)[y(\phi_{k-1}, \gamma) - \widetilde{w}_1^N] + \delta(1 - \delta)\left\{ \phi_{k-1}\gamma[y(1, \gamma) - \overline{w}] + (1 - \phi_{k-1}\gamma)[y(\phi_k^*, \gamma) - \widetilde{w}_1^N] \right\} \\
&\quad + \phi_{k-1}[\gamma + \gamma(1 - \gamma)]\delta^2(1 - \delta^{T-k-1})\left[y(1, \alpha) - \overline{w} - V\right] + \delta^2 V.
\end{aligned}
$$

Consider then the following deviation for the firm: offer $(w^0, 0)$ with $w^0 = \widetilde{w}_1^N$ to the available age 1 worker and dismiss him after poor performance. Since

$$
V \geq V' = (1 - \delta)[y(\phi_0, \alpha) - \widetilde{w}_1^N] + \delta V + \phi_0\alpha\delta(1 - \delta^{T-2})\left[y(1, \alpha) - \overline{w} - V\right],
$$

a lower bound on the firm's payoff from this deviation is

$$
\begin{aligned}
\widetilde{V} &= (1 - \delta)[y(\phi_0, \alpha) - \widetilde{w}_1^N] + \phi_0\alpha\delta \left\{ (1 - \delta^{T-1})[y(1, \alpha) - \overline{w}] + \delta^{T-1}V \right\} + (1 - \phi_0\alpha)\delta V' \\
&= (1 - \delta)[y(\phi_0, \alpha) - \widetilde{w}_1^N] + \delta(1 - \delta)\{\phi_0\alpha[y(1, \alpha) - \overline{w}] + (1 - \phi_0\alpha)[y(\phi_0, \alpha) - \widetilde{w}_1^N]\} \\
&\quad + [\phi_0\alpha + (1 - \phi_0\alpha)\phi_0\alpha] \left\{ \delta^2(1 - \delta^{T-2})[y(1, \alpha) - \overline{w} - V]\right\} + \delta^2 V.
\end{aligned}
$$

Now let $\Delta_w = \overline{w} - \widetilde{w}_1^N < \overline{w}$, and observe that

$$\phi_0\alpha + (1 - \phi_0\alpha)\phi_0\alpha = \phi_0\alpha(2 - \phi_0\alpha) > \phi_{k-1}\gamma(2 - \gamma) = \phi_{k-1}[\gamma + \gamma(1 - \gamma)];$$

$$\phi_{k-1}\gamma[y(1, \gamma) - \overline{w}] + (1 - \phi_{k-1}\gamma)[y(\phi_k^*, \gamma) - \widetilde{w}_1^N] = \underline{y} - \widetilde{w}_1^N + \phi_{k-1}\gamma\left(\Delta_y - \Delta_w\right);$$

$$\phi_0\alpha[y(1, \alpha) - \overline{w}] + (1 - \phi_0\alpha)[y(\phi_0, \alpha) - \widetilde{w}_1^N] = \underline{y} - \widetilde{w}_1^N + \phi_0\alpha\left[(1 + \alpha - \phi_0\alpha)\Delta_y - \Delta_w\right].$$

Thus, given that $\Delta_y > \overline{w}$ by (A1), we have that $\overline{V}_1^k < \widetilde{V} \le V$ for all $k \ge 3$. To finish, notice that

$$\overline{V}_\ell^k = (1 - \delta)[y(\phi_{k-1}, \gamma) - \widetilde{w}_1^N] + \phi_{k-1}\gamma\delta\left\{(1 - \delta)[y(1, \gamma) - \overline{w}] + \delta(1 - \delta^{T-k-1})[y(1, \alpha) - \overline{w}]\right.$$
$$\left. + \delta^{T-k}V\right\} + (1 - \phi_{k-1}\gamma)\delta\overline{V}_{\ell-1}^{k+1}$$

for all $k \ge 3$ and $\ell \in \{2, \ldots, T - k\}$. Hence, if $\overline{V}_{\ell-1}^{k+1} < V$ for $k \ge 3$ and $\ell \ge 2$, then $\overline{V}_\ell^k < V$ as well. This completes Step 1.

Step 2. Suppose the firm retains an age 2 worker despite not being committed to doing so and let ℓ be the number of periods that elapse before the firm dismisses the worker if he only produces low output. Consider first the case in which $\ell = 0$, so that the worker does not exert effort, and once again let V be the payoff to the firm after it hires an age 1 worker. Since $V < y(1, \alpha) - \overline{w}$ and an age 2 worker accepts a one-period wage of zero, an upper bound on the firm's payoff is

$$V' = (1 - \delta)y(\phi_1, \gamma) + \phi_1\gamma\left\{\delta(1 - \delta^{T-2})[y(1, \alpha) - \overline{w}] + \delta^{T-1}V\right\} + (1 - \phi_1\gamma)\delta V$$
$$= \delta V + (1 - \delta)\underbrace{\left\{y(\phi_1, \gamma) + \phi_1\gamma\frac{\delta(1 - \delta^{T-2})}{1 - \delta}[y(1, \alpha) - \overline{w} - V]\right\}}_{V_+}.$$

This upper bound is achieved when the worker exerted effort in the previous period.

Now observe that an option for the firm is to always offer $(w^0, (1, w^1))$ with $w^0 = w^1 = 0$ to age 1 workers. Since, by Step 1, the firm never retains a worker of age 3 who has only produced low output, Lemma 4 implies that an age 1 worker accepts $(w^0, (1, w^1))$ with $w^0 = w^1 = 0$ and exerts effort. Thus, $V \ge V_2$, where V_2 is the firm's payoff when it always offers $(w^0, (1, w^1))$ with $w^0 = w^1 = 0$ to age 1 workers. It is easy to see that V_2 satisfies the following recursion:

$$V_2 = (1 - \delta)y(\phi_0, \alpha) + \phi_0\alpha\left\{\delta(1 - \delta)[y(1, \gamma) - \overline{w}] + \delta^2(1 - \delta^{T-2})[y(1, \alpha) - \overline{w}] + \delta^T V\right\}$$
$$+ (1 - \phi_0\alpha)\left\{\delta(1 - \delta)y(\phi_1, \gamma) + \phi_1\gamma\left\{\delta^2(1 - \delta^{T-2})[y(1, \alpha) - \overline{w}] + \delta^T V\right\} + (1 - \phi_1\gamma)\delta^2 V\right\}.$$

53

Let $\Delta_1 = [y(1, \alpha) - \overline{w}] - y(\phi_1, \gamma)$. Solving the above recursion for V_2 we obtain that

$$V_2 = [y(1, \alpha) - \overline{w}] - \frac{(1 - \delta)\Delta + \delta(1 - \delta)\{\phi_0\alpha[y(1, \alpha) - y(1, \gamma)] + (1 - \phi_0\alpha)\Delta_1\}}{1 - \delta^2 + \phi_0[\alpha + \gamma(1 - \alpha)]\delta^2(1 - \delta^{T-2})}. \quad (24)$$

We claim that $V_+ < V_2$ when the firm is patient enough, so that in this case the firm is better off dismissing the age 2 worker. In order to show this, observe that

$$A = \lim_{\delta \to 1} V_2 = [y(1, \alpha) - \overline{w}] - \frac{\Delta + \phi_0\alpha[y(1, \alpha) - y(1, \gamma)] + (1 - \phi_0\alpha)\Delta_1}{2 + \phi_0[\alpha + \gamma(1 - \alpha)](T - 2)}$$

and that

$$B = \lim_{\delta \to 1} V_+ \leq y(\phi_1, \gamma) + \frac{\Delta + \phi_0\alpha[y(1, \alpha) - y(1, \gamma)] + (1 - \phi_0\alpha)\Delta_1}{2 + \phi_0[\alpha + \gamma(1 - \alpha)](T - 2)}\phi_1\gamma(T - 2).$$

We are done if we show that

$$A - B = \Delta_1 - \frac{1 + \phi_1\gamma(T - 2)}{2 + \phi_0[\alpha + \gamma(1 - \alpha)](T - 2)}[\Delta - \phi_0\alpha(\gamma - \alpha)\Delta_y + (1 - \phi_0\alpha)\Delta_1] > 0.$$

Now notice that $A > B$ if, and only if,

$$\{2 + \phi_0[\alpha + \gamma(1 - \alpha)](T - 2)\}\Delta_1 > [1 + \phi_1\gamma(T - 2)][\Delta - \phi_0\alpha(\gamma - \alpha)\Delta_y + (1 - \phi_0\alpha)\Delta_1],$$

which reduces to

$$[1 + \phi_0\alpha(T - 1)]\Delta_1 - [1 + \phi_1\gamma(T - 2)][\Delta - \phi_0\alpha(\gamma - \alpha)\Delta_y] > 0.$$

Since $\Delta_1 = \Delta + (\phi_0\alpha - \phi_1\gamma)\Delta_y$, we have that $A > B$ if, and only if,

$$\phi_1\gamma\{[1 + \phi_0\alpha(T - 1)]\Delta_y + [\Delta - \phi_0\alpha(\gamma - \alpha)\Delta_y](T - 2)\}$$
$$< \phi_0\alpha\{(T - 1)\Delta + [1 + \phi_0\alpha(T - 1)]\Delta_y + (\gamma - \alpha)\Delta_y\}.$$

Straightforward algebra shows that this last inequality is satisfied by (7).

Consider now the case in which $\ell \geq 1$. Since this is useful in Appendix B, we proceed without assuming (A2). As in Step 1, an upper bound on the firm's payoff is obtained when the worker: (*i*) exerts effort until he is of age $2 + \ell - 1$ if he does not produce high output by then; and (*ii*) accepts a wage offer of $\widetilde{w}_1^N$ as long as he does not reveal himself to be of high ability. Let $\widetilde{V}_\ell$ be the firm's payoff in this case. We claim that $\widetilde{V}_\ell \geq \widetilde{V}_{\ell+1}$ for all $\ell \geq 1$.

In order to prove the claim, let $\phi^*_{\ell+1}(\xi) = (1-\xi)(1-\gamma)^{\ell-1}\phi_1/[(1-\xi)(1-\gamma)^{\ell-1}\phi_1 + 1 - \phi_1]$, where $\xi = \alpha$ if the worker did not exert effort when of age 1 and $\xi = \gamma$ otherwise. By construction, $\phi^*_{\ell+1}(\xi)$ is the worker's reputation when of age $2+\ell$ if he behaves as in (*i*) and never produces high output. Moreover, let

$$C = (1-\delta)[y(\phi^*_{\ell+1}(\xi),\gamma) - \widetilde{w}^N_1] + \phi^*_{\ell+1}(\xi)\gamma\left\{\delta(1 - \delta^{T-\ell-2})[y(1,\alpha) - \overline{w}] + \delta^{T-\ell-1}V\right\}$$
$$+ (1 - \phi^*_{\ell+1}(\xi)\gamma)\delta V$$

and

$$D = (1-\delta)[y(\phi^*_{\ell+1}(\xi),\gamma) - \widetilde{w}^N_1] + \phi^*_{\ell+1}(\xi)\gamma\left\{\delta(1 - \delta)[y(1,\gamma) - \overline{w}]\right.$$
$$+ \delta^2(1 - \delta^{T-\ell-3})[y(1,\alpha) - \overline{w}] + \delta^{T-\ell-1}V\right\} + (1 - \phi^*_{\ell+1}(\xi)\gamma)\left\{\delta(1 - \delta)[y(\phi^*_{\ell+2}(\xi),\gamma) - \widetilde{w}^N_1]\right.$$
$$+ \phi^*_{\ell+2}(\xi)\gamma\left\{\delta^2(1 - \delta^{T-\ell-3})[y(1,\alpha) - \overline{w}] + \delta^{T-\ell-1}V\right\} + (1 - \phi^*_{\ell+2}(\xi)\gamma)\delta^2 V\right\}.$$

It is easy to see that $C \geq D$ implies that $\widetilde{V}_\ell \geq \widetilde{V}_{\ell+1}$. Given that

$$V \geq (1-\delta)[y(\phi_0,\alpha) - \widetilde{w}^N_1] + \phi_0\alpha\left\{\delta(1 - \delta^{T-2})[y(1,\alpha) - \overline{w}] + \delta^{T-1}V\right\} + (1 - \phi_0\alpha)\delta V$$
$$\geq (1-\delta)[y(\phi_0,\alpha) - \widetilde{w}^N_1] + \phi_0\alpha\left\{\delta(1 - \delta^{T-\ell-3})[y(1,\alpha) - \overline{w}] + \delta^{T-\ell-2}V\right\}$$
$$+ (1 - \phi_0\alpha)\delta V, \tag{25}$$

the same algebra as in the proof of Step 1 shows that $C \geq D$ if

$$\phi^*_{\ell+1}(\xi)\gamma[y(1,\alpha) - \overline{w}] + (1 - \phi^*_{\ell+1}(\xi)\gamma)y(\phi_0,\alpha)$$
$$\geq \phi^*_{\ell+1}(\xi)\gamma[y(1,\gamma) - \overline{w}] + (1 - \phi^*_{\ell+1}(\xi)\gamma)y(\phi^*_{\ell+2}(\xi),\alpha). \tag{26}$$

Since $y(1,\alpha) - \overline{w} > y(\phi_0,\alpha)$ by (A1) and $(1 - \phi^*_{\ell+1}(\xi))\phi^*_{\ell+2}(\xi) = \phi^*_{\ell+1}(\xi)(1-\gamma)$, straightforward algebra shows that (26) holds given that $\phi^*_{\ell+1}(\xi)\gamma \leq \phi_2\gamma \leq \phi_0\alpha$. This shows that $\widetilde{V}_\ell \geq \widetilde{V}_{\ell+1}$.

To finish the argument, suppose the firm deviates as follows: it offers $(w^0, 0)$ with $w^0 = \widetilde{w}^N_1$ to the worker under consideration and replaces him with an age 1 worker if he produces low output. Denote the firm's payoff from this deviation by $\widetilde{V}_0$. By (25),

$$\widetilde{V}_0 \geq (1-\delta)[y(\phi_1,\gamma) - \widetilde{w}^N_1] + \delta(1-\delta)\left\{\phi_1\xi[y(1,\alpha) - \overline{w}] + (1 - \phi_1\xi)[y(\phi_0,\alpha) - \widetilde{w}^N_1]\right\}$$
$$+ (\phi_1\xi + (1 - \phi_1\xi)\phi_0\alpha)\{\delta^2(1 - \delta^{T-3})[y(1,\alpha) - \overline{w} - V]\} + \delta^2 V.$$

Given that

$$\widetilde{V}_1 = (1 - \delta)[y(\phi_1, \xi) - \widetilde{w}_1^N] + \phi_1 \xi \delta \left\{ (1 - \delta)[y(\phi_1, \gamma) - \overline{w}] + \delta(1 - \delta^{T-2})[y(1, \alpha) - \overline{w}] \right.$$

$$\left. + \delta^{T-3} V \right\} + (1 - \phi_1 \xi)\delta \left\{ (1 - \delta)[y(\phi_2^*(\xi), \gamma) - \widetilde{w}_1^N] + \phi_2^*(\xi)\gamma \left\{ \delta(1 - \delta^{T-2})[y(1, \alpha) - \overline{w}] \right. \right.$$

$$\left. \left. + \delta^{T-3} V \right\} + (1 - \phi_2^*(\xi)\gamma)\delta V \right\},$$

the same argument as above shows that $\widetilde{V}_0 \geq \widetilde{V}_1$ if

$$\phi_1 \xi y(1, \gamma) + (1 - \phi_1 \xi)y(\phi_2^*(\xi), \gamma) < \phi_1 \xi y(1, \alpha) + (1 - \phi_1 \xi)y(\phi_0, \alpha).$$

Since the right side of the last inequality is strictly decreasing in ξ, straightforward algebra shows that the inequality is satisfied by (7). Thus, the firm has a profitable deviation. □

Proof of Corollary 2: We claim that the firm's payoff from making an offer with $q > 1$ to an age 1 worker is smaller than the firm's payoff from making an offer of $q = 1$ to the same worker. By Step 1 in the proof of Proposition 2, this is clearly the case if an age 1 worker does not exert effort when the firm makes him an offer with $q > 1$. Suppose, then, that an age 1 worker exerts effort when the firm makes him an offer with $q > 1$. Since an age 1 worker also exerts effort when offered $q = 1$, Step 2 in the proof of Proposition 2 shows that the firm's payoff when it offers $q = 1$ to an age 1 worker is greater than when it offers $q > 1$ to the same worker. □

Appendix B: Omitted Details

Non-IID Case When $e_- = \overline{e}$ For Age 1 Workers

Here we show that our results in the non-IID case survive when $e_- = \overline{e}$ for age 1 workers if we impose the following strengthening of (A1):

(A1') $y(\phi_0, \gamma) + \phi_0[y(1, \gamma) - y(1, \alpha)] < y(1, \alpha) - \overline{w}$.

Recall that (A1) is equivalent to $y(\phi_0, \alpha) + \phi_0[y(1, \gamma) - y(1, \alpha)] < y(1, \alpha) - \overline{w}$. Since now age 1 workers are more productive, it is natural that we need a stronger condition than (A1) to ensure that the firm always wants to retain a worker known to be of the high type. Notice that (A1') is satisfied if ϕ_0 is sufficiently small.

First observe that Lemma 3 is still valid when we impose (A1'). Indeed, (A1') just states that the present value of the output generated by a worker who has not revealed himself to be of the high type is less than $y(1, \alpha) - \overline{w}$. Now consider the following auxiliary environment: $e_- = \underline{e}$ for age 1 workers and the probability that a worker is of high ability is ϕ_0' such that

$$\frac{(1 - \gamma)\phi_0}{1 - \phi_0\gamma} = \frac{(1 - \alpha)\phi_0'}{1 - \phi_0'\alpha}. \tag{27}$$

For fixed α and γ, equation (27) defines a one-to-one map between ϕ_0 and ϕ_0'. By construction, the auxiliary environment is equivalent to the original one. Indeed, the reputation of an age 1 worker who produces low output is the same in both cases. Moreover, since $\phi_0' < \phi_0$, (A1') implies that $y(\phi_0', \gamma) < y(1, \alpha) - \overline{w}$, so that (A1) holds in the auxiliary environment. Since $\lim_{\phi_0 \to 0} \phi_0' = 0$, we are done if all model's restrictions are satisfied when the ex ante probability that workers are of high ability is small.

Non-IID Case Without (A3)

Here we consider the non-IID case without Assumption (A3). For ease of exposition, we assume that $v(0) = -\infty$. We can adapt the analysis to the case in which $v(0)$ is finite, but low enough that (A3) is violated. We also assume that v is strictly concave. The first result we establish is the equivalent of Proposition 2, namely, conditions under which an age 1 worker exerts effort only if offered probation. Then, we establish the equivalent of Proposition 3, namely, conditions under which the use of probation has informational value.

Suppose the firm offers $(w^0, (1, w^1))$ to an age 1 worker. We know from the proof of Step 1 of Proposition 2 that the worker is dismissed when of age 3 if he has not revealed himself to be of high ability by then. Thus, the worker has no incentive to exert effort when of age 2 if he produces low output when of age 1. Now let: (*i*) $R_+ = (1 - \delta)^{-1}(1 - \delta^{T-2})v(\overline{w})$; (*ii*) $R_- = (1 - \delta)^{-1}(1 - \delta^{T-2})v(\underline{w})$; and (*iii*) $R_0 = (1 - \delta)^{-1}(1 - \delta^{T-1})v(\underline{w})$. The worker's IC constraint for effort exertion when of age 1 is given by

$$-c + (1 - \phi_0\alpha)\delta \max\left\{v(w^1) + \phi_1\gamma\delta R_+ + (1 - \phi_1\gamma)\delta R_-, R_0\right\}$$
$$\geq (1 - \phi_0\alpha)\delta \max\left\{v(w^1) + \phi_1\alpha\delta R_+ + (1 - \phi_1\alpha)\delta R_-, R_0\right\}. \tag{28}$$

Notice that (A2) implies that (28) is satisfied when $v(w^1) + \phi_1 \alpha \delta R_+ + (1 - \phi_1 \alpha) \delta R_- \geq R_0$.

We claim that a necessary condition for the firm to minimize the cost of employing the worker and inducing him to exert effort when of age 1 is:

$$v(w^1) + \phi_1 \gamma \delta R_+ + (1 - \phi_1 \gamma) \delta R_- > R_0; \tag{29}$$

$$v(w^1) + \phi_1 \alpha \delta R_+ + (1 - \phi_1 \alpha) \delta R_- \leq R_0. \tag{30}$$

That (29) is necessary is immediate, for otherwise the left side of (28) is smaller than its right side. Next, suppose that (30) is not satisfied. We claim that in this case the worker accepts $(w^0, (1, w^1))$ with $w^0 = w^1$. Indeed, if the worker accepts this offer, (28) and (30) imply that his lifetime payoff is bounded below by

$$v(w^1) + \phi_0 \alpha \delta (1 - \delta)^{-1} (1 - \delta^{T-1}) v(\overline{w}) + (1 - \phi_0 \alpha) \delta R_0$$
$$\geq v(w^1) + \phi_0 \alpha \delta R_+ + (1 - \phi_0 \alpha) \delta R_- + \delta (1 - \delta)^{-1} (\delta^{T-2} - \delta^{T-1}) v(\underline{w})$$
$$\geq (1 - \delta)^{-1} (1 - \delta^T) v(\underline{w}),$$

where the second inequality follows from (30) and the fact that $\phi_0 \alpha > \phi_1 \alpha$. Thus, it must be that $w^0 \leq w^1$, for otherwise the firm can reduce w^0 without violating both (28) and the worker's participation constraints. Now define $\varepsilon > 0$ and κ to be such that:

$$v(w^1 - \varepsilon) + \phi_1 \alpha \delta (R_+ - R_-) + \delta R_- = R_0;$$
$$v(w^0 + \kappa) - v(w^0) = (1 - \phi_0 \alpha) \delta [v(w^1) - v(w^1 - \varepsilon)].$$

Since $w^0 \leq w^1$, the same argument as in the proof of Claim 2 in the proof of Proposition 1 shows that $\kappa \leq (1 - \phi_0 \alpha) \delta \varepsilon < \delta \varepsilon$. Hence, if the firm reduces w^1 to $w^1 - \varepsilon$ and increases w^0 to $w^0 + \kappa$, it reduces the cost of employing the worker while still satisfying (28) and his participation constraints.

Consider then the following minimization problem, which we denote by (M):

$$\min \quad w^0 + \delta (1 - \phi_0 \alpha) w^1$$

$$\text{s.t.} \quad -c + (1 - \phi_0 \alpha) \delta \left[v(w^1) + \phi_1 \gamma \delta R_+ + (1 - \phi_1 \gamma) \delta R_- \right] \geq (1 - \phi_0 \alpha) \delta R_0,$$

$$v(w^0) - c + \phi_0 \alpha \delta (1 - \delta)^{-1} (1 - \delta^{T-1}) v(\overline{w})$$
$$+ (1 - \phi_0 \alpha) \delta \left[v(w^1) + \phi_1 \gamma \delta R_+ + (1 - \phi_1 \gamma) \delta R_- \right] \geq (1 - \delta)^{-1} (1 - \delta^T) v(\underline{w}).$$

Since v is strictly concave, it is easy to see that (M) has a unique solution, denoted by $(\widehat{w}_1^N, \widehat{w}_2^N)$. The argument in the previous paragraph shows that the offer with $q = 1$ which minimizes the cost of employing the worker while inducing him to exert effort when of age 1 is the one with $w^0 = \widehat{w}_1^N$ and $w^1 = \widehat{w}_2^N$. Let λ and μ be, respectively, the multipliers associated with the first and second inequality constraints of (M). The (necessary and sufficient) Kuhn-Tucker conditions for (M) are the following:

$$-1 + \mu v'(w^0) = 0; \tag{31}$$

$$-1 + \mu v'(w^1) + \lambda v'(w^1) = 0. \tag{32}$$

Equation (31) implies that $\mu = 1/v'(w^0) > 0$, in which case we can rewrite (32) as

$$\frac{v'(w^1)}{v'(w^0)} + \lambda v'(w^1) = 1.$$

Since a necessary condition for the last equation is that $v'(w^1) \leq v'(w^0)$, we can then conclude that $\widehat{w}_2^N \geq \widehat{w}_1^N$. Since $\mu > 0$, straightforward algebra shows that

$$v(\widehat{w}_1^N) \leq v(\underline{w}) - \phi_0 \alpha \delta (1 - \delta)^{-1} (1 - \delta^{T-1})[v(\overline{w}) - v(\underline{w})]; \tag{33}$$

$$v(\widehat{w}_2^N) \geq v(\underline{w}) - \phi_1 \gamma \delta (1 - \delta)^{-1} (1 - \delta^{T-2})[v(\overline{w}) - v(\underline{w})] + \frac{c}{(1 - \phi_0 \alpha)\delta}. \tag{34}$$

The following result shows that $\widehat{w}_1^N = \widehat{w}_2^N$ when c is small enough.

Lemma 5. *There exists* $\overline{c} = \overline{c}(\delta) > 0$ *such that if* $c < \overline{c}$, *then* $\widehat{w}_1^N = \widehat{w}_2^N = \widehat{w}$, *where* $\widehat{w}$ *is such that*

$$v(\widehat{w}) = v(\underline{w}) - \frac{(1 - \phi_0 \alpha)\phi_1 \gamma}{1 + \delta(1 - \phi_0 \alpha)} \delta^2 (1 - \delta)^{-1} (1 - \delta^{T-2})[v(\overline{w}) - v(\underline{w})] + \frac{c}{1 + \delta(1 - \phi_0 \alpha)}.$$

Proof. Suppose that $\lambda > 0$, in which case both (33) and (34) hold with equality. Given that $(1 - \delta^{T-1})/(1 - \delta^{T-2})$ is increasing in δ, we have that

$$v(\widehat{w}_1^N) - v(\widehat{w}_2^N) \geq \phi_0 \alpha \psi \frac{T - 1}{T - 2} \delta[v(\overline{w}) - v(\underline{w})] - \frac{c}{(1 - \phi_0 \alpha)\delta}.$$

Thus, if $c < \overline{c}$, where

$$\overline{c} = (1 - \phi_0 \alpha)\phi_0 \alpha \psi \frac{T - 1}{T - 2} \delta^2 [v(\overline{w}) - v(\underline{w})],$$

we have that $\widehat{w}_2^N < \widehat{w}_1^N$, a contradiction. Therefore, $\lambda = 0$ when $c < \overline{c}$, in which case we know $\widehat{w}_1^N = \widehat{w}_2^N$. The desired result follows from the fact that $\mu > 0$. $\qquad\square$

Let V_2 denote the firm's payoff when it always offers $(\widehat{w}_1^N, (1, \widehat{w}_2^N))$ to age 1 workers. The same algebra as in Appendix A shows that

$$V_2 = y(1, \alpha) - \overline{w} - \frac{(1-\delta)\Delta + \delta(1-\delta)\{\phi_0\alpha[y(1,\alpha) - y(1,\gamma)] + (1 - \phi_0\alpha)\Delta_1\}}{1 - \delta^2 + \phi_0[\alpha + \gamma(1-\alpha)]\delta^2(1 - \delta^{T-2})},$$

where $\Delta = y(1,\alpha) - \overline{w} - [y(\phi_0, \alpha) - \widehat{w}_1^N]$ and $\Delta_1 = y(1,\alpha) - \overline{w} - [y(\phi_1, \gamma) - \widehat{w}_2^N]$. By construction, V_2 is the highest payoff possible to the firm when it always makes an offer with $q = 1$ to the age 1 workers that induces them to exert effort. We can now establish the equivalent of Proposition 2 to the case under consideration. In order to do so, let $\underline{w}_2^N$ be such that

$$v(\underline{w}_2^N) = v(\underline{w}) - \phi_1\gamma\delta(1-\delta)^{-1}(1 - \delta^{T-2})[v(\overline{w}) - v(\underline{w})].$$

Notice that $\underline{w}_2^N < \widehat{w}_2^N$. By construction, $\underline{w}_2^N$ is the smallest wage an age 2 worker who exerted effort and produced low output when of age 1 accepts if he is not retained after poor performance. Moreover, let $\Delta_w = \widehat{w}_1^N - \widehat{w}_2^N \leq 0$ and $\Delta_w' = \widehat{w}_2^N - \underline{w}_2^N > 0$. Notice that if $\widehat{w}_1^N = \widehat{w}_2^N$, so that $\Delta_w = 0$, then the second condition in Proposition 4 below is more stringent than condition (7) in Proposition 2.

Proposition 4. *Suppose that $\phi_2\gamma \leq \phi_0\alpha$ and*

$$\phi_1\gamma < \min\left\{\phi_0\alpha\left[1 + \frac{\eta + \Upsilon + \phi_0\alpha\eta(T-2)}{1 + \phi_0\alpha + [\Upsilon + \phi_0\alpha(1-\eta)](T-2)}\right.\right.$$
$$\left.\left. - \frac{[1 + \phi_0\alpha(T-1)](\Delta_w/\Delta_y) + \{2 + \phi_0[\alpha + \gamma(1-\alpha)](T-2)\}(\Delta_w'/\Delta_y)}{1 + \phi_0\alpha + [\Upsilon + \phi_0\alpha(1-\eta)](T-2)}\right], \frac{\alpha(1-\alpha)(1-\phi_0)}{1 - \phi_0\alpha}\right\}.$$

There exists $\delta^* \in (0,1)$ such that if $\delta > \delta^*$, then in equilibrium the firm retains a worker of age $k \geq 2$ who has never produced high output only if committed to doing so.

Proof. From Step 2 in the proof of Proposition 2, we just need to show that there exists $\delta^* \in (0,1)$ such that if $\delta > \delta^*$, then the firm's payoff from employing an age 1 worker is greater than the firm's payoff from making an offer with $q = 0$ to the incumbent. Let V' be the highest payoff possible for the firm if it makes an offer with $q = 0$ to the age 2 incumbent. It is easy to see that

$$V' = \delta V + (1-\delta)\left\{[y(\phi_1, \gamma) - \underline{w}_2^N] + \phi_1\gamma\frac{\delta(1 - \delta^{T-2})}{1 - \delta}[y(1,\alpha) - \overline{w} - V]\right\},$$

where V is the firm's payoff after it hires an age 1 worker. Since $V \geq V_2$, we are done if we show that there exists $\delta^* \in (0, 1)$ such that if $\delta > \delta^*$, then

$$[y(\phi_1, \gamma) - \underline{w}_2^N] + \phi_1 \gamma \frac{\delta(1 - \delta^{T-2})}{1 - \delta}[y(1, \alpha) - \overline{w} - V_2] < V_2.$$

The same algebra as in the proof of Proposition 2 shows that when $\delta = 1$, the last condition reduces to the second condition in the statement of the proposition. This completes the proof. $\square$

Proposition 4 admits the following corollary. Its proof is identical to the proof of Corollary 2.

Corollary 3. *In equilibrium, the firm offers at most one period of probation to age 1 workers.*

Let V_{nc} be the firm's greatest payoff when it always makes an offer with $q = 0$ to age 1 workers. It follows immediately from the above corollary that if the conditions of Proposition 4 hold, then commitment to employment is beneficial if, and only, if $V_c := V_2 > V_{nc}$. Straightforward algebra shows that $V_c - V_{nc} = \rho_O + \rho_I$, where ρ_O is the same as in the main text and

$$\rho_I = \left\{ \frac{(1 - \delta)(1 + \xi)}{1 - \delta + \phi_0 \alpha \delta(1 - \delta^{T-1})} - \frac{(1 - \delta)[1 + \delta(1 - \phi_0 \alpha)(1 + \psi - \chi)]}{1 - \delta^2 + \phi_0[\alpha + \gamma(1 - \alpha)]\delta^2(1 - \delta^{T-2})} \right\} \Delta,$$

where $\xi = \Delta''_w / \Delta$, $\chi = \Delta_w / \Delta$, and ψ is defined as in the main text. Straightforward algebra shows that commitment to employment has informational value if, and only if,

$$\phi_1 \gamma > \phi_0 \alpha \frac{(1 + \psi - \chi)}{1 + \xi} \frac{1 - \delta^{T-1}}{1 - \delta^{T-2}} + \frac{(\psi - \chi)}{1 + \xi} \frac{(1 - \delta)}{\delta(1 - \delta^{T-2})}.$$

Notice that when the firm is patient, the above condition reduces to

$$\phi_1 \gamma > \phi_0 \alpha \frac{(1 + \psi - \chi)}{1 + \xi} \frac{T - 1}{T - 2}.$$

We can now state the equivalent of Proposition 3 for the case under consideration.

Proposition 5. *Suppose that $\phi_2 \gamma \leq \phi_0 \alpha$ and that*

$$\phi_0 \alpha \frac{(1 + \psi - \chi)}{1 + \xi} \frac{T - 1}{T - 2} < \phi_1 \gamma$$

$$< \min \left\{ \phi_0 \alpha \left[1 + \frac{\eta + \Upsilon + \phi_0 \alpha \eta(T - 2)}{1 + \phi_0 \alpha + [\Upsilon + \phi_0 \alpha(1 - \eta)](T - 2)} \right. \right.$$

$$\left. - \frac{[1 + \phi_0 \alpha(T - 1)](\Delta_w / \Delta_y) + \{2 + \phi_0[\alpha + \gamma(1 - \alpha)](T - 2)\}(\Delta'_w / \Delta_y)}{1 + \phi_0 \alpha + [\Upsilon + \phi_0 \alpha(1 - \eta)](T - 2)} \right], \left. \frac{\alpha(1 - \alpha)(1 - \phi_0)}{1 - \phi_0 \alpha} \right\}.$$

There exists $\delta^{**} \in (0, 1)$ such that if $\delta > \delta^{**}$, then: (i) the firm always offers probation to age 1 workers in equilibrium; (ii) commitment to employment has informational value.

Appendix C: Evidence on Probationary Practices in Academia

Here, we include the questionnaire we submitted to the Office for Faculty Affairs of a private U.S. research university (whose Economic Department is ranked among the top ten in the U.S.) and to the Dean's Office of the business school of a public U.S. research university (whose Economic Department is ranked among the top thirty in the U.S.) in order to inquire about their probationary practices for hires at the faculty level.

In the case of the private U.S. institution, we interviewed a senior staff member of the Office for Faculty Affairs. In the case of the business school of the public U.S. institution, we interviewed a senior staff member of the Dean's Office with responsibilities in human resource management. Representatives from both institutions required that their identities and the identities of their institutions remain confidential as a pre-condition for answering our questionnaire. We report the actual questionnaire below.

Questionnaire

We are seeking evidence in favor or against the following interpretation of hiring, retention, and promotion practices in academia and other professional service industries:

(a) Organizations like law firms and academic institutions rarely fire employees before the end of the probationary period of employment, as the probationary contract entails a binding commitment on the part of the organization to employ an individual;

(b) A probationary contract of employment does not necessarily entail a binding commitment on the part of an organization to consider an employee for subsequent promotion or different assignment based on certain criteria being satisfied. When the contract also entails this latter type of commitment, it is a commitment disjoint from the commitment under (a).

We are interested in the probationary employment contract offered to assistant professors at your institution. In this regard, our questions are:

1) What is the length of the probationary appointment offered to newly hired assistant professors?

2) Are different types of employment contracts offered to tenure-track and non-tenure track assistant professors?

3) Are assistant professors who are full-time employees ever hired under a different contract than the tenure-track probationary one?

4) Most academic institutions with so-called six-year tenure clocks do not offer new assistant professors six-year contracts, but rather grant employment for, say, the first three years. However, in this case the contract explicitly states that at the end of the first three years, the person will be considered for reappointment for another three years.

- How do probationary appointments for assistant professors work at your university?

- What are their main characteristics in terms of the commitment to employ/inability to fire an individual that the university assumes under such contracts?

- Which rules or regulations apply to probationary appointments?

- From the university viewpoint, what is the purpose of the first probationary phase of three years? Is the possibility to dismiss candidates who do not seem promising enough at the end of these first three years an important consideration affecting the renewal decision?

5) Can assistant professors on tenure-track be fired during the probationary period of employment at a date different from the renewal dates of the probationary appointment, that is, at a date different from the end of the first three or six years of probation?

- If so, under which conditions?

- How common/rare is this type of termination?

- Do rules and practices differ across assistant professors on tenure-track and not on tenure-track?

6) As for promotion and tenure:

- Does the probationary contract offered to assistant professors explicitly commit the university to consider an assistant professor for promotion and tenure at the end of the probationary period of employment or consideration for promotion and tenure is part of an informal agreement that the university customarily honors? In this latter case, is consideration for promotion and tenure at all envisioned in the probationary contract offered to assistant professors?

- When the commitment to consider an assistant professor for promotion and tenure at the end of the probationary period is explicit, what are the specific rules or regulations that force the university to consider an assistant professor for promotion and tenure?

- More generally, which rules and regulations govern consideration of an assistant professor for promotion and tenure?

- Do these rules differ in the case of promotion and tenure of an assistant professor currently employed at another institution and considered at your university for employment in a tenured position? If so, how?