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. The ability of a worker is initially unknown to both the worker and the firm, and a worker’s effort affects the information on ability that is conveyed by 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 encourages workers to exert effort, thus improving the firm’s ability to identify their talent. We argue that the incentive value of commitment to retention helps explain the use of fixed probationary appointments in environments in which there exists uncertainty about ability.

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

*We are grateful to Steve Matthews and Jan Eeckhout for their generous comments. We also benefited from conversations with Maria Goltsman, Hari Govindan, and Matt Mitchell, and from the input of various seminar participants. Enoch Hill provided excellent research assistance. Braz Camargo gratefully acknowledges financial support from CNPq. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.
1 Introduction

It has long been 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—to determine whether new workers are suited to handle the responsibilities associated with their jobs.¹

The use of probationary appointments is common in management consulting, the legal profession, academia, and government bureaucracies. 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, even though the performance of a worker is informative about ability, the quality of this information is affected by the worker’s behavior on the job. For instance, whether a researcher is successful on a project depends not only on his talent, but also on the nature and scope of the project, which help signal the researcher’s potential. Likewise, whether a restructuring

¹Probationary 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 employ a worker: the firm can dismiss the worker only at the end of this period. Probationary contracts in academia and in professional service industries are a primary example.
project is successful in addressing the needs of a client firm depends both on the ability of the consultant working on it and on his effort. Second, information about ability may be revealed only gradually by performance. The completion of a research paper normally requires a few months or years. 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 probation as a short-term commitment to employment and assume that incentive pay is not feasible. There are two reasons for not considering incentive pay. First, as mentioned above, the use of probationary appointments is common in academia and bureaucracies, where the use of incentive pay is quite limited. Thus, 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. Thus, when a worker exerts effort, good performance is a more precise signal of high ability, so the worker’s output becomes more informative about his ability.

Each worker in the market has an outside option available that increases with his reputation, which we model as the firm’s belief that the worker is of high ability. Then, workers

---

2Note 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.
of higher reputation are perceived to be more productive, but can be employed only at a higher wage. The value of a worker to the firm nevertheless increases with his 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.

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 in case his performance is good, as good performance increases his reputation. Thus, a worker effectively “invests” in his reputation when he exerts effort.

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. In our environment, effort also makes performance more informative about ability, which allows the firm to better sort the workers that it employs. In this case, the firm benefits from offering probation to a new worker if commitment to employment strengthens the worker’s motive to exert effort due to career concerns. Commitment to employment, however, is costly because it prevents the firm from dismissing the worker if he performs poorly.

In this paper, we identify circumstances under which the use of probationary appointments can be beneficial to a firm by contrasting two cases 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 affects both current and future output, but the effect on future output is stronger. We refer to the second case as the non-IID case.

We begin by showing that probation never increases reputational incentives for effort in the IID case. More precisely, we show that a worker who has not revealed himself to be of high ability has the strongest incentive to exert effort if he is dismissed after low output. Intuitively, if such a worker is retained after low output, he has an additional opportunity to prove his ability. This possibility weakens the worker’s incentive to exert effort in the first place. Thus, the firm cannot benefit from commitment to employment in the IID case.
The firm can benefit from offering probation to new hires in the non-IID case, though. Commitment to employment is valuable in this case because in the non-IID case, the incentive problem of a new worker is compounded by the time separation of costs and benefits typical of investment problems. When effort affects mostly future output, the worker can gain from investing 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, as a worker whose initial performance is poor is less attractive to the firm than a new worker. Thus, when it offers probation, the firm overcomes its incentive to dismiss underperforming workers, thereby inducing new hires to generate more precise information about their ability through effort. In other words, the firm can benefit from probation in the non-IID case, since commitment to employment solves a time inconsistency problem. Moreover, 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 benefit from commitment to employment can be positive, even ignoring the resulting increase in output.

The rest of the paper is organized as follows. We discuss the related literature in the remainder of this section and introduce the model in the next section. Section 3 contains some auxiliary results. We consider the IID case in Section 4 and the non-IID case in Section 5. We discuss some of our modeling choices in Section 6 and conclude in Section 7. The Appendices contain omitted proofs and details.

*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 (1982, 1999); see also Scharfstein and Stein (1990). A few papers have studied the interplay between career concerns and incentive contracts. For instance, Gibbons and Murphy (1992) analyze the optimal com-
bination of explicit incentives (from piece-rate contracts) and career concerns incentives.\textsuperscript{3} Mukherjee (2008) investigates the substitutability between incentives from career concerns and incentives from implicit bonus contracts when a firm can decide whether to 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 that it considers a framework in which the key decision for the firm is whether to retain the worker it currently employs.\textsuperscript{4}

The retention problem that the firm faces in our setting is an example of a multi-armed bandit, the sequential sampling problem of a decision maker choosing between a number of alternatives with uncertain rewards (here, the different workers the firm can employ).\textsuperscript{5} 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.\textsuperscript{6} This difference is crucial for our analysis, since the firm can never gain from offering commitment to employment if the incentive problem is absent, that is, if the workers’ 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

\textsuperscript{3}Andersson (2002) extends the analysis in Gibbons and Murphy (1992) to the case in which contracts are unobservable.

\textsuperscript{4}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.

\textsuperscript{5}See Berry and Fristedt (1985) for an exposition of the theory of multi-armed bandits. In our 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.

\textsuperscript{6}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.
by a strategy that does not involve commitment to employment.

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. 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 an unverifiable investment in firm-specific human capital. Although it is ex ante optimal for a firm to reward such an investment, the decision to do so is not ex post optimal. Anticipating this, workers underinvest in human capital. 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. 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 when contracts are incomplete.

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

---

7 See Guasch and Weiss (1981, 1982) for early references on labor market models of hiring and compensation in the presence of adverse selection. Bull and Tedeschi (1989) model probation as a period during which a firm commits to monitor 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 (1998) 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 not accept employment.

8 O’Flaherty and Siow (1992) provide an alternative explanation for the use of up-or-out rules. They analyze a model of on-the-job screening and firm growth, and show that the optimal retention decision is an up-or-out rule. Carmichael (1988) shows how the institution of tenure in academia can induce incumbent professors to hire individuals who are potentially more talented than themselves.
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 promote the most qualified workers, so as to reduce the incentive to engage in influence activities. Waldman (2003) analyzes a time inconsistency problem that arises due to the dual role of promotions: they serve both to reward performance and to 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. Ghosh and Waldman (2010) study the choice between standard promotion practices and up-or-out contracts when promotions serve the dual role described above. They show that up-or-out contracts are superior when the level of a worker’s firm-specific human capital is low.9

Our paper differs from the previous literature 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 our 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 work that studies how promotion practices affect incentives (Prendergast (1993), Waldman (2003), and Ghosh and Waldman (2010)), we assume there exists a single task and many workers who are heterogeneous in their ability to perform this task. Thus, the firm faces an experimentation problem. Finally, the moral hazard problem we study is nonstandard. A worker’s current effort not only affects current output, as commonly assumed, but also affects the quality of the information about ability that performance conveys, as well as future output.10

9Our 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.
10In Ghosh and Waldman (2010), a worker’s choice of effort affects the mean of the posterior belief about his ability, but not the variance. Thus, unlike our framework, effort does not make a worker’s performance more informative about his ability.
2 Environment

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

Workers. Workers enter the market sequentially, one in each period. They have a concave and strictly increasing utility function $v : \mathbb{R}_+ \rightarrow \mathbb{R}$, live for $T \geq 3$ periods once they enter the market, and discount future utility at rate $\delta \in (0, 1)$. 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 the high-ability type as the worker’s reputation. Thus, all workers enter the market with the same reputation.

In each period of employment, a worker can either exert effort ($e$), incurring a cost $c > 0$, or not ($\epsilon$), and can produce either high ($\gamma$) or low ($y$) output. A worker’s choice of effort, which is unobservable, affects his output. So, a worker’s belief about his ability need not coincide with his reputation. We consider two cases. In the first case, our benchmark, the output produced by a worker is affected only by the worker’s current choice of effort. In the second case, a worker’s choice of effort has an impact on both current and future output. 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 current and previous choice of effort, 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 $y \in \{y, \gamma\}$. In the IID case,

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

where $\alpha \in (0, 1)$, $\eta(e) = 0$, and $\eta(\tau) = \eta > 0$. In the non-IID case,

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

\[11\] 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.
where $\eta(\bar{e}, e) \geq \eta(e, \epsilon)$, $\eta(e, \epsilon) > \eta(\epsilon, e)$, and $\eta(\epsilon, e) = 0$. So, in the non-IID case a worker of high ability who exerts effort in a period increases his expected output in the next period.

Notice that in both the IID and non-IID cases, a low-ability 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 probability that a low-ability type worker produces high output is greater than zero but still sufficiently small, so that good performance is a strong signal of high ability.

Also observe from (1) and (2) that in both cases, effort increases the probability of good performance only if the worker is of the high-ability type. Thus, effort increases the likelihood that a high-ability worker reveals himself to be of high ability. In this precise sense, effort makes a worker’s performance more informative about his ability, and this is beneficial to the firm. The informational role of effort is central to our analysis.

As in the career concerns literature, workers in the market have available an outside option paying a wage $\omega_R$ 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 reality it may not always be profitable or possible for a firm to rehire a worker who has previously separated from it. We return to this point in Section 6. We assume that $\omega_R(1) = \bar{\omega}$ and $\omega_R(\phi) = \omega_R(\phi_0) = \overline{\omega}$ for all $\phi \leq \phi_0$. So, as long as a worker fails to produce high output, his outside option remains constant at $\overline{\omega}$, but it increases to $\bar{\omega}$ the first time he produces high output. The restriction that a worker’s outside option cannot decrease below $\overline{\omega}$ captures situations in which a worker can seek employment in an alternative labor market where his ability has no value. 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 and the workers discount future payoffs at the same rate. 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 $\Pi < y - \omega$.\footnote{More generally, the key assumption is that the firm is capacity constrained, and so cannot absorb all}
So, the firm prefers to employ a worker it knows is of the low-ability type rather than not employ 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 (and who is still alive), 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 wage \(w^0\) at the end of the period if he accepts employment. The firm can also commit to make minimum one-period wage offers to a worker of age \(k \leq T - 1\) for the 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 \((w^0, (q, \{w^\sigma\}_{s=1}^q))\) 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 \(\{w^\sigma\}_{s=1}^q\) of minimum future one-period wage offers, where \(w^\sigma\) is the minimum one-period wage offer \(s\) periods into the future. Observe that the list \(\{w^\sigma\}_{s=1}^q\) needs to be specified only if \(q \geq 1\). Notice also that if the firm offers \((w^0, (q, \{w^\sigma\}_{s=1}^q))\) with \(q \geq 1\) to a worker of age \(k\), then it must propose \((w^0, 0)\) with \(w^0 \geq w^\sigma\) to him when he is of age \(k + s\).

Observe that a promise of future wage offers to a worker does not necessarily represent a commitment to employ him, as the promised minimum one-period wage offers may not be sufficiently attractive to induce the worker to accept them. In such case, the firm can induce the worker to quit before the end of the commitment period. However, as we show in the next section, in equilibrium the firm never makes an offer that it expects to be rejected. Thus, a commitment to future wage offers is indeed a commitment to employment: if on the path of play 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. We say the firm offers “probation” to an age 1 worker when it offers him \((w^0, (q, \{w^\sigma\}_{s=1}^q))\) with \(q \geq 1\).

\(^{13}\)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 to 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.

workers in the market. We can extend our analysis to the case where the firm has a finite number of vacancies.
Let $y(\phi, \xi) = \phi \xi \bar{y} + (1 - \phi \xi) y = y + \phi \xi \Delta y$, where $\Delta y = \bar{y} - y$, be the expected output of a worker of reputation $\phi$ when the probability that he produces high output is $\xi$ if he is of the high-ability type. Moreover, let $\bar{\eta} = \max_{\xi, e} \{ \eta(e), \eta(e_{-}, e) \}$. By construction, $y(\phi_0, \alpha + \bar{\eta})$ is an upper bound on the expected output of an age 1 worker. Since the reputation of an age $k \geq 2$ worker who has never produced high output is smaller than $\phi_0$, $y(\phi_0, \alpha + \bar{\eta})$ is also an upper bound on the expected output of such a worker. The following restriction is a maintained assumption:

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

Since the firm cannot hire a worker known to be of the high-ability type for less than $\bar{w}$, (A1) implies that the firm’s flow payoff from employing such a worker is always greater than its flow payoff from employing a worker who has not revealed himself to be of high ability.\footnote{As we show in the next section, 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.}

A necessary condition for (A1) is that $\phi_0(\alpha + \bar{\eta}) < \alpha$ and $\Delta y > \bar{w}$. Notice that (A1) is satisfied if high-ability workers are sufficiently scarce, that is, if $\phi_0$ is sufficiently small.

Assumption (A1) implies that the firm always prefers a worker of the high-ability type to an age 1 worker, regardless of effort. This assumption plays an important role in our analysis as 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 it 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.
worker who either does not receive an offer or rejects one collects his outside option, and so does the firm 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 $\Sigma_f$ 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, the sets $\Sigma_w(t)$ are the same for all workers. We denote them by $\Sigma_w$. 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 $\sigma_w : \mathbb{N} \rightarrow \Sigma_w$, where $\sigma_w(t)$ is the behavior strategy of the worker who enters the market in period $t$. A strategy profile $(\sigma_w, \sigma_f)$ is worker symmetric if $\sigma_w(t)$ is independent of $t$. We restrict attention to worker symmetric perfect Bayesian equilibria.\(^{15}\) In what follows, we use the expression “in equilibrium” as a shorthand for the expression “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. Intuitively, by accepting employment at the firm, the worker has the opportunity to prove that he is of high ability and thus increase his future compensation. Indeed, 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

\(^{15}\)One alternative to the assumption that workers follow symmetric strategies is to assume that in each period, the identity of the worker who becomes available is random. The approach we follow is simpler.
high-ability 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

$$R = v(w) + \pi \alpha \delta (1 - \delta)^{-1}(1 - \delta^{T-k})v(\overline{w}) + (1 - \pi \alpha)\delta (1 - \delta)^{-1}(1 - \delta^{T-k})v(\overline{w})$$

when he accepts a one-period wage of $w$. Now observe that if

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

then $R$ is at least equal to $(1 - \delta)^{-1}(1 - \delta^{T-k+1})v(w)$. Given that the right side of the above inequality is smaller than $v(\overline{w})$, the worker is willing to accept (an offer with) a one-period wage smaller than his outside option. In particular, if $v$ is bounded below, then

$$v(w) - \phi \alpha(T - k)[v(\overline{w}) - v(w)] < v(0)$$

when $T - k$ is large. Thus, an age $k$ worker accepts a one-period wage of zero if he is sufficiently young and patient.

As it turns out, 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 the ones 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 an age 1 worker is also less expensive to employ.

### 3.2 Auxiliary Results

We now prove three results that are useful for the analysis that follows. See Appendix A for proofs. The first result we establish, which follows from the assumption that $\Pi < \underline{y} - \underline{w}$, is that the firm makes an offer in every period and never makes an offer that it believes will be rejected. In particular, the firm never collects its outside option on the path of play.

**Lemma 1.** In equilibrium the firm always makes an offer and never makes an offer that it expects to be rejected.
A consequence of the argument in the proof of Lemma 1 is that an incumbent known to be of the high-ability type always accepts an offer if the one-period wage is at least \( \bar{w} \). Note also that in our setting, a worker has no incentive to exert effort if there exists no uncertainty about his ability. Hence, whenever the firm makes an offer to an incumbent it knows is of the high-ability type, the one-period wage it offers is the lowest possible. The following result formalizes these two observations.

**Lemma 2.** Suppose the firm has an incumbent it knows is of the high-ability type, and let \( w' \) be the smallest one-period wage the firm can offer him if it is committed to do so. The following holds in equilibrium: (i) the firm never offers the incumbent a one-period wage greater than \( \max\{\bar{w}, w'\} \); (ii) if the firm makes an offer to the incumbent, then it never commits to future one-period wage offers greater than \( \bar{w} \); (iii) the incumbent never exerts effort.

A sketch of the proof of Lemma 2 is as follows. Consider a worker known to be of the high-ability 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, for 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 \( \max\{\bar{w}, w'\} \) to the worker. The desired result now follows from a backward induction argument.

The last result we establish in this section follows from the fact that workers use symmetric strategies. It states that the firm’s (expected present discounted) continuation payoff from hiring an age 1 worker is independent of calendar time. For ease of notation, let \( V(h|\sigma) \) denote the firm’s lifetime payoff after a history \( h \) when the strategy profile under play is \( \sigma \).

**Lemma 3.** If \( \sigma \) is an equilibrium, then \( V(h|\sigma) = V(h'|\sigma) \) for any two histories \( h \) and \( h' \) for the firm after which it makes an offer to the available age 1 worker.
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 is that the firm does not benefit from commitment to employment in this case.\(^\text{16}\)

We start by characterizing the firm’s retention decision. By Lemma 2, the firm always obtains the same (flow) payoff when it employs a worker of high ability at wage \(\overline{w}\). By assumption (A1), this payoff is also the highest flow payoff the firm can obtain. Moreover, by Lemma 3, the firm’s continuation payoff when it hires an age 1 worker is independent of calendar time. Thus, the firm always retains an incumbent it knows is of high ability. The result below formalizes this discussion; see Appendix B for a proof. For convenience, in the remainder of the paper we sometimes say that the firm “offers \(w\)” to a worker whenever it makes an offer in which the one-period wage is \(w\).

**Lemma 4.** 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 do so. In equilibrium, the firm offers \(\max\{\overline{w}, w'\}\) to this worker, who accepts the offer.

We now establish the main result of this section. In order to do so, consider an age 1 worker who is not retained if his performance is poor, that is, if he produces low output. Since this is possible only if the firm has not offered probation to the worker, Lemma 2 implies that if the worker produces high output, then his wage in every subsequent period is \(\overline{w}\). Hence, \(\phi_0 \eta \delta (1 - \delta)^{-1}(1 - \delta^{T-1})[v(\overline{w}) - v(w)]\) 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 \(\delta(1 - \delta)^{-1}(1 - \delta^{T-1})[v(\overline{w}) - v(w)]\) is the increase in the worker’s lifetime payoff if he produces high output. Thus, when

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

\(^{16}\)We show in Section 6 that our results hold as long as high output is a strong enough signal of high ability. 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.
an age 1 worker who is not retained after low output has an incentive to exert effort. We can then establish the following result.

**Proposition 1.** Suppose (3) holds. There exists an equilibrium $\sigma^*$ 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 $\sigma^*$. Suppose now (3) does not hold. The firm dismisses age 1 workers after low output in every equilibrium.

The following corollary is an immediate consequence of Proposition 1.

**Corollary 1.** Suppose (3) holds. The highest payoff the firm can attain in equilibrium is under $\sigma^*$.

Thus, whether condition (3) is satisfied or not, the greatest payoff the firm can obtain in equilibrium is when age 1 workers are dismissed after low output. In other words, commitment to employment is not beneficial to the firm.

Here we just sketch the proof of Proposition 1. The details are in Appendix B. Suppose first that (3) is satisfied, so that it is incentive compatible for an age 1 worker to exert effort if he is not retained after poor performance. In this case, we prove that there exists an equilibrium in which the firm dismisses age 1 workers after low output and these workers exert effort. The desired result now 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, 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, an age $k \geq 2$ worker is worse than an age 1 worker in continuation payoff terms since the likelihood that he reveals himself to be of high ability is lower, and in case he proves himself to be of the high-ability type, he can work at the firm for a shorter period of time.

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 2 that we only need to consider workers who have not produced high output. Next 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

$$-c + \pi(\alpha + \eta)\delta R(y, e|\pi, k) + [1 - \pi(\alpha + \eta)]\delta R(y, e|\pi, k) \geq \pi \alpha \delta R(y, e|\pi, k) + (1 - \pi \alpha)\delta R(y, e|\pi, k),$$

where $R(y, e|\pi, k)$ is his continuation payoff if he chooses $e$ and produces $y$, which depends on $\pi$ and $k$. Since the worker reveals himself to be of high ability if he produces high output, Lemma 2 implies that $R(y, e|\pi, k) = R(y, e|\pi, k)$. Then we can rewrite (4) as

$$\pi \eta \delta [R(y, e|\pi, k) - R(y, e|\pi, k)] + (1 - \pi \alpha)\delta [R(y, e|\pi, k) - R(y, e|\pi, k)] \geq c.$$

As it turns out, $\Delta_1$ is bounded above by $(1 - \delta)^{-1}(1 - \delta^{T-k})[v(\bar{w}) - v(w)]$. Indeed, since $R(y, e|\pi, k) \geq (1 - \delta)^{-1}(1 - \delta^{T-k})v(w)$, given that the worker can always collect his outside option, we immediately obtain the above bound on $\Delta_1$ if the worker’s wage after he reveals himself to be of high ability is always $\bar{w}$. Now observe that once the worker reveals himself to be of high ability, the firm pays him a wage $w$ greater than $\bar{w}$ in some period $t$ only if it committed to do 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 $\Delta_1$ 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 $R(y, e|\pi, k) \leq R(y, e|\pi, k)$. 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, $\Delta_2$ is at most zero. We can then conclude that a necessary condition for (5) is

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

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 2.** 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 $\phi_1$ is the highest reputation possible for an age 2 worker who produced low output when of age 1. By assumption, we have that

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

But then, a straightforward modification of the argument preceding Corollary 2 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.
Suppose now that (3) holds with a strict inequality. A natural question to ask is whether it is still the case that there exists no equilibrium in which age 1 workers are retained after low output. We show in Appendix B that the answer is no, that is, there exists an equilibrium in which age 1 workers are retained after poor performance.

In order to understand why multiple equilibria arise, first notice that any equilibrium in which an age 1 worker is retained after poor performance has the feature that such a worker does not exert effort if the firm offers him \( q = 0 \). Otherwise, the firm would be able to increase its payoff by offering the worker \( q = 0 \) and dismissing him after low output. Now observe that when (3) holds strictly, it must be that an age 1 worker is retained after poor performance if he is not to exert effort when offered \( q = 0 \). Otherwise, the worker would have an incentive to exert effort. As the equilibrium in Appendix B shows, it is optimal for the firm to retain an age 1 worker after low output precisely when such a worker does not exert effort when offered \( q = 0 \). Thus, when (3) holds strictly, the belief that an age 1 worker does not exert effort if he is retained after low output can be self-fulfilling.

A consequence of the above discussion is that if effort is strictly optimal for an age 1 worker who is not retained after low output, then the firm would actually benefit from committing to dismiss age 1 workers after poor performance. Indeed, this commitment eliminates the multiplicity just discussed 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 on both current and future output. We show that in this case, there exists scope for the offer of probation to age 1 workers when effort affects mostly future output.\(^\text{17}\) Recall that \( \eta(e, e_-) \) is the increase in the probability

\(^{17}\)A few papers have considered the problem of repeated moral hazard with effort persistence. Jarque (2010) shows that this problem is observationally equivalent to a problem without persistence if the agent’s utility is linear in effort and the distribution of outcomes in a period is a function of the discounted sum of efforts. 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.
that a high-ability worker produces high output when his current choice of effort is \( e \) and his previous choice of was \( e_- \). For simplicity, we assume that \( \eta(\bar{e}, \underline{e}) = 0 \) and \( \eta(\bar{e}, \bar{e}) = \eta(\bar{e}, \bar{e}) \). So, only effort in the previous period affects the probability of high output. Our analysis makes it clear that the results we obtain also hold when \( \eta(\bar{e}, e_-) - \eta(\bar{e}, e_-) \) is positive but small. Furthermore, we assume that \( e_- = \bar{e} \) for age 1 workers. We show at the end of Appendix C that this assumption is not crucial for our results.

Let \( \alpha + \eta(\bar{e}, \bar{e}) = \gamma > \alpha \). When \( \eta(\bar{e}, \underline{e}) = 0 \) and \( \eta(\bar{e}, \bar{e}) = \eta(\underline{e}, \bar{e}) \), we have that

\[
\Pr\{y|\tau, e, \underline{e}\} = \begin{cases} 
\alpha & \text{if } \tau = H \\
0 & \text{if } \tau = L
\end{cases}
\]

and

\[
\Pr\{y|\tau, e, \bar{e}\} = \begin{cases} 
\gamma & \text{if } \tau = H \\
0 & \text{if } \tau = L
\end{cases}
\]

Notice that both an age 1 worker and an age \( k \geq 2 \) worker who has only produced low output have no incentive to exert effort if they are dismissed after low output. This fact remains true as long as \( \eta(\bar{e}, e_-) - \eta(\underline{e}, e_-) \) is small regardless of \( e_- \).

We divide the analysis of the non-IID case in two parts. In the first part of our analysis, Subsection 5.1, we identify a necessary and sufficient condition for an age 1 worker to exert effort if he is retained after low output. So, the firm can benefit from retaining an age 1 worker who performs poorly. We then derive conditions under which the firm retains such a worker only if it commits to do so. Thus, commitment to employment is potentially beneficial to the firm.

As discussed in the introduction, unlike in a standard moral hazard setting, effort plays an additional role in our environment: it makes a worker’s performance more informative about his ability. Thus, the potential gain from commitment to employment is not only due to the fact that by offering probation the firm can increase output, but also due to the fact that commitment to employment allows the firm to better identify a worker’s type.

In the second part of our analysis, Subsection 5.2, we determine conditions under which the firm benefits from offering commitment to employment. We also determine conditions
under which the resulting gain is greater than just the increase in output, so as to make explicit the informational role of effort when commitment to employment is valuable.

Before we start, notice, as in the IID case, that the firm always makes an offer to an incumbent it knows is of the high-ability type and always offers the lowest one-period wage possible. The proof of Lemma 5 and all other omitted proofs are in Appendix C.

**Lemma 5.** Suppose the firm has an incumbent it knows is of the high-ability type, and let \( w' \) be the smallest one-period wage the firm can offer him if it is committed to do so. In equilibrium, the firm offers \( \max\{\overline{w}, w'\} \) to this worker, who accepts this offer.

### 5.1 Commitment Is Necessary for Incentives

Let \( \overline{w}_2^N \) be such that

\[
v(\overline{w}_2^N) = \max \left\{ v(w) - \phi_1 \alpha \delta (1 - \delta)^{-1} (1 - \delta^{T-2})[v(\overline{w}) - v(w)], v(0) \right\},
\]

where \( \phi_1 = (1 - \alpha)\phi_0/(1 - \phi_0\alpha) \) is the reputation (and also the private belief) of an age 2 worker who produced low output when of age 1. We know from the discussion in Subsection 3.1 that such a worker accepts a one-period wage of \( \overline{w}_2^N \) regardless of his effort choice when of age 1. Consider then an age 1 worker who is employed by the firm and suppose that: (i) the firm offers him \((\overline{w}_2^N, 0)\) 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; and (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

\[
-(1 - \delta)c + \phi_0 \alpha \delta (1 - \delta^{T-1})v(\overline{w})
+ (1 - \phi_0 \alpha) \delta \left[ (1 - \delta)v(\overline{w}_2^N) + \phi_1 \gamma \delta (1 - \delta^{T-2})v(\overline{w}) + (1 - \phi_1 \gamma) \delta (1 - \delta^{T-2})v(w) \right]
\geq \phi_0 \alpha \delta (1 - \delta^{T-1})v(\overline{w})
+ (1 - \phi_0 \alpha) \delta \left[ (1 - \delta)v(\overline{w}_2^N) + \phi_1 \alpha \delta (1 - \delta^{T-2})v(\overline{w}) + (1 - \phi_1 \alpha) \delta (1 - \delta^{T-2})v(w) \right],
\]

21
which is equivalent to

\[ \phi_0(1 - \alpha)(\gamma - \alpha)\delta^2(1 - \delta^{\tau - 2})[v(\bar{w}) - v(w)] \geq (1 - \delta)c. \] (7)

By construction, (7) implies that an age 1 worker can be induced to exert effort if the firm promises—either implicitly or through commitment to employment—that it will not dismiss the worker after poor performance. A natural question to ask is whether it is possible for an age 1 worker to exert effort if (7) is not satisfied. The next result shows that workers never exert effort if (7) is violated.

**Lemma 6.** No worker exerts effort in equilibrium if (7) is not satisfied.

In light of Lemma 6, from now on we take condition (A2) as given:

(A2) \( \phi_0(1 - \alpha)(\gamma - \alpha)\delta^2(1 - \delta^{\tau - 2})[v(\bar{w}) - v(w)] > (1 - \delta)c. \)

Without this assumption, except for the knife-edged case in which (7) holds with equality, there exists no scope for commitment to employment in the non-IID case. The same argument as in the proof of Proposition 1 shows that in equilibrium the firm always dismisses an age 1 worker after poor performance when (7) does not hold.

In order to simplify the exposition, we also take the following condition as given:

(A3) \( v(w) + \phi_1 \alpha \delta(1 - \delta)^{-1}(1 - \delta^{\tau - 2})[v(w) - v(\bar{w})] \leq v(0). \)

Condition (A3), which can be satisfied only if \( v \) is bounded below, implies that \( \bar{w}_2^N \) is zero. In other words, under (A3) 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. It is easy to see that (A3) also implies that an age 1 worker accepts any offer by the firm. We discuss the non-IID case without assumption (A3) in Appendix D.

In the remainder of this subsection, we show that there exist situations in which the firm retains an age 1 worker who fails to produce high output only if it is committed to do so. In such cases, the firm must offer probation to an age 1 worker if the worker is to exert effort.

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 an age \( k + 1 \) worker who has only produced low output so far produces high output. The first result we establish is that if \( \phi_2 \gamma \leq \phi_0 \alpha \), then the firm dismisses a worker of age \( k \geq 3 \) who has only produced low output unless it is committed to employ him. 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-ability type.

**Lemma 7.** Suppose that
\[
\phi_2 \gamma \leq \phi_0 \alpha.
\]  
(8)

The firm retains a worker of age \( k \geq 3 \) who has not revealed himself to be of high ability only if it is committed to do so.

We now prove the main result of this section: namely, that under conditions, the firm can induce an age 1 worker to exert effort only if it offers him probation. In order to do so, let \( \Delta = [y(1, \alpha) - \bar{w}] - y(\phi_0, \alpha) = \alpha(1 - \phi_0)\Delta_y - \bar{w} \) be the expected (flow) payoff gain from employing a high-ability worker instead of an age 1 worker.

**Proposition 2.** Suppose that \( \phi_2 \gamma \leq \phi_0 \alpha \) and that
\[
\phi_1 \gamma < \phi_0 \alpha \left[ 1 + \min \left\{ \frac{(\gamma - \alpha)\Delta_y + \Delta + \phi_0 \alpha(\gamma - \alpha)\Delta_y(T - 2)}{(1 + \phi_0 \alpha)\Delta_y + \{\Delta + \phi_0 \alpha[1 - (\gamma - \alpha)]\Delta_y\}(T - 2)} \alpha(1 - \phi_0) \right\} \right].
\]  
(9)

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 do so.

Note that \( \phi_1 \gamma > \phi_0 \alpha \) is compatible with (9). When \( \phi_1 \gamma > \phi_0 \alpha \), the firm’s flow payoff when it employs an age 2 worker who exerted effort when of age 1 but produced low output is greater than the firm’s flow payoff when it employs an age 1 worker. Hence, Proposition 2 can hold only if the firm is patient enough.

The restriction in Proposition 2 that \( \phi_1 \gamma \) cannot be much greater than \( \phi_0 \alpha \) is intuitive. Otherwise, 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 but produced low output is more attractive to the firm than an age 1 worker regardless of the firm’s discount factor. In this case, the firm
retains the age 2 worker even if it is not committed to do so. 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 sufficiently large.

5.2 Commitment Is Beneficial

Here we determine conditions under which the firm benefits from offering commitment to employment. We also determine conditions under which the gain from commitment is not just due to the increase in output when workers exert effort.

The first result we establish is a necessary and sufficient condition for commitment to employment to be beneficial. In order to do so, suppose first that the firm is constrained to make offers with $q = 0$ to age 1 workers and let $V_1$ be its payoff. Recall that $q$ denotes the number of periods the firm commits to employ the worker. Moreover, suppose that (8) and (9) hold and let $\delta > \delta^*$, so that commitment to employment is necessary to induce an age 1 worker to exert effort. By Proposition 2, when $q = 0$ the firm dismisses an age 1 worker who produces low output (and so the worker does not exert effort). Moreover, by (A3), an age 1 worker accepts a one-period wage of zero. Thus, the payoff $V_1$ satisfies

$$V_1 = (1 - \delta) y(\phi_0, \alpha) + \phi_0 \alpha \left\{ \delta (1 - \delta^{T-1}) [y(1, \alpha) - \bar{w}] + \delta^T V_1 \right\} + (1 - \phi_0 \alpha) \delta V_1.$$ 

Solving for $V_1$, we obtain that

$$V_1 = \lambda_1 y(\phi_0, \alpha) + (1 - \lambda_1) [y(1, \alpha) - \bar{w}],$$

where $\lambda_1 = (1 - \delta) \left\{ 1 - \delta + \phi_0 \alpha \delta (1 - \delta^{T-1}) \right\}^{-1}$.

Now suppose the firm is not constrained in the offers it can make to workers. An option for the firm is to offer $(w^0, (1, w^1))$ to an age 1 worker. By Proposition 2, the worker knows he is dismissed when of age 3 if he has not revealed himself to be of high ability by then. Thus, the worker does not exert effort when of age 2 if he produced low output when of age 1. As a result, (A2) and (A3) imply that an age 1 worker exerts effort if the firm makes him an offer with $q = 1$. Let $V_2$ be the firm’s payoff when it always offers $(w^0, (1, w^1))$ with
when it always makes an offer with \( q = 1 \) to age 1 workers. It is easy to see that \( V_2 \) satisfies

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

Solving for \( V_2 \), we obtain that

\[
V_2 = \lambda_2 y(\phi_0, \alpha) + (1 - \lambda_2) [y(1, \alpha) - \bar{w}] - \lambda_2 \delta \{ \phi_0 \alpha [y(1, \alpha) - y(1, \gamma)] + (1 - \phi_0 \alpha) \Delta_1 \}, \tag{11}
\]

where \( \Delta_1 = [y(1, \alpha) - \bar{w}] - y(\phi_1, \gamma) \) and \( \lambda_2 = (1 - \delta) \left\{ 1 - \delta^2 + \phi_0 [\alpha + \gamma (1 - \alpha)] \delta^2 (1 - \delta T^{-2}) \right\}^{-1} \).

For the algebra leading to \( V_2 \), see the derivation of (29) in the proof of Proposition 2 in Appendix C.

The firm benefits from commitment to employment when this allows it to obtain a payoff greater than \( V_1 \). Hence, a sufficient condition for commitment to employment to be beneficial is that \( V_2 > V_1 \). As it turns out, \( V_2 > V_1 \) is also necessary for commitment to employment to be beneficial. To summarize, we have the following result.

\textbf{Lemma 8.} \textit{Suppose that (8) and (9) are satisfied. If} \( \delta > \delta^* \), \textit{then commitment to employment is beneficial if, and only if,} \( V_2 > V_1 \).

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, he increases his output and makes his performance 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 \( \rho \) be the lifetime value to the firm of the extra output it obtains when age 1 workers exert effort. Notice that \( \rho \) satisfies

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

Indeed, when an age 1 worker exerts effort, which is possible only if he is retained after poor performance, the (expected) increase in output in the following period is either \( y(1, \gamma) - \bar{w} \),
\( y(1, \alpha) \) or \( y(\phi_1, \gamma) - y(\phi_1, \alpha) \), depending on whether the worker produces high or low output. Solving for \( \rho \), we obtain that

\[
\rho = \lambda_2 \delta \{ \phi_0 \alpha [y(1, \gamma) - y(1, \alpha)] + (1 - \phi_0 \alpha) [y(\phi_1, \gamma) - y(\phi_1, \alpha)] \}.
\]

Since \( \Delta_1 = \Delta + [y(\phi_0, \alpha) - y(\phi_1, \gamma)] \), equations (10) and (11) imply that

\[
V_2 - V_1 = \{ \lambda_1 - \lambda_2 [1 + \delta (1 - \phi_0 \alpha)(1 + \psi)] \} \Delta + \rho,
\]

where \( \psi = [y(\phi_0, \alpha) - y(\phi_1, \alpha)]/\Delta \). Recall that \( \Delta = [y(1, \alpha) - \bar{y}] - y(\phi_0, \alpha) \) is the (non-normalized) payoff gain to the firm from employing a worker of high ability. Equation (12) thus decomposes the gain from commitment to employment in two parts: the output gain \( \rho \) and the informational gain net of the commitment cost \( \{ \lambda_1 - \lambda_2 [1 + \delta (1 - \phi_0 \alpha)(1 + \psi)] \} \Delta \).

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 \( V_2 - V_1 \) is greater than \( \rho \), so that the benefit of probation is greater than the extra output it generates. When \( V_2 - V_1 > \rho \), we have that even if the gain in output due to the use of probation were lost, commitment to employment would still be beneficial to the firm. We say that commitment to employment is information beneficial when \( V_2 - V_1 > \rho \). Equation (12) implies that \( V_2 - V_1 > \rho \) if, and only if, \( \lambda_1 > \lambda_2 [1 + \delta (1 - \phi_0 \alpha)(1 + \psi)] \). Straightforward algebra shows that this last condition is equivalent to

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

A necessary condition for (13) is that \( \phi_1 \gamma > \phi_0 \alpha (1 - \delta^{T-1})/(1 - \delta^{T-2}) \). 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. The term \( (1 - \delta^{T-1})/(1 - \delta^{T-2}) \) reflects the fact that even if the performance of an age 2 worker is more informative about his ability than the performance of an age 1 worker, the age 2 worker lives for one less period. Notice that when the firm is
patient, which is the case of interest, condition (13) reduces to
\[ \phi_1 \gamma > \phi_0 \alpha (1 + \psi) \frac{T - 1}{T - 2}. \] (14)

We can now state the main result of this section. It follows immediately from Lemma 8 and the discussion that follows it.

**Proposition 3.** Suppose that \( \phi_2 \gamma \leq \phi_0 \alpha \) and that
\[ \phi_0 \alpha (1 + \psi) \frac{T - 1}{T - 2} < \phi_1 \gamma \]
\[ < \phi_0 \alpha \left[ 1 + \min \left\{ \frac{(\gamma - \alpha) \Delta_y + \Delta + \phi_0 \alpha (\gamma - \alpha) \Delta_y (T - 2)}{(1 + \phi_0 \alpha) \Delta_y + \{\Delta + \phi_0 \alpha [1 - (\gamma - \alpha)] \Delta_y\} (T - 2)}, \alpha (1 - \phi_0) \right\} \right]. \] (15)

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 is information beneficial.

Condition (15), which combines (9) and (14), has a clear interpretation. On the one hand, \( \phi_1 \gamma \) needs to be high enough for the informational gain from commitment to employment to be positive. On the other hand, \( \phi_1 \gamma \) cannot be too high, otherwise probation is not necessary to induce age 1 workers to exert effort. Notice the tension between the conditions \( \phi_2 \gamma \leq \phi_0 \alpha \) and (15). On the one hand, \( \phi_2 \gamma \) must be low enough for the firm to prefer an age 1 worker over an age \( k \geq 3 \) worker who has only produced low output. On the other hand, we need \( \phi_1 \gamma \) to be greater than \( \phi_0 \alpha \), which implies that \( \phi_2 \gamma \) cannot be too low. We analyze the conditions \( \phi_2 \gamma \leq \phi_0 \alpha \), (15), and (A1) to (A3) in more detail in Appendix C.\(^{18}\)

To summarize our results in this section, when effort has a delayed impact on output and high output is a strong signal of high ability, an age 1 worker can benefit from exerting effort only if he is retained after low output. Thus, if the firm cannot credibly promise to retain an age 1 worker after low output, then the use of probationary appointments can provide such workers with the appropriate incentive to exert effort. Moreover, since effort makes performance more informative about ability, the firm can benefit from offering commitment to employment purely from an informational point of view.

\(^{18}\)It can be shown that \( \phi_2 \gamma \leq \phi_0 \alpha \), (15), and (A1) to (A3) can be jointly satisfied for a broad range of parameter values.
Lastly, notice that in equilibrium there is no scope for more than two periods of probation. This result depends in part on the assumption that there is a one-period delay in the impact of effort on output. It also depends on the condition that $\phi \gamma \leq \phi_0 \alpha$, which constrains the gain to the firm from employing a worker of age 3 or more. In principle, our analysis can allow for more than two periods of probation as an equilibrium outcome if, for instance, it takes more than one period for effort to have an impact on output.

6 Discussion

In this section, we discuss some of our modeling assumptions. We also discuss the extent to which our results change if incentive pay is possible.

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 such a labor market setting, a worker obtains a higher wage once he reveals himself to be of high ability. However, unlike standard models of career concerns, we consider a market in which either: (i) firms possess enough monopsony 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. With perfect competition and no accumulation of firm-specific human capital for high-ability workers, a firm would not strictly benefit from identifying workers of high ability, in which case commitment to employment would not be valuable.

Nondecreasing Outside Option. As discussed in Section 2, the assumption that a worker’s outside option cannot decrease below $w$ captures situations in which a worker can seek employment in an alternative labor market where his ability has no value. The model can accommodate the case in which $w_R(\phi)$, the worker’s outside option, falls below $w$ when $\phi < \phi_0$ as long as this decrease is not too large. Indeed, by continuity, 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 simply complicate the analysis.

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 hiring him back. Our assumption of no recall rules out the possibility for the firm to do so. This assumption is consistent with the fact that the firm may not be willing or able to rehire a worker. The assumption of no recall is important in the non-IID case, as 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. Considering a more general environment with recall would provide no additional insights.

Good Performance Reveals High Ability. We assume that only high-ability type workers can produce high output. Here, we argue that our results still hold if the probability that a low-ability worker produces high output is positive but small. For simplicity, we assume that this probability, which we denote by $\alpha_L$, does not depend on a low-ability worker’s current or (in the non-IID case) previous choice of effort. We can adapt our argument to the case in which a low-ability worker’s effort has an impact on the probability that he produces high output.

Observe first that there exists $\phi^* \in (0, 1)$, which depends on $T$, such that Lemma 2 applies to any worker with reputation in the interval $[\phi^*, 1]$. Moreover, by increasing $\phi^*$ if necessary, we can take $\phi^*$ to be such that $y(\phi^*, \alpha) - \bar{w} > y(\phi_0, \alpha + \eta)$, in which case a straightforward modification of the proofs of Lemmas 4 and 5 shows that it is optimal for the firm to retain an incumbent with reputation in the interval $[\phi^*, 1]$. Now observe that if $\alpha_L$ is sufficiently small, then producing high output 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 $[\phi^*, 1]$.

For instance, this will be the case if the firm currently employs a younger worker who has revealed himself to be of high ability.
[\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-ability type. Moreover, by reducing \( \alpha'_L \) 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.\(^{20}\)

**Incentive Pay.** We assume that wage payments 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 provided to workers by career concerns. This allows us to analyze the trade-offs involved in the use of probation in a more transparent way. Moreover, as pointed out in the introduction, 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 common, 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. If (3) holds, and so career concerns are strong enough to induce an age 1 worker who is dismissed after low output to exert effort, then 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-ability type to exert effort. 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 to the firm in the IID case.

Consider now the non-IID case. We know that when performance pay is not possible, the

\(^{20}\) Notice that as in the environment of Section 2, effort increases the likelihood that a high-ability worker produces high output. Hence, effort still makes performance more informative about ability.
firm can induce an age 1 worker to exert effort only by retaining him after poor performance. Now, 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 the explicit promise of retention.

7 Conclusion

In this paper, we provide a rationale for the use of short-term commitment to employment in markets in which a worker’s ability is uncertain. We prove that a firm can benefit from committing to employ workers of unknown ability if this commitment encourages them 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. The reason is that 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. This inability to credibly promise retention undermines a worker’s incentives for effort. Hence, in this case it can be ex ante optimal for a firm to commit to ex post inefficient outcomes. Finally, 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.
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] Carmichael, H.L. (1988): “Incentives in Academics: Why Is There Tenure?,” *Journal of Political Economy*, 96(3), 453-472.

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

[7] Eeckhout, J., and X. Weng (2010): “Assortative Learning,” mimeo.

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

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

[10] 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.

[11] Guasch, J. L., and A. Weiss (1981): “Self-Selection in the Labor Market,” *American Economic Review*, 71(3), 275-284.
[12] Guasch, J. L., and A. Weiss (1982): “An Equilibrium Analysis of Wage-Productivity Gaps,” *Review of Economic Studies*, 49(4), 485-497.

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

[14] Holmström, B. (1982): “Managerial Incentives: A Dynamic Perspective.” In *Essays in Economics and Management in Honor of Lars Wahlbeck*, Helsinki: Swedish School of Economics. Reprinted in *Review of Economic Studies*, 66(1), 169-182, 1999.

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

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

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

[18] Kahn, C., and G. Huberman (1988): “Two-Sided Uncertainty and ‘Up-or-Out’ Contracts,” *Journal of Labor Economics*, 6(4), 423-444.

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

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

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

[22] Moscarini, G. (2005): “Job Matching and the Wage Distribution,” *Econometrica*, 73(2), 481-516.
[23] Mukherjee, A. (2008): “Sustaining Implicit Contracts When Agents Have Career Concerns: The Role of Information Disclosure,” *RAND Journal of Economics*, 39(2), 469-490.

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

[25] 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.

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

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

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

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

[30] Wang, R., and A. Weiss (1998): “Probation, Layoffs, and Wage-Tenure Profiles: A Sorting Explanation,” *Labour Economics*, 5(3), 359-383.
Appendix A: Proofs Preliminary Results

PROOF OF LEMMA 1

Suppose not, and let \( \tau \) be a period in which the firm does not employ a worker. The firm’s lifetime payoff in period \( \tau \) is then \( (1 - \delta) \Pi + \delta V \), where \( V \) is the firm’s expected continuation payoff. There are two possibilities in period \( \tau \): (i) the firm is not committed to make an offer to its incumbent or it has no incumbent; or (ii) the firm is committed to make an offer to its incumbent. Suppose (i) holds and consider the following deviation for the firm: offer \((w, 0)\) to the age 1 worker in period \( \tau \) and behave from period \( \tau + 1 \) on as if it did not deviate in \( \tau \). Since an age 1 worker always accepts a one-period wage of \( w \), the firm’s payoff from this deviation is at least \( (1 - \delta) (y - w) + \delta V \), which is greater than \( (1 - \delta) \Pi + \delta V \) by assumption. Thus, the firm has a profitable deviation, a contradiction. Consider now case (ii). There are two alternatives. Either the incumbent is known to be of the high-ability type or not. Since, by (A1), \( y(1, \alpha) - w > y(\phi_0, \alpha + \eta) > y - w \), and a worker known to be of high ability accepts any wage greater than \( w \), an argument similar to the one used in case (i) shows that the firm has a profitable deviation, a contradiction. 

PROOF OF LEMMA 2

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 \( \ell \). Note that if \( \ell = 0 \), then: (i) the firm never offers the incumbent a one-period wage greater than \( \max\{w, w'\} \); (ii) if the firm makes the incumbent an offer, then it never commits to future one-period wage offers greater than \( w \) (trivially satisfied); and (iii) the incumbent does not exert effort if employed. Now observe, by the argument in the proof of Lemma 1, that if the firm offers the incumbent a one-period wage of \( w \), then he accepts the offer (no matter \( k \) and \( \ell \)). In particular, the incumbent will never punish the firm for a deviation by rejecting a one-period wage offer of \( w \). 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. It is now easy to see that (i) and (ii) must also hold, for otherwise the firm can profitably deviate either by lowering its one-period wage offer to the incumbent or by lowering the future one-period wage offers that it promises to the incumbent.

PROOF OF LEMMA 3

Suppose there exist histories $h$ and $h'$ for the firm after which it makes an offer to the available age 1 worker with $V(h'|\sigma) > V(h|\sigma)$. Consider now the deviation for the firm where it behaves after $h$ as if $h'$ had happened. Since workers follow symmetric strategies, this deviation increases the firm’s payoff after $h$ by $V(h'|\sigma) - V(h|\sigma)$, a contradiction.

Appendix B: Proofs IID Case

PROOF OF LEMMA 4

Suppose $\sigma$ is an equilibrium. By Lemma 2, an incumbent known to be of the high-ability type never exerts effort. Moreover, such a worker rejects any offer with a one-period wage smaller than $\bar{w}$. Hence, by (A1), $V(h|\sigma) < y(1, \alpha) - \bar{w}$ if $h$ is the initial history of the game. Lemma 3 then implies that $V(h'|\sigma) < y(1, \alpha) - \bar{w}$ for every history $h'$ for the firm after which it hires the available age 1 worker. The desired result now follows from the fact that a worker known to be of the high-ability type always accepts an offer of $\bar{w}$.

PROOF OF PROPOSITION 1

We divide the proof of Proposition 1 into two parts. First, we consider the case in which (3) holds, so that it is incentive compatible for an age 1 worker to exert effort if he is not retained after low output. Then, we consider the case in which (3) is violated.

CASE 1. Suppose that (3) holds. We divide the proof of 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 $\sigma^*$ in which the firm dismisses age 1 workers after poor performance and these workers exert effort. Finally, 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 smaller than its payoff in $\sigma^*$.

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

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

By construction, $\bar{w}_1$ is the smallest one-period wage an age 1 worker accepts if he exerts effort and is not retained after poor performance. Recall from the main text that Lemma 2 implies that if the firm does not offer probation to an age 1 worker and he produces high output, then his wage in every subsequent period is $\bar{w}$. We have the following result.

**Claim 1.** 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-ability 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 $\bar{w}_1$ as long as he does not produce high output and a wage of $\bar{w}$ once he reveals himself to be of high ability.

**Proof:** We start with some notation. Let $w_k$, $e_k$, and $\phi_k$ 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 $\phi_k$ is strictly decreasing in $k$. Now define $\tilde{\zeta}(e_k)$ and $\tilde{c}(e_k)$ to be such that: (i) $\tilde{\zeta}(e_k) = \alpha$ and $\tilde{c}(e_k) = 0$ if $e_k = \bar{e}$; and (ii) $\tilde{\zeta}(e_k) = \alpha + \eta$ and $\tilde{c}(e_k) = c$ if $e_k = \bar{e}$.

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

$$v(w_k + \kappa) - v(w_k) = [1 - \phi_k \tilde{\zeta}(e_k)]\delta\{v(\bar{w} + \varepsilon) - v(\bar{w})\},$$

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

$$v(\bar{w} + \kappa) \leq v(\bar{w}) + v(w_k + \kappa) - v(w_k)$$

$$\leq [1 - \delta + \delta \phi_k \tilde{\zeta}(e_k)]v(\bar{w}) + [1 - \phi_k \tilde{\zeta}(e_k)]\delta v(\bar{w} + \varepsilon) \leq v(\bar{w} + (1 - \phi_k \tilde{\zeta}(e_k))\delta \varepsilon).$$
where the last inequality also follows from the concavity of \( v \). Hence, \( \kappa \leq [1 - \phi_k \xi(e_k)]\delta \varepsilon \), as \( \varepsilon \) 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 
\[
\delta^k \varepsilon - [1 - \phi_k \xi(e_k)]\delta^k \varepsilon = \phi_k \xi(e_k) \delta^k \varepsilon \text{ while still satisfying his participation constraints (the worker has the option of not changing his behavior after the change in wages). The desired result follows by an induction argument.}^{21}
\]

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 \( s \in \{1, \ldots, T - k\} \). Then let 
\[
\varepsilon = w_{k,k+s} - \overline{w} \text{ and define } \kappa \text{ to be such that}
\]
\[
v(w_k + \kappa) - v(w_k) = \phi_k \xi(e_k) \delta^s[v(\overline{w} + \varepsilon) - v(\overline{w})].
\]

Given that \( w_k \leq \underline{w} < \overline{w} \), the same argument as in the previous paragraph shows that 
\[
\kappa \leq \phi_k \xi(e_k) \delta^s \varepsilon.
\]

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
\[
v(w_q) + \phi_q \xi(e_q)\delta(1 - \delta)^{-1}(1 - \delta^{T-q+1})[v(\overline{w}) - v(\underline{w})] - c(e_q) \geq v(\underline{w}). \tag{16}
\]

Since \( \phi_q < \phi_0 \) and (3) implies that
\[
\phi_0(\alpha + \eta)\delta(1 - \delta)^{-1}(1 - \delta^T)[v(\overline{w}) - v(\underline{w})] - c \geq
\phi_0 \xi(e_q)\delta(1 - \delta)^{-1}(1 - \delta^T)[v(\overline{w}) - v(\underline{w})] - c(e_q),
\]
\[
21\text{Note that if the firm and the workers have different discount factors, then decreasing } w_{k+1} \text{ to } \underline{w} \text{ and increasing } w_k \text{ to } w_k + \kappa \text{ changes the present discount wage bill from employing the worker by } \delta_f \varepsilon - \delta_w \kappa, \text{ where } \delta_f \text{ is the firm’s discount factor and } \delta_w \text{ is the worker’s discount factor. Thus, the conclusion that } w_k \leq \underline{w} \text{ for all } k \in \{1, \ldots, q\} \text{ is necessary to minimize the present discounted wage bill from employing the worker holds as long as } \delta_f \geq \delta_w. \text{ More generally, Claim 1 holds as long as } \delta_f \geq \delta_w.\]
we then have that $w_q > \tilde{w}_1$. Now note that

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

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 $\kappa$ be such that

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

Given that $w_{q-1} \leq w_q$, the same argument as in the first paragraph of the proof shows that $\kappa \leq [1 - \phi_{q-1} \xi(e_{q-1})] \delta \varepsilon$. Hence, if $w_q$ is such that (16) 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 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

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

In particular, $w_k \geq \tilde{w}_1$ 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. Notice that the firm’s payoff in any such equilibrium is the same. Let $\sigma^*$ 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 make him an offer; (II) the firm offers $(\tilde{w}_1, 0)$ 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 employ him; (III) an incumbent who has revealed himself to be of the high-ability type does not exert effort; (IV) an age 1 worker who accepts an offer of $(w, 0)$ exerts effort; and (V) a worker accepts an offer if indifferent between taking it and collecting his outside option. Notice that
we have not specified the off-the-path of play behavior of the firm and the workers. We do this later.

Observe that it is incentive compatible for an age 1 worker to exert effort if the firm offers him \((w,0)\). In what follows, we prove that: \((A)\) regardless of his behavior, it is optimal for the firm to dismiss an incumbent who has never produced high output if it is not committed to employ him; and \((B)\) if the firm makes an offer to the available age 1 worker, then it is optimal for the firm to offer \((\tilde{w}_1,0)\). The following facts will be useful. First, if \(V^*\) is the firm’s payoff when play is given by \(\sigma^*\), then \(V^*\) satisfies

\[
V^* = (1 - \delta)[y(\phi_0, \xi) - \tilde{w}_1] + \phi_0 \xi \{\delta(1 - \delta^{-1})[y(1, \alpha) - \bar{w}] + \delta^TV^*\} + (1 - \phi_0 \xi)\delta V^* \\
= (1 - \delta)[y(\phi_0, \xi) - \tilde{w}_1] + \delta V^* + \phi_0 \xi [1 - \delta^{-1}][y(1, \alpha) - \bar{w} - V^*],
\]

where \(\xi = \alpha + \eta\). Second, \(y(\phi_0, \xi) - \tilde{w}_1 < V^* < y(1, \alpha) - \bar{w}\) by (A1). Third, the reputation of an incumbent who has only produced low output is smaller than \(\phi_0\).

We begin with result \((A)\). 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 \(\tilde{w}_1\) in order for the firm’s offer to be accepted; we know from the proof of Lemma 1 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

\[
V_0 = (1 - \delta)[y(\phi_0, \xi) - \tilde{w}_1] + \phi_0 \xi \{\delta(1 - \delta^{-k})[y(1, \alpha) - \bar{w}] + \delta^{k+1}V^*\} + (1 - \phi_0 \xi)\delta V^*,
\]

which is smaller than \(V^*\) since \(V^* < y(1, \alpha) - \bar{w}\) 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 1 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 $\tilde{w}_1$ as long as he does not produce high output and a wage of $w$ 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

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

Now observe that

$$(1 - \delta^{q+1})[y - \tilde{w}_1] + \delta^{q+1}V^* > (1 - \delta^{q+1})[y - \tilde{w}_1] + \delta^{q+1} \phi_0 \xi \{ (1 - \delta)[\overline{y} - \tilde{w}_1] + \delta^{q+1} \phi_0 \xi \{ (1 - \delta)[y - \tilde{w}_1] + \delta V^* \} \}
+ \delta(1 - \delta^{T-k-q-1})[y(1, \alpha) - \overline{w}] + \delta^{T-k-q}V^* \} + \delta^{q+1}(1 - \phi_0 \xi \{ (1 - \delta)[y - \tilde{w}_1] + \delta V^* \} \}
= \phi_0 \xi \{ (1 - \delta^{q+1})[y - \tilde{w}_1] + \delta^{q+1}(1 - \delta)[\overline{y} - \tilde{w}_1] + \delta^{q+2}(1 - \delta^{T-k-q-1})[y(1, \alpha) - \overline{w}] + \delta^{T-k+1}V^* \} + (1 - \phi_0 \xi \{ (1 - \delta^{q+2})[y - \tilde{w}_1] + \delta^{q+2}V^* \} ,
$$

where the inequality follows from (17) 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

$$[\phi_0(1 - \xi)^{q+1} + 1 - \phi_0] \{ (1 - \delta^{q+1})[y - \tilde{w}_1] + \delta^{q+1}V^* \} >
(1 - \phi_0)[1 - (1 - \xi)^{q+1}] \{ (1 - \delta^{q+1})[y - \tilde{w}_1] + \delta^{q+1}V^* \} + (1 - \xi)^{q+1} \phi_0 \xi \{ (1 - \delta^{q+1})[y - \tilde{w}_1] + \delta^{q+1}V^* \}
+ \delta^{q+1}(1 - \delta)[\overline{y} - \tilde{w}_1] + \delta^{q+2}(1 - \delta^{T-k-q-1})[y(1, \alpha) - \overline{w}] + \delta^{T-k+1}V^* \} + (1 - \xi)^{q+1}(1 - \phi_0 \xi \{ (1 - \delta^{q+2})[y - \tilde{w}_1] + \delta^{q+2}V^* \} .
$$

Given that $(1 - \xi)^{q+1}(1 - \phi_0 \xi) + (1 - \phi_0)[1 - (1 - \xi)^{q+1}] = 1 - \phi_0 + \phi_0(1 - \xi)^{q+2}$ and $y - \tilde{w}_1 < V^*$, we can then conclude that

$$[\phi_0(1 - \xi)^{q+1} + 1 - \phi_0] \{ (1 - \delta^{q+1})[y - \tilde{w}_1] + \delta^{q+1}V^* \} >
[1 - \phi_0 + \phi_0(1 - \xi)^{q+2}] \{ (1 - \delta^{q+2})[y - \tilde{w}_1] + \delta^{q+2}V^* \} + (1 - \xi)^{q+1} \phi_0 \xi \{ (1 - \delta^{q+1})[y - \tilde{w}_1] + \delta^{q+1}V^* \}
+ \delta^{q+1}(1 - \delta)[\overline{y} - \tilde{w}_1] + \delta^{q+2}(1 - \delta^{T-k-q-1})[y(1, \alpha) - \overline{w}] + \delta^{T-k+1}V^* \} .
$$

Thus, $V_q > V_{q+1}$, and so (A) is true.
We now establish result (B). The above argument together with Claim 1 shows that if the firm makes an offer with \( q \geq 1 \) to an age 1 worker, then its payoff is bounded above by
\[
y(\phi_0, \xi) - \tilde{w}_1 + \phi_0 \xi \left\{ \delta(1 - \delta^{T-1})[y(1, \alpha) - \overline{w}] + \delta^T V^* \right\} + (1 - \phi_0 \xi) V_0,
\]
which is smaller than \( V^* \). This implies the desired result since the firm’s payoff in case it offers \((\tilde{w}_1, 0)\) to an age 1 worker is \( V^* \) and \( \tilde{w}_1 \) 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 specify the off-the-equilibrium-path behavior of the firm and the workers. We know from Lemma 2 that the firm must offer \((\min\{\overline{w}, w'\}, 0)\) to an incumbent it knows is of high ability if it is committed to offer him a one-period wage of at least \( w' \) (and these workers do not exert effort). Thus, we are left with the task of determining: (i) the one-period wage offers the firm makes when it is committed to employ 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; and (iii) the effort choice of an age 1 worker who receives an offer with \( q \geq 1 \). Recall that the behavior given by (II) is optimal regardless of how we specify (i) and (ii).

Consider a worker of age \( k \geq 2 \) who has not revealed himself to be of the high-ability type, and let \( \phi \) and \( \pi \) be, respectively, his reputation and private belief that he is of high ability. We proceed by induction in \( \ell \), 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, \((1 - \delta)^{-1}(1 - \delta^{T-k})[v(\overline{w}) - v(\overline{\pi})] \geq (1 - c)\). If committed to employ him, the firm offers the worker a one-period wage of \( \max\{\tilde{w}, w'\} \), where \( w' \) is the smallest one-period wage the firm can offer and \( \tilde{w} \) 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 make an offer to the worker when he is of age \( k + 1 \). The worker’s choice of effort is optimal given \( \pi \) and his and the firm’s behavior after he produces low output (in which case \( \ell = 0 \)). As when \( \ell = 0 \), if committed to employ him, the firm offers the worker a one-period wage of \( \max\{\tilde{w}, w'\} \), where \( w' \) is the smallest one-period wage the firm can offer and \( \tilde{w} \) 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 \( \lambda_1 \) to \( \lambda_q \), with \( \lambda_1 < 1 \), such that with probability \( \lambda_k \) 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 that it hires exert effort as long as they do not reveal themselves to be of high ability, Lemma 3 and Claim 1 imply that \( V^{**} \leq V \), where \( V \) is such that \( V = \sum_{k=1}^q \lambda_k T_k V \) and

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

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 (17), that \( T_1 V^* = V^* \). Moreover, \( T_k V^* \) is strictly decreasing in \( k \) by the argument in the proof that \( \sigma^* \) is an equilibrium. 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 \( V \) by the contraction mapping theorem, we can then conclude that \( V < V^* \), which implies the desired result.

**CASE 2.** Consider now the case in which (3) does not hold, so that an age 1 worker who is not retained after poor performance has no incentive to exert effort. First, we establish that if not committed to do 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. Then, 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 do 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 proof of Lemma 1 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 employ 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-ability type when he is of age $k + j$, with $j \in \{0, \ldots, q\}$; and (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

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

(18)

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

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

and, by Lemma 2, that

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

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

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 $\phi$. 44
Consider then the following deviation for the firm: offer \((w^0, (q, \{w_s^q\}_{s=1}^q))\) 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 \(w^*\) 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 2, 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

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

(19)

where \(\pi \leq \phi_0\) is the worker’s belief that he is of high ability and \(R(\overline{y}, \overline{e} | \pi, k)\) is the worker’s continuation payoff if he chooses \(e\) and produces \(y\), which depends on \(\pi\) and \(k\). Note that \(\pi\) need not coincide with the worker’s reputation. Given that \(R(\overline{\gamma}, \overline{e} | \pi, k) = R(\overline{\gamma}, \overline{e} | \pi, k)\), we can rewrite (19) as

\[\pi\eta\delta[\overline{R}(\overline{\gamma}, \overline{e} | \pi, k) - \overline{R}(\overline{y}, \overline{e} | \pi, k)] + \delta(1 - \pi\alpha)[\overline{R}(\overline{\gamma}, \overline{e} | \pi, k) - \overline{R}(\overline{y}, \overline{e} | \pi, k)] \geq c.\]

(20)

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 \(\ell\).

1) Suppose that \(\ell = 1\), that is, the worker is dismissed after low output (and so \(\Delta_2 = 0\)). In this case, Lemma 2 implies that (20) reduces to

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

45
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 $\ell'$ periods. Let $\pi(e) = [1 - \xi(e)]\pi/\pi(0)$ 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 $\ell(e)$ 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 $k + \ell(e) - 1$ if he only produces low output. Since $\pi(e) > \pi(\tau)$, the induction hypothesis implies that $\ell(e) \leq \ell(e) = \ell$.

Following the same logic used in the derivation of (18), we have that

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

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 $j \in \{0, \ldots, \ell(e) - 2\}$; and (ii) $R(0, e)$ is the worker’s continuation payoff after low output if he never produces high output afterward, 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 2 then implies that

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

and

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

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

\[ R(y, e|\pi, k) \geq \pi(e) \sum_{j=0}^{\hat{l}(\pi) - 2} (1 - \alpha)^j [R(j) - R(\emptyset, \pi)] + R(\emptyset, \pi). \]

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

To finish the induction argument, observe that \( \Delta_1 \leq (1 - \delta)^{-1}(1 - \delta_{w}^{T-k})[v(\bar{w}) - v(\underline{w})] \). It is immediate to see that this bound on \( \Delta_1 \) holds if \( w_s \leq \bar{w} \) for all \( s \in \{0, \ldots, \ell' - 1\} \), given that \( R(y, e|\pi, k) \geq (1 - \delta)^{-1}(1 - \delta_{w}^{T-k})v(\bar{w}) \). Now recall that if \( w_s > \bar{w} \) for some \( s \in \{0, \ldots, \hat{l}(\pi) - 2\} \), then Lemma 2 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 \( w_s \geq \bar{w} \), which implies the desired result. Consequently, a necessary condition for \( (20) \) is that

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

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.

EQUILIBRIUM WITH AGE 1 WORKERS RETAINED AFTER LOW OUTPUT

Here we show that when \( (3) \) holds with strict inequality, we can have an equilibrium in the IID case in which the firm retains age 1 workers after low output. For this, let \( \phi_1(e) = [1 - \xi(e)]\phi_0/[1 - \phi_0\xi(e)] \), where \( \xi(e) = \alpha \) and \( \xi(\pi) = \alpha + \eta \), be the reputation of an age 2 worker who chooses \( e \) and produces low output when of age 1. Moreover, let \( \phi_2 = (1 - \alpha)^2\phi_0/(1 - \alpha)^2\phi_0 + 1 - \phi_0 \) be the highest reputation possible for a worker of age
$k \geq 3$ who has never produced high output. Suppose then that

\[ v(w) - \phi_1(\omega)\alpha(1-\delta)^{-1}(1-\delta^{T-2})[v(\bar{w}) - v(w)] < v(0) \]  
(21)
\[ \phi_1(\omega)\eta\delta(1-\delta)^{-1}(1-\delta^{T-2})[v(\bar{w}) - v(w)] > c, \]  
(22)
\[ \max\{\phi_2, \phi_1(\omega)\}\eta\delta(1-\delta)^{-1}(1-\delta^{T-2})[v(\bar{w}) - v(w)] < c, \]  
(23)
\[ \phi_1(\omega)(\alpha + \eta) > \phi_0\alpha(T-1)/(T-2), \]  
(24)
\[ \phi_0\eta\delta[v(\bar{w}) - v(0)] < c. \]  
(25)

Condition (21), which can be satisfied only 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 (22) 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 (23) 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 $\sigma^{**}$ where: (I) the firm offers $(\bar{w}, 0)$ to an incumbent it knows is of high ability if it is not committed to make 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 employ 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 make him an offer; (IV) the firm offers $(\max\{\bar{w}, w\}, 0)$ to an incumbent if it is committed to offer 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 $\sigma^{**}$ is an equilibrium if $\delta$ is sufficiently close to one.
Proof: 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).

Now observe, by Lemma 2 and (23), 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

\[
-(1-\delta)c + \phi_0(\alpha + \eta)\delta(1-\delta^{T-1})v(\bar{w}) + [1-\phi_0(\alpha + \eta)]\delta \left\{ (1-\delta)v(0) + \phi_1(\tau)\alpha \delta(1-\delta^{T-2})v(\bar{w}) + [1-\phi_1(\tau)\alpha]\delta(1-\delta^{T-2})v(\bar{w}) \right\}
\]

Straightforward algebra shows that (26) reduces to

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

which is not satisfied by (25). Now observe that the same argument as in the proof of Lemma 6 in the non-IID case shows that if (26) 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 (see the proof of Proposition 2 in Section 5 for a similar argument) shows that

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

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

\[
V' = (1-\delta)y(\phi_1(e), \alpha + \eta) + \phi_1(e)(\alpha + \eta)\delta(1-\delta^{T-2})[y(1, \alpha) - \bar{w} - V] + \delta V.
\]
Hence, $V' > V$ if, and only if,

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

which reduces to

$$
[1 + \phi_1(\epsilon)(\alpha + \eta)(T - 2)]\Delta > [1 + \phi_0 \alpha + \phi_0 \alpha(T - 2)]\Delta'
$$

when $\delta \approx 1$. Since $\Delta > \Delta'$, as $\phi_1(\epsilon)(\alpha + \eta) > \phi_0 \alpha$ by (24), the last inequality is satisfied if $\phi_1(\epsilon)(\alpha + \eta)(T - 2) > \phi_0 \alpha(T - 1)$, which is true by (24). □

Observe that $\sigma^{**}$ 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 employ 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 do so.

**Appendix C: Proofs Non-IID Case**

**PROOF OF LEMMA 5**

The present value of the output generated by a worker who has not revealed himself to be of the high-ability type is at most $\xi^* = y(\phi_0, \alpha) + \delta \phi_0[y(1, \gamma) - y(1, \alpha)]$. 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 $\xi^* < y(\phi_0, \gamma)$, and so $\xi^* < y(1, \alpha) - \bar{w}$ by (A1). The desired result now follows from the same argument as in the proof of Lemma 4. □

**PROOF OF LEMMA 6**

By Lemma 2, 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 and
assume that (7) is not satisfied. The worker’s incentive-compatibility constraint for effort exertion is

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

where \(\pi \leq \phi_0\) is the worker’s private belief that he is of high ability, \(\xi\) 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 \(\pi\) and \(k\) for ease of notation. Notice that, unlike the IID case, the probability \(\xi\) is independent of the worker’s current choice of effort. Since \(R(\overline{y}, \overline{e}) = R(\overline{y}, \overline{e})\), we can rewrite the last inequality as

\[\delta (1 - \pi \xi) [R(\overline{y}, \overline{e}) - R(y, e)] \geq c.\] (27)

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 \(R(\overline{y}, \overline{e}) = R(y, e) = (1 - \delta)^{-1} (1 - \delta^{T - k}) v(w)\) if \(\ell = 1\), in which case (27) 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,

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

where \(\xi'(\overline{e}) = \gamma\) and \(\xi'(\overline{e}) = \alpha\). The same argument as in the proof of Case 2 of Proposition 1 shows that \(R_{-}(\overline{e}) \leq R_{-}(\overline{e})\).\(^2\) This implies that a necessary condition for (27) is that

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

\(^2\)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 \(k + 2\).
To finish, notice, also by the same argument as in the proof of Case 2 of Proposition 1, that
\[ R_+ - R_-(e) \leq (1 - \delta)^{-1}(1 - \delta^{T-k-1})[v(\omega) - v(\mu)]. \]
Hence, (27) cannot be satisfied by assumption. Thus, by induction, the worker never exerts effort when employed if (7) is not satisfied. \[ \blacksquare \]

**PROOF OF LEMMA 7**

Since this is useful in Appendix D, we prove Lemma 7 without assuming (A3). For this, let \( \tilde{w}_1^N \) be such that
\[ v(\tilde{w}_1^N) = \max \left\{ v(\omega) - \phi_0 \alpha \delta (1 - \delta)^{-1}(1 - \delta^{T-1})[v(\mu) - v(\mu)], v(0) \right\}. \]
We know from Section 3 that an age 1 worker accepts a one-period wage of \( \tilde{w}_1^N \). Moreover, \( \tilde{w}_1^N \) is the smallest one-period wage an age 1 worker accepts if he is not retained after poor performance. Indeed, since an age 1 worker can be dismissed after low output only if the firm does not offer probation to him, Lemma 2 implies that the worker’s lifetime continuation payoff after high output is \( (1 - \delta)^{-1}(1 - \delta^{T-1})v(\mu) \). Notice that \( \tilde{w}_1^N = 0 \) if (A3) holds.

Suppose now, by contradiction, that the firm retains a worker of age \( k \geq 3 \) who has never produced high output despite not being committed to do 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 that 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 \( \tilde{w}_1^N \). Thus, an upper bound to the firm’s payoff (obtained if the worker exerted effort in the previous period) is
\[ \nabla_0^k = (1 - \delta)[y(\phi_{k-1}, \gamma) - \tilde{w}_1^N] + \phi_{k-1} \gamma \delta \left\{ (1 - \delta^{T-k})[y(1, \alpha) - \mu] + \delta^{T-k} V \right\} + (1 - \phi_{k-1} \gamma) \delta V \]
\[ = (1 - \delta)[y(\phi_{k-1}, \gamma) - \tilde{w}_1^N] + \delta V + \phi_{k-1} \gamma \delta (1 - \delta^{T-k})[y(1, \alpha) - \mu - V], \]
where \( V \) is the firm’s payoff after it hires an age 1 worker. Given that \( V < y(1, \alpha) - \mu \) by the proof of Lemma 5 and once more using the fact that \( \phi_{k-1} \gamma \leq \phi_0 \alpha \), we then have that
\[ \nabla_0^k < (1 - \delta)[y(\phi_0, \alpha) - \tilde{w}_1^N] + \delta V + \phi_0 \alpha \delta (1 - \delta^{T-1})[y(1, \alpha) - \mu - V]. \]
However, since an age 1 worker accepts \((\tilde{w}_1^N, 0)\), the payoff on the right-hand side of the above inequality is a lower bound on the firm’s payoff if it offers \((\tilde{w}_1^N, 0)\) to an age 1 worker and dismisses this worker 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 \(\ell\) 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 1 in the proof of Proposition 1 then shows that (1) a lower bound on the firm’s present discounted wage bill from employing the worker is if it pays him a wage of \(\tilde{w}_1^N\) as long as he does not produce high output and a wage of \(\bar{w}\) once he reveals himself to be of the high-ability type. Now observe that (2) 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 \(V^k_\ell\) be the firm’s payoff in the case in which the worker behaves as in (2) and wage payments are as in (1). It is easy to see that

\[
V^k_\ell = (1 - \delta)[y(\phi_{k-1, \gamma} - \tilde{w}_1^N) + \phi_{k-1} \gamma \delta \{(1 - \delta)[y(1, \gamma) - \bar{w}] + \delta(1 - \delta^{T-k-1})[y(1, \alpha) - \bar{w}] + \delta^{T-k} V\} + (1 - \phi_{k-1} \gamma) \delta V^k_{\ell-1}.\]

We claim that \(V^k_\ell < V\) for all \(k \geq 3\) and \(\ell \in \{1, \ldots, T - k\}\), so that the firm can profitably deviate by employing the age 1 worker.

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

\[
V^k_1 = (1 - \delta)[y(\phi_{k-1, \gamma} - \tilde{w}_1^N) + \phi_{k-1} \gamma \delta \{(1 - \delta)[y(1, \gamma) - \bar{w}] + \delta(1 - \delta^{T-k-1})[y(1, \alpha) - \bar{w}] + \delta^{T-k} V\} + (1 - \phi_{k-1} \gamma) \delta \{(1 - \delta)[y(\phi_k^*, \gamma) - \tilde{w}_1^N] + \phi_k^* \gamma \delta \{(1 - \delta^{T-k-1})[y(1, \alpha) - \bar{w}] + \delta^{T-k} V\} + (1 - \phi_k^* \gamma) \delta V\},
\]
where \( \phi_k^* = (1 - \gamma)\phi_{k-1}/(1 - \phi_{k-1}\gamma) \). It is immediate to see from the above equation that
\[
\nabla_1^k = (1 - \delta)[y(\phi_{k-1}, \gamma) - \tilde{w}_1^N] + \delta(1 - \delta)\{\phi_{k-1}\gamma[y(1, \gamma) - \bar{w}] + (1 - \phi_{k-1}\gamma)[y(\phi_k^*, \gamma) - \tilde{w}_1^N]\}
+ \phi_{k-1}[\gamma + (1 - \gamma)]\delta^2(1 - \delta^{T-k-1})[y(1, \alpha) - \bar{w} - V] + \delta^2V.
\]
Consider then the following deviation for the firm: offer \((\tilde{w}_1^N, 0)\) to the available age 1 worker and dismiss him after poor performance. Since
\[
V \geq V' = (1 - \delta)[y(\phi_0, \alpha) - \tilde{w}_1^N] + \delta V + \phi_0\alpha\delta(1 - \delta^{T-2})[y(1, \alpha) - \bar{w} - V],
\]
a lower bound on the firm’s payoff from this deviation is
\[
\tilde{V} = (1 - \delta)[y(\phi_0, \alpha) - \tilde{w}_1^N] + \phi_0\alpha\delta \{ (1 - \delta^{T-1})[y(1, \alpha) - \bar{w}] + \delta^{T-1}V \} + (1 - \phi_0\alpha)\delta V'
= (1 - \delta)[y(\phi_0, \alpha) - \tilde{w}_1^N] + \delta(1 - \delta)\{\phi_0\alpha[y(1, \alpha) - \bar{w}] + (1 - \phi_0\alpha)[y(\phi_0, \alpha) - \tilde{w}_1^N]\}
+ \phi_0\alpha + (1 - \phi_0\alpha)\phi_0\alpha \{ \delta^2(1 - \delta^{T-2})[y(1, \alpha) - \bar{w} - V] \} + \delta^2V.
\]
Now observe that
\[
\phi_0\alpha + (1 - \phi_0\alpha)\phi_0\alpha = \phi_0\alpha(2 - \phi_0\alpha) > \phi_{k-1}\gamma(2 - \gamma) = \phi_{k-1}[\gamma + (1 - \gamma)],
\]
\[
\phi_{k-1}\gamma[y(1, \gamma) - \bar{w}] + (1 - \phi_{k-1}\gamma)[y(\phi_k^*, \gamma) - \tilde{w}_1^N] = y - \tilde{w}_1^N + \phi_{k-1}\gamma[\Delta_y - \Delta_w],
\]
and
\[
\phi_0\alpha[y(1, \alpha) - \bar{w}] + (1 - \phi_0\alpha)[y(\phi_0, \alpha) - \tilde{w}_1^N] = y - \tilde{w}_1^N + \phi_0\alpha[(1 + \alpha - \phi_0\alpha)\Delta_y - \Delta_w],
\]
where \( \Delta_w = \bar{w} - \tilde{w}_1^N \leq \bar{w} \). Thus, given that \( \Delta_y > \bar{w} \) by (A1), we have that \( \nabla_1^k < \tilde{V} \) for all \( k \geq 3 \). Since \( \tilde{V} \leq V \), we can conclude that \( \nabla_1^k < V \) for all \( k \geq 3 \). To finish, notice that if \( \nabla_{\ell-1}^k < V \) for \( k \geq 3 \) and \( \ell \geq 2 \), then \( \nabla_{\ell}^k < V \) as well. This establishes the lemma. \( \blacksquare \)

**PROOF OF PROPOSITION 2**

We break the proof of Proposition 2 into two parts.

**Step 1.** The firm dismisses a worker of age 2 who did not exert effort and produced low output when of age 1 if it is not committed to employ him.
As in the proof of Lemma 7, we establish Step 1 without assuming \((A3)\). Suppose not and let \(V\) be the firm’s payoff after it hires an age 1 worker. The same argument used in the proof of Lemma 7 shows that if the worker does not exert effort (so that at best he exerts effort when of age 3), then the firm can profitably deviate by replacing him with an age 1 worker. Consider then the case in which the worker exerts effort, which is possible only if the firm employs him for \(\ell \geq 1\) more periods. Notice that if the worker has only produced low output before he is of age \(2 + j\), with \(j \in \{0, \ldots, \ell\}\), then an upper bound on the probability that he produces high output when of age \(2 + j\) is \(\max\{\phi_1 \alpha, \phi_2 \gamma\} \leq \phi_0 \alpha\). Thus, as in the proof of Lemma 7, a lower bound on the firm’s present discounted wage bill from employing the worker is if it pays him a wage of \(\tilde{w}_1^n\) as long as he does not produce high output and a wage of \(w\) once he reveals himself to be of the high-ability type. Since, by Lemma 7, the firm’s payoff when it employs a worker of age 4 who has not revealed himself to be of high ability is smaller than \(V\), straightforward algebra shows that an upper bound on the firm’s payoff is given by

\[
\nabla = (1 - \delta)[y(\phi_1, \alpha) - \tilde{w}_1^n] + \delta(1 - \delta)\{\phi_1 \alpha[y(1, \gamma) - w] + (1 - \phi_1 \alpha)[y(\phi_2, \gamma) - \tilde{w}_1^n]\} \\
+ (\phi_1 \alpha + (1 - \phi_1 \alpha)\phi_2 \gamma)\delta^2(1 - \delta^{T-3})[y(1, \alpha) - w - V] + \delta^2 V.
\]

Consider now the following deviation for the firm: offer \((\tilde{w}_1^n, 0)\) to the available age 1 worker and dismiss him after poor performance. We know from the proof of Lemma 7 that a lower bound on the firm’s payoff from this deviation is

\[
\tilde{V} = (1 - \delta)[y(\phi_0, \alpha) - \tilde{w}_1^n] + \delta(1 - \delta)\{\phi_0 \alpha[y(1, \gamma) - w] + (1 - \phi_0 \alpha)[y(\phi_0, \alpha) - \tilde{w}_1^n]\} \\
+ (\phi_0 \alpha + (1 - \phi_0 \alpha)\phi_0 \alpha)\{\delta^2(1 - \delta^{T-2})[y(1, \alpha) - w - V]\} + \delta^2 V.
\]

Given that \(\phi_0 \alpha > \phi_2 \gamma\), we know that \(\tilde{V} > \nabla\) if

\[
y(\phi_1, \alpha) + \delta \{\phi_1 \alpha[y(1, \gamma) - w] + (1 - \phi_1 \alpha)[y(\phi_2, \gamma) - \tilde{w}_1^n]\} \\
< y(\phi_0, \alpha) + \delta \{\phi_0 \alpha[y(1, \alpha) - w] + (1 - \phi_0 \alpha)[y(\phi_0, \alpha) - \tilde{w}_1^n]\}.
\]

(28)
Since \( \phi_1 \alpha \gamma + (1 - \phi_1 \alpha) \phi_2 \gamma = \phi_1 \gamma \) and (as \( \Delta_y > \overline{w} \geq \Delta_w = \overline{w} - \overline{w}_1^N \) by (A1))

\[
y(\phi_1, \alpha) - \delta \phi_1 \alpha \overline{w} - \delta(1 - \phi_1 \alpha) \overline{w}_1^N = y - \delta \overline{w}_1^N + \phi_1 \alpha [\Delta_y - \delta \Delta_w] < y - \delta \overline{w}_1^N + \phi_0 \alpha [\Delta_y - \delta \Delta_w] = y(\phi_0, \alpha) - \delta \phi_0 \alpha \overline{w} - \delta(1 - \phi_0 \alpha) \overline{w}_1^N,
\]

a sufficient condition for (28) is that \( \phi_1 \gamma < \phi_0 \alpha (1 + \alpha - \phi_0 \alpha) \), which is satisfied by (9). Thus, the firm has a profitable deviation, a contradiction.

**Step 2.** There exists \( \delta^* \in (0, 1) \) such that if \( \delta > \delta^* \), then the firm employs a worker of age 2 who exerted effort and produced low output when of age 1 only if committed to do so.

Suppose, by contradiction, that regardless of its discount factor, the firm retains an age 2 worker who exerted effort and produced low output when of age 1 despite not being committed to do so, and let \( \ell \) 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, under (A3), an age 2 worker accepts a one-period wage of zero, the firm’s payoff is bounded above by

\[
V' = (1 - \delta)y(\phi_1, \gamma) + \phi_1 \gamma \left\{ \delta(1 - \delta^{T-2})[y(1, \alpha) - \overline{w}] + \delta^{T-1}V \right\} + (1 - \phi_1 \gamma)\delta V
= \delta V + (1 - \delta) \frac{y(\phi_1, \gamma) + \phi_1 \gamma \delta(1 - \delta^{T-2})[y(1, \alpha) - \overline{w} - V]}{\overline{w}_1^N}.
\]

We know from the main text that an age 1 worker exerts effort if the firm offers him \((w^0, (1, w^1))\) with \( w^0 = w^1 = 0 \). Since an option for the firm is to always make such an offer to the age 1 workers, we have that \( V \geq V_2 \), where \( V_2 \) satisfies the following recursion:

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

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

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

56
Thus,
\[ V_2 = [y(1, \alpha) - \bar{w}] - \frac{(1-\delta)\Delta + \delta(1-\delta)\{\phi_0\alpha[y(1, \alpha) - y(1, \gamma)] + (1-\phi_0\alpha)\Delta_1\}}{1-\delta^2 + \phi_0[\alpha + \gamma(1-\alpha)]\delta^2(1-\delta^{T-2})}. \]  
(29)

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, a contradiction. In order to show this, observe that
\[ A = \lim_{\delta \rightarrow 1} V \geq \lim_{\delta \rightarrow 1} V_2 = [y(1, \alpha) - \bar{w}] - \frac{\Delta + \phi_0\alpha[y(1, \alpha) - y(1, \gamma)] + (1-\phi_0\alpha)\Delta_1}{2 + \phi_0[\alpha + \gamma(1-\alpha)](T-2)} \]
and that
\[ B = \lim_{\delta \rightarrow 1} V_+ \leq y(\phi_1, \gamma) + \frac{\Delta + \phi_0\alpha[y(1, \alpha) - y(1, \gamma)] + (1-\phi_0\alpha)\Delta_1}{2 + \phi_0[\alpha + \gamma(1-\alpha)](T-2)} \phi_1 \gamma(T-2). \]

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 (9).

Consider now the case in which \( \ell \geq 1 \) and recall that (A3) implies that \( \tilde{w}_1^N = 0 \). The same argument used in Step 1 shows that the firm’s payoff is bounded above by
\[ V' = (1-\delta)y(\phi_1, \gamma) + \delta(1-\delta)\phi_1 \gamma[y(1, \gamma) - \bar{w}] + (1-\phi_1 \gamma)y(\phi_2^*, \gamma) \]
\[ + \phi_1(\gamma + (1-\gamma))\delta^2(1-\delta^{T-3})[y(1, \alpha) - \bar{w} - V] + \delta^2 V, \]

57
where \( \phi_2^* = (1 - \alpha)(1 - \gamma)\phi_0 / [(1 - \alpha)(1 - \gamma)\phi_0 + 1 - \phi_0] < \phi_2 \) is the reputation of an age 3 worker who exerts effort when of age 1 and fails to produce high output in his first two periods of employment. Consider the following deviation for the firm: offer \((w, 0)\) with \(w = 0\) to the worker under consideration and replace him with an age 1 worker if he produces low output. A lower bound on the firm’s payoff from this deviation is

\[
\tilde{V} = (1 - \delta)y(\phi_1, \gamma) + \delta(1 - \delta)\{\phi_1\gamma[y(1, \alpha) - \overline{w}] + (1 - \phi_1\gamma)y(\phi_0, \alpha)\}
+ (\phi_1\gamma + (1 - \phi_1\gamma)\phi_0\alpha)\{\delta^2(1 - \delta^{T-3})[y(1, \alpha) - \overline{w} - V]\} + \delta^2V.
\]

Recall from Step 1 that a lower bound on the firm’s payoff when it employs an age 1 worker is

\[
(1 - \delta)y(\phi_0, \alpha) + \delta V + \phi_0\alpha\delta(1 - \delta^{T-2})[y(1, \alpha) - \overline{w} - V].
\]

Since \(\phi_1(\gamma + \gamma(1 - \gamma)) = \phi_1\gamma + (1 - \phi_1\gamma)\phi_2^* \gamma\) and \(\phi_0\alpha > \phi_2^* \gamma\), we then have that \(\tilde{V} > V'\) if

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

The last condition is equivalent to \(\phi_1\gamma < \phi_1\gamma\alpha + (1 - \phi_1\gamma)\phi_0\alpha\), which is satisfied by (9); just note that \(\phi_0\alpha(1 - \alpha + \phi_0\alpha)^{-1} > \phi_0\alpha[1 + \alpha(1 - \phi_0)]\). Thus, once again the firm is better off replacing the age 2 worker, a contradiction. It is immediate to see that this last case also holds without (A3).

**PROOF OF LEMMA 8**

We are done if we show that it is never optimal for the firm to make an offer with \(q \geq 2\) to an age 1 worker. Suppose, by contradiction, that the firm makes an offer with \(q \geq 2\) to an age 1 worker. There are two cases to consider: (i) the worker does not exert effort when of age 1; and (ii) the worker exerts effort when of age 1. Consider case (i) first. In this case, Step 1 of the proof of Proposition 2 shows that the firm can profitably deviate by offering \(q = 0\) to the worker and then dismissing him after low output. Consider now case (ii). There are two possibilities: either the worker does not exert effort when of age 2 or the worker exerts effort when of age 2 if he produces low output when of age 1. Since an age 1 worker exerts effort when offered \(q = 1\), in both subcases the firm can profitably deviate by
making an offer with \( q = 1 \) to the worker and dismissing him when he is of age 2 if he has not revealed himself to be of the high-ability type by then. The first subcase follows from Lemma 7, and the second subcase follows from Step 2 of Proposition 2.

CONDITIONS (8), (15), AND (A1) TO (A3)

Let \( \gamma = \kappa \alpha \) with \( \kappa > 1 \) and \( \phi_0 \kappa < 1 \). Recall that a necessary condition for (A1) is that \( \phi_0 \gamma < \alpha \). Straightforward algebra shows that (8) is equivalent to

\[
\kappa \leq \kappa^1_+ = \phi_0 + \frac{1 - \phi_0}{(1 - \alpha)^2}.
\]

When \( T \) is large, condition (15) becomes

\[
1 + \psi \leq \frac{(1 - \alpha) \kappa}{1 - \phi_0 \alpha} \leq 1 + \min \left\{ \frac{\phi_0 (\kappa - 1) \alpha^2 \Delta_y}{\Delta + \phi_0 \alpha [1 - (\kappa - 1) \alpha] \Delta_y}, \alpha (1 - \phi_0) \right\}. \quad (30)
\]

Given that \( \psi = \alpha^2 \phi_0 (1 - \phi_0) \Delta_y / (1 - \phi_0 \alpha) \Delta \) and \( \Delta = \alpha (1 - \phi_0) \Delta_y - \eta < \alpha (1 - \phi_0) \Delta_y \), we then have that \( (1 - \phi_0 \alpha) (1 + \psi) > 1 \). Hence, a necessary condition for (30) is that \( \kappa (1 - \alpha) > 1 \).

Since we also need \( \kappa \alpha < 1 \), condition (30) can then be satisfied only if \( \alpha < 1 - \alpha \).\(^{23}\) We assume that \( \alpha \) is such that \( \alpha < 1/2 \) in what follows.

Now assume that \( \Delta \) is such that

\[
\Delta = \phi_0 (1 + \lambda) \alpha (\kappa - 1) \Delta_y, \quad (31)
\]

with \( 0 < \lambda < (1 - \phi_0 \kappa) / (\kappa - 1) \). In this case, condition (30) becomes

\[
1 + \frac{\alpha (1 - \phi_0)}{(1 - \phi_0 \alpha)(\kappa - 1)(1 + \lambda)} \leq \frac{(1 - \alpha) \kappa}{1 - \phi_0 \alpha} \leq \min \left\{ 1 + \frac{(\kappa - 1) \alpha}{\kappa + (\lambda - \alpha)(\kappa - 1)}, 1 + \alpha (1 - \phi_0) \right\}. \quad (32)
\]

We can ensure the above value of \( \Delta \) by adjusting \( \Delta_y \) accordingly. Indeed, given that \( \Delta = \alpha (1 - \phi_0 \kappa) \Delta_y - \eta + \phi_0 \alpha (\kappa - 1) \Delta_y \), we obtain the desired value of \( \Delta \) if we set \( \Delta_y \) to be such that \( \phi_0 \lambda \alpha (\kappa - 1) \Delta_y = \alpha (1 - \phi_0 \kappa) \Delta_y - \eta \); the restriction on \( \lambda \) implies that this is possible.

In particular, (A1) is satisfied when \( \Delta \) is given by (31).

\(^{23}\)In fact, given that \( \phi_0 \alpha (1 + \psi) (T - 1) / (T - 2) > \phi_0 \alpha (1 + \psi) \) for all \( T \geq 3 \), \( \alpha < 1/2 \) is necessary for (15) regardless of the value of \( T \).
Suppose then that $\phi_0 \approx 0$ and $\lambda \approx 0$. In this case, the first inequality in (32) becomes

$$1 + \frac{\alpha}{\kappa - 1} \leq (1 - \alpha)\kappa,$$

while the second inequality in (32) becomes

$$\kappa \leq \kappa^2_+ = \max \left\{ \frac{1}{1 - \alpha} + \frac{\alpha^2}{(1 - \alpha)^2}, \frac{1}{1 - \alpha} + \frac{\alpha}{1 - \alpha} \right\}.$$

Since the left-hand side of (33) is strictly decreasing in $\kappa$, while its right-hand side is strictly increasing in $\kappa$, (33) is equivalent to $\kappa \geq \kappa_-$, where $\kappa_-$ is the only value of $\kappa$ for which (33) holds with equality. Now observe that $\alpha < 1 - \alpha$ implies that $\kappa^2_+ < \kappa^1_+$ and that $\kappa^2_+ > \kappa_-$ if, and only if,

$$(1 - \alpha)^3 < \alpha^2. \quad (34)$$

Moreover, observe that $\kappa^2_+ \alpha < 1$ if, and only if,

$$\alpha^3 < (1 - \alpha)(1 - 2\alpha). \quad (35)$$

It is easy to see that (34) and (35) can be jointly satisfied; for instance, $\alpha = 2/5$ satisfies both conditions.

We have then established the following result: for all $\alpha \in (0, 1/2)$ satisfying (34) and (35), there exist $1 < \kappa < \kappa^*$ with $\kappa \alpha < 1$, $\bar{\phi}_0 > 0$ with $\bar{\phi}_0 \kappa < 1$, and $0 < \bar{\lambda} < 1/(\kappa - 1)$, such that if $\gamma \in (\kappa \alpha, \kappa \alpha)$, $\phi_0 \in (0, \bar{\phi}_0)$, and $\Delta$ is given by (31) with $\lambda \in (0, \bar{\lambda})$, then (A1) holds and there exists $T^* \geq 3$ with the property that (8) and (15) are satisfied for all $T \geq T^*$.

Finally, notice that, by increasing $T^*$ if necessary, conditions (A2) and (A3) are satisfied for all $T \geq T^*$ when $\delta \approx 1$.

NON-IID CASE WHEN $e_- = e_-$ FOR AGE 1 WORKERS

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

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

Recall that (A1) is equivalent to $y(\phi_0, \alpha) + \phi_0[y(1, \gamma) - y(1, \alpha)] < y(1, \alpha) - \bar{\omega}$. 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-ability type. Notice that (A1') is satisfied if $\phi_0$ is sufficiently small.

First observe that Lemma 5 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-ability type is less than $y(1, \alpha) - \overline{w}$. Now consider the following auxiliary environment: $e_\alpha = e$ for age 1 workers and the probability that a worker is of high ability is $\phi_0$ such that

$$\frac{(1 - \gamma)\phi_0}{1 - \phi_0 \gamma} = \frac{(1 - \alpha)\phi'_0}{1 - \phi'_0 \alpha}.$$  \hspace{1cm} (36)

For fixed $\alpha$ and $\gamma$, equation (36) defines a one-to-one map between $\phi_0$ and $\phi'_0$. 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 (8), (15), and (A1) to (A3) can be satisfied when the ex ante probability that workers are of high ability is small. We know from the above analysis of these five conditions that this is indeed the case.

**Appendix D: 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 is information beneficial.

Suppose the firm offers $(w^0, (1, w^1))$ to an age 1 worker. We know from Appendix C that Lemma 7 is valid without (A3). Hence, the worker is dismissed when of age 3 if he has not revealed himself to be of high ability by then, so that he has no incentive to exert effort when of age 2 if he produces low output when of age 1. Now let: \( R_+ = (1 - \delta)^{-1}(1 - \delta^{T-2})v(\overline{w}); \)
(ii) \( R_- = (1 - \delta)^{-1}(1 - \delta^{T-2})v(w) \); and (iii) \( R_0 = (1 - \delta)^{-1}(1 - \delta^{T-1})v(w) \). The worker’s incentive compatibility constraint for effort exertion when of age 1 is given by

\[
-c + (1 - \phi_0 \alpha)\delta \max \left\{ v(w_1) + \phi_1 \gamma R_1 + (1 - \phi_1 \gamma)\delta R_-, R_0 \right\} \\
\geq (1 - \phi_0 \alpha)\delta \max \left\{ v(w_1) + \phi_1 \alpha R_1 + (1 - \phi_1 \alpha)\delta R_-, R_0 \right\}.
\]

(37)

Notice that (A2) implies that (37) is satisfied when

\[
v(w_1) + \phi_1 \alpha R_1 + (1 - \phi_1 \alpha)\delta R_0 \geq R_0.
\]

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

\[
v(w_1) + \phi_1 \gamma R_1 + (1 - \phi_1 \gamma)\delta R_0 > R_0;
\]

(38)

\[
v(w_1) + \phi_1 \alpha R_1 + (1 - \phi_1 \alpha)\delta R_0 \leq R_0.
\]

(39)

That (38) is necessary is immediate, for otherwise the left-hand side of (37) is smaller than its right-hand side. Next, suppose that (39) 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, (37) and (39) imply that his lifetime payoff is bounded below by

\[
v(w_1) + \phi_0 \alpha (1 - \delta)^{-1}(1 - \delta^{T-1})v(w_1) + (1 - \phi_0 \alpha)\delta R_0 \\
\geq v(w_1) + \phi_0 \alpha R_1 + (1 - \phi_0 \alpha)\delta R_0 + \delta(1 - \delta)^{-1}(1 - \delta^{T-2} - \delta^{T-1})v(w) \\
\geq (1 - \delta)^{-1}(1 - \delta^{T})v(w),
\]

where the second inequality follows from (39) 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 (37) and the worker’s participation constraints. Now define \( \varepsilon > 0 \) and \( \kappa \) to be such that

\[
v(w_1 - \varepsilon) + \phi_1 \alpha \delta(R_1 - R_-) + \delta R_- = R_0;
\]

\[
v(w_0 + \kappa) - v(w_0) = (1 - \phi_0 \alpha)\delta[v(w_1) - v(w_1 - \varepsilon)].
\]

Since \( w_0 \leq w_1 \), the same argument as in the proof of Claim 1 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
where \( \omega^0 \) to \( \omega^0 + \kappa \), it reduces the cost of employing the worker while still satisfying (37) and his participation constraints.

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

\[
\begin{align*}
\min & \quad \omega^0 + \delta(1 - \phi_0 \alpha)w^1 \\
\text{s.t.} & \quad -c + (1 - \phi_0 \alpha)\delta \left[ v(w^1) + \phi_1 \gamma \delta R_+ + (1 - \phi_1 \gamma) \delta R_- \right] \geq (1 - \phi_0 \alpha)\delta R_0, \\
& \quad v(w^0) - c + \phi_0 \alpha \delta (1 - \delta)^{-1}(1 - \delta^{T-1})v(\omega) \\
& \quad + (1 - \phi_0 \alpha)\delta \left[ v(w^1) + \phi_1 \gamma \delta R_+ + (1 - \phi_1 \gamma) \delta R_- \right] \geq (1 - \delta)^{-1}(1 - \delta^T)v(w).
\end{align*}
\]

Since \( v \) is strictly concave, it is easy to see that \((M)\) has a unique solution, denoted by \((\hat{\omega}_1^N, \hat{\omega}_2^N)\). 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 \( \omega^0 = \hat{\omega}_1^N \) and \( \omega^1 = \hat{\omega}_2^N \). Let \( \lambda \) and \( \mu \) be, respectively, the multipliers associated with the first and second inequality constraints of \((M)\). The (necessary and sufficient) Kuhn-Tucker conditions for \((M)\) are

\[
\begin{align*}
-1 + \mu v^0 & = 0; \quad (40) \\
-1 + \mu v^1 + \lambda v^1 & = 0. \quad (41)
\end{align*}
\]

Equation (40) implies that \( \mu = 1/v^0 > 0 \), in which case we can rewrite (41) as

\[
\frac{v^1}{v^0} + \lambda v^1 = 1.
\]

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

\[
\begin{align*}
v(\hat{\omega}_1^N) & \leq v(\omega) - \phi_0 \alpha \delta (1 - \delta)^{-1}(1 - \delta^{T-1})[v(\omega) - v(\omega)]; \quad (42) \\
v(\hat{\omega}_2^N) & \geq v(\omega) - \phi_1 \gamma \delta (1 - \delta)^{-1}(1 - \delta^{T-2})[v(\omega) - v(\omega)] + \frac{c}{(1 - \phi_0 \alpha)\delta}. \quad (43)
\end{align*}
\]

The following result shows that \( \hat{\omega}_1^N = \hat{\omega}_2^N \) when \( c \) is small enough. Recall that \( \psi = [y(\phi_0, \alpha) - y(\phi_1, \alpha)]/\Delta \), where \( \Delta = [y(1, \alpha) - \omega] - y(\phi_0, \alpha) \).
Lemma 9. Suppose that \( \phi_1 \gamma > \phi_0 \alpha (1 + \psi) (T - 1) / (T - 2) \). There exists \( \tau = \tau (\delta) > 0 \) such that if \( c < \tau \), then \( \hat{w}^N_1 = \hat{w}^N_2 = \hat{w} \), where \( \hat{w} \) is such that

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

Proof: Suppose that \( \lambda > 0 \), in which case both (42) and (43) hold with equality. Given that \( (1 - \delta^{T - 1}) / (1 - \delta^{T - 2}) \) is increasing in \( \delta \), we have that

\[
v(\hat{w}^N_1) - v(\hat{w}^N_2) \geq \phi_0 \alpha \psi \frac{T - 1}{T - 2} [v(w) - v(\hat{w})] - \frac{c}{(1 - \phi_0 \alpha) \delta}.
\]

Thus, if \( c < \tau \), where

\[
\tau = (1 - \phi_0 \alpha) \phi_0 \alpha \psi \frac{T - 1}{T - 2} \delta^2 [v(w) - v(\hat{w})],
\]

we have that \( \hat{w}^N_2 < \hat{w}^N_1 \), a contradiction. Therefore, \( \lambda = 0 \) when \( c < \tau \), in which case we know \( \hat{w}^N_1 = \hat{w}^N_2 \). The desired result follows from the fact that \( \mu > 0 \). \( \blacksquare \)

Let \( V_2 \) denote the firm’s payoff when it always offers \((\hat{w}^N_1, (1, \hat{w}^N_2))\) to age 1 workers. The same algebra as in Appendix C shows that

\[
V_2 = y(1, \alpha) - \bar{w} - \frac{(1 - \delta) \Delta + \delta(1 - \delta) \{ \phi_0 \alpha [y(1, \alpha) - y(1, \gamma)] + (1 - \phi_0 \alpha) \Delta_1 \}}{1 - \delta^2 + \phi_0 [\alpha + \gamma (1 - \alpha)] \delta^2 (1 - \delta^{T - 2})}
= \lambda_2 [y(\phi_0, \alpha) - \hat{w}^N_1] + (1 - \lambda_2) [y(1, \alpha) - \bar{w}] - \lambda_2 \delta \{ \phi_0 \alpha [y(1, \alpha) - y(1, \gamma)] + (1 - \phi_0 \alpha) \Delta_1 \},
\]

where \( \Delta = y(1, \alpha) - \bar{w} - [y(\phi_0, \alpha) - \hat{w}^N_1] \), \( \Delta_1 = y(1, \alpha) - \bar{w} - [y(\phi_1, \gamma) - \hat{w}^N_2] \), and \( \lambda_2 \) is the same as in the main text. 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 for the case under consideration. In order to do so, let \( \hat{w}^N_2 \) be such that

\[
v(\hat{w}^N_2) = v(w) - \phi_1 \gamma \delta (1 - \delta)^{-1} (1 - \delta^{T - 2}) [v(w) - v(\hat{w})].
\]

Notice that \( \hat{w}^N_2 < \hat{w}^N_2 \). By construction, \( \hat{w}^N_2 \) 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 \( \Delta_w = \hat{w}^N_1 - \hat{w}^N_2 \leq 0 \) and \( \Delta'_w = \hat{w}^N_2 - \hat{w}^N_2 > 0 \). Notice that if \( \hat{w}^N_2 = \hat{w}^N_2 \), so that \( \Delta_w = 0 \), then condition (45) in Proposition 4 below is more stringent than condition (9) in Proposition 2.
Proposition 4. Suppose that $\phi_2 \gamma \leq \phi_0 \alpha$ and

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

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 do so.

Proof: We know from Step 1 in the proof of Proposition 2 in Appendix C that if $\phi_1 \gamma < \phi_0 \alpha [1 + \alpha (1 - \phi_0)]$, then the firm employs a worker of age 2 who did not exert effort and produced low output when of age 1 only if committed to do so. Next, consider the case in which the firm has an incumbent of age 2 who exerted effort and produced low output when of age 1. 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

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

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

$$
[y(\phi_1, \gamma) - \bar{w}_1^N] + \phi_1 \gamma \frac{\delta(1 - \delta^{T-2})}{1 - \delta} [y(1, \alpha) - \bar{w} - V_2] < V_2.
$$

The same algebra as in the proof of Proposition 2 shows that the last condition reduces to (45) when $\delta = 1$. This completes the proof of Proposition 4. 

We now establish the equivalent of Lemma 8 to the case under consideration. In order to do so, let $V_1$ be the firm’s greatest payoff when it always makes an offer with $q = 0$ to the age 1 workers. It is easy to see that

$$
V_1 = \lambda_1 [y(\phi_0, \alpha) - \bar{w}_1^N] + (1 - \lambda_1) [y(1, \alpha) - \bar{w}],
$$

where $\lambda_1$ is the probability that the firm makes an offer with $q = 0$ to the age 1 workers.
where $\lambda_1$ is the same as in the main text and $\bar{w}_1^N$ is the same as in the proof of Lemma 7. Notice that $\bar{w}_1^N \leq \bar{w}_1^N$. The same argument as in the proof of Proposition 1 shows that $V_1$ is greater than the firm’s payoff when it always offers $(w^0, (1, w^1))$ to age 1 workers, but it does not induce them to exert effort (that is, $w^1$ is such that (37) is not satisfied). We then have the following result; the proof is the same as the proof of Lemma 8.

**Lemma 10.** Suppose that $\phi_2^{\gamma} \leq \phi_0^{\alpha}$ and (45) holds. If $\delta > \delta^*$, then commitment to employment is beneficial if, and only if, $V_2 > V_1$.

Given that

$$V_1 = \lambda_1[y(\phi_0, \alpha) - \bar{w}_1^N] + (1 - \lambda_1)[y(1, \alpha) - \bar{w}] - \lambda_1 \Delta''_w,$$

where $\Delta''_w = \bar{w}_1^N - \bar{w}_1^N$ and $\Delta_1 = \Delta + y(\phi_0, \alpha) - y(\phi_1, \gamma) - \Delta_w$, we then have that

$$V_2 - V_1 = \{\lambda_1(1 + \xi) - \lambda_2[1 + \delta(1 - \phi_0^{\alpha})(1 + \psi - \chi)]\} \Delta + \rho,$$

where $\xi = \Delta''_w / \Delta$, $\chi = \Delta_w / \Delta$, and $\psi$ is the same as in the main text. Straightforward algebra shows that commitment to employment is information beneficial, that is, $V_2 - V_1 > \rho$, if, and only if,

$$\phi_1^{\gamma} > \phi_0^{\alpha} \frac{(1 + \psi - \chi)}{1 + \xi} \frac{1 - \delta^{T-1}}{1 - \delta^{T-2}} + \frac{(\psi - \chi)}{1 + \xi} \frac{(1 - \delta)}{\delta(1 - \delta^{T-2})}.$$  \(\text{(46)}\)

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

$$\phi_1^{\gamma} > \phi_0^{\alpha} \frac{(1 + \psi - \chi)}{1 + \xi} \frac{1 - \delta}{T - 2}.$$  \(\text{(47)}\)

We can now state the equivalent of Proposition 3 to the case under consideration. It follows immediately from Lemma 10 and the discussion that follows it.

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

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

$$< \min \left\{ \phi_0^{\alpha} \left[ \frac{(1 - \phi_0^{\alpha})[\Delta''_w + \Delta + \phi_0^{\alpha}(\gamma - \alpha)\Delta_y(T - 2)]}{(1 + \phi_0^{\alpha})\Delta_y + \{\Delta + \phi_0^{\alpha}[1 - (\gamma - \alpha)]\Delta_y\}(T - 2)} \right] \right\}.$$  \(\text{(48)}\)
There exists \( \delta^{**} \in (0,1) \) such that if \( \delta > \delta^{**} \), then: (i) the firm always offers probation to age 1 workers in equilibrium; and (ii) commitment to employment is information beneficial.

Suppose \( \phi_0, \alpha, \gamma, \Delta_y, \) and \( T \) are such that: (i) (A1) and the conditions of Proposition 3 are satisfied; and (ii) (A2) is satisfied when \( \delta \approx 1 \). We know from Appendix C that such values exist. Now assume that \( \delta > \delta^{*} \) and let \( c < \tau(\delta^{*}) \). Given that \( v \) is concave implies that

\[
v(\hat{w}_2^N) - v(\overline{w}_2^N) \geq v'(\overline{w})[\hat{w}_2^N - \overline{w}_2^N] = v'(\overline{w})\Delta_w',\]

Lemma 9 implies that \( \Delta_w = \hat{w}_1^N - \hat{w}_2^N = 0 \) and that

\[
\Delta_w' \leq \frac{1}{v'(\overline{w})} \left\{ \frac{\phi_1 \gamma \delta (1-\delta)^{-1}(1-\delta^{T-2})[v(\overline{w}) - v(\hat{w})] + c}{1 + \delta (1 - \phi_0 \alpha)} \right\} \leq A(\phi_0, \alpha, \gamma, T) \frac{[v(\overline{w}) - v(w)]}{v'(\overline{w})},
\]

where \( A(\phi_0, \alpha, \gamma, T) \) is a constant term that depends on \( \phi_0, \alpha, \gamma, \) and \( T \). The second inequality follows from (44). Now observe that since (15) is satisfied by assumption, the fact that \( \xi \geq 0 \) and \( \chi = 0 \) (given that \( \Delta_w = 0 \)) implies that (47) is satisfied as long as \( \Delta_w' \) is sufficiently close to zero. Since we can take \( v'(\overline{w}) \) as large as we want while adjusting \( \overline{w} \) to keep \( v(\overline{w}) - v(w) \) constant, and thus preserving (A1) and (A2), it is possible to have \( \Delta_w' \) as small as one wants.

Summarizing: there exist \( \phi_0, \alpha, \gamma, \Delta_y, \) and \( T \) such that if \( \delta \) is close to one, \( c \) is small enough, and \( v'(\overline{w}) \) is large enough, then (A1), (A2), and the conditions in Proposition 5 are simultaneously satisfied.