Escola de
ECONOMIA
de São Paulo

LEARNING–BY–EMPLOYING: THE VALUE OF COMMITMENT

UNDER UNCERTAINTY

BRAZ CAMARGO

ELENA PASTORINO

FGV

EESP

Learning–by–Employing: The Value of Commitment Under Uncertainty[*]

Braz Camargo[†]

University of Western Ontario

Elena Pastorino[‡]

University of Iowa

August 2008

Abstract

We analyze a dynamic principal–agent model where an infinitely–lived principal faces a sequence of finitely–lived agents who differ in their ability to produce output. The ability of an agent is initially unknown to both him and the principal. An agent's effort affects the information on ability that is conveyed by performance. We characterize the equilibrium contracts and show that they display short–term commitment to employment when the impact of effort on output is persistent but delayed. By providing insurance against early termination, commitment encourages agents to exert effort, and thus improves on the principal's ability to identify their talent. We argue that this helps explain the use of probationary appointments in environments in which there exists uncertainty about individual ability.

Keywords: dynamic principal–agent model, learning, commitment.

JEL Classification: C73, D21, D83, J41, M12, M51, M54.

1 Introduction

It has long been recognized that skilled individuals can only be identified through careful selection. Given the well–known difficulty to identify talent, firms usually employ a range of methods to evaluate job candidates.[1] 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 promise employment for a pre–specified period of time—to determine whether new workers are suited to handle the duties and challenges associated with their jobs.[2]

The use of probation is common to many occupations and professions, including management consulting, the legal profession, and academia. It is generally agreed that in these latter instances a worker's output critically depends on his skill, but the qualities that distinguish a successful individual are typically only revealed through time. Why then should an employer commit to retain a worker of *uncertain* talent for a *certain* period of time rather than decide on employment as the relationship unfolds? Intuitively, if performance on the job provides information about talent, and thus is a signal of future productivity, the flexibility to replace workers whose performance is unsatisfactory should always be valuable to a firm. In this paper we show, on the contrary, that when the talent of new hires is uncertain, an employer might benefit from restricting its ability to dismiss workers early in their careers. The reason is that such commitment can induce workers to invest in generating information about their ability and thus help a firm identify their talent.

Indeed, a firm cannot benefit from offering probation when a worker's talent is uncertain, but the extent to which performance reveals ability is independent of his dedication to his job. In this

---

[1]The following quote, which discusses the limitations of behavioral interviews, illustrates this point: "Despite their advantages, behavioral interviews really only establish a candidate's minimum qualifications; they don't identify star talent. A candidate's experience, for example, is obviously an important hiring factor, but we all know seasoned executives who aren't stars. Similarly, being likable doesn't mean you have the intellectual horsepower to be a stellar leader. (...) Knowledge is information acquired through experience or formal training. Intelligence is the skill with which someone uses knowledge to solve a problem." [Justin Menkes, managing director of the *Executive Intelligence Group*, a New York–based consulting firm focused on the assessment of executive talent, as quoted in "*Hiring for Smart*", Harvard Business Review, Vol. 83, No. 11, November 2005.]

[2]The terminology is not uniform. Probationary periods are also understood as the stage at the beginning of an employment relationship during which an employer has greater discretion to dismiss workers. This is common in unionized industries and is not the object of our study.

case, the only effect of commitment to employment is to prevent the firm from dismissing a worker once he is perceived to be ill–suited to his job. In many instances, however, not only performance conveys information about ability, but the precision of this information is directly affected by a worker's effort on the job. For instance, whether a restructuring project is successful in addressing the needs of a client firm depends both on the talent of the consultants involved and on their commitment. Likewise, a talented lawyer is more likely to develop a successful legal strategy for a client if he is fully engaged in his work. In all of these circumstances, the prospect of an early dismissal might discourage a worker from dedicating himself to his job despite the fact that exerting effort helps reveal talent. Offering probation may then be valuable to a firm if insurance against early failure stimulates workers to produce informative signals about their ability.

In order to analyze the tradeoffs involved in the use of probation, we consider a labor market where an infinitely–lived firm faces a constant inflow of finitely–lived workers. At any date, the firm can employ at most one worker.[3] We model probation as a short–term commitment to employment. Workers differ in their ability to produce output and the talent of a worker is initially unknown to both him and the firm. The performance of an employed worker also depends on his choice of effort. We assume that effort increases the probability of good performance only if the worker is talented. Thus, effort makes a worker's output more informative about his ability.

Every worker in the market has available an outside option whose value increases with his reputation, which we model as the firm's belief that the worker is of high ability. Then, workers of higher reputation are perceived as more productive, but can only be employed at a higher wage. The value of a worker nevertheless increases with his reputation. As a consequence, the firm faces an opportunity cost by retaining a worker whose initial performance is poor: such a worker is less likely to be talented than a worker who is new to the market.

In the paper we identify circumstances under which probation is beneficial by contrasting two cases distinguished by the effect of effort on output. In the first case, which is 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 the current and the future performance of a worker,

---

[3]The key assumption is that the firm is capacity–constrained, and so cannot create vacancies to absorb all workers in the market.

but the effect is mostly on the latter. In other words, effort has a persistent but delayed impact on output. We refer to the second case as the non–IID case.

In this environment, a worker who has yet to prove his talent—in particular, a worker who is new to the market—has an incentive to exert effort due to a concern for his future career. Indeed, effort makes performance more informative about ability and a worker is rewarded with a wage increase if he proves his talent—a worker 'invests' in his reputation when he exerts effort. The gain to the firm if workers exert effort is an improved ability to screen them. The firm then benefits from offering probation to a worker who is new to the market if this strengthens his 'career concerns' motive to exert effort. This commitment, however, can be costly, for it prevents the firm from dismissing the worker after he performs poorly.

We establish that in the IID case probation never increases, and can in fact decrease, a worker's reputational incentive to exert effort. The firm can benefit from offering probation in the non–IID case, though. The reason for this result is that in the non–IID case the incentive problem of a new worker is compounded by the time separation of costs and returns typical of investment problems. When the impact of effort is mostly on future output, the worker can gain from investing in his reputation only if he is guaranteed to participate in its return, that is, only if employment lasts until the impact of effort on output materializes. The firm then destroys a worker's incentive to exert effort if it dismisses him after poor performance. As a result, the firm can gain from offering probation exactly when it would otherwise dismiss a newly hired worker who exerted effort but failed to perform well. In other words, probation can help a firm overcome its incentive to dismiss under–performing workers and thus induce new hires to generate valuable information about their ability. Hence, offering insurance against bad performance can actually *increase* a worker's incentive to exert effort.

The rest of the paper is organized as follows. We discuss the related literature in the next section and introduce the model in Section 3. Section 4 contains some preliminary results. We consider the IID case in Section 5 and the non–IID case in Section 6. Section 7 shows that explicit–output contingent contracts cannot substitute for the use of probation. Section 8 concludes the paper. Appendix A contains all the omitted proofs and Appendix B contains omitted details.

2 Related Literature

Formally, the problem of the firm in our setting is a multi–armed bandit, the sequential sampling problem of a decision maker who has to choose between a (finite) number of alternatives with uncertain rewards.[4] The first application of the multi–armed bandit framework to the analysis of employer learning in labor markets is Jovanovic (1979). Harris and Weiss (1984) and Eeckhout (2006) extend Jovanovic's analysis to the case where productivity is general instead of match–specific. A fundamental difference between these papers and ours is that in our framework rewards are endogenous: the behavior of the firm affects the decisions of workers.[5]

A small literature considers the problem of repeated moral hazard with effort persistence. Jarque (2005) 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. A few papers have also analyzed dynamic incentives in the presence of both adverse selection and moral hazard. Laffont and Tirole (1988) show, in a two–period model, that short–term contracts might not be sufficient to induce an informed agent to reveal his private information. More recently, Jeitschko and Mirman (2002a), also in a two–period setting, analyze how optimal short–term contracts trade off up–front payments for self–selection, and thus the amount of information transmission, against the second–period informational rents to better types of agents. Banks and Sundaram (1998) consider the problem of optimal agent–retention by a long–lived principal in the absence of commitment when agents live for two periods.[6]

A related paper on contracting in the presence of uncertainty is Manso (2007), which investigates the extent to which contracts can motivate innovation: the discovery, through experimentation and learning, of alternatives that are better than the currently known actions. He shows, in a two–period model, that optimal long–term wage contracts that stimulate innovation can be non–monotone in performance, that is, they might reward early failure and late success.

---

[4]See Berry and Fristedt (1985) for an exposition of the theory of multi–armed bandits.

[5]The firm faces an experimentation problem with signal–dependence when effort is persistent. See Datta, Mirman, and Schlee (2002) for an analysis of such a model.

[6]Osband (1989) investigates the trade-off between information acquisition and incentives for effort exertion when the principal relies on an expert of unknown ability to forecast the mean of an uncertain outcome.

Two papers that also analyze the interplay between explicit and implicit incentives in dynamic principal–agent problems are Gibbons and Murphy (1992) and Levin (2003). The first derives the optimal combination of explicit and implicit contracts in a model of career concerns. The second characterizes optimal stationary self–enforcing contracts in an environment in which explicit contracts can extend for only one period.

In our setting the nature of the relationship between the firm and a worker resembles that of two parties in a hold–up problem (see Klein, Crawford, and Alchian (1978) for an early reference). In the non–IID case, despite the value of information about ability being general, workers underinvest in their reputation unless protected by a probationary appointment. Che and Sakovics (2004) analyze the role of contracts in a dynamic model of hold–up and provide conditions under which trade contracts are beneficial.

Equilibrium contracts in our setting have the 'up–or–out' feature that after probation a worker is either permanently retained or dismissed. Kahn and Huberman (1988) and Waldman (1990) show how up–or–out contracts that specify that an individual should be fired if not promoted within some set time can induce workers to invest in the acquisition of human capital. O'Flaherty and Siow (1992) analyze a model of on–the–job screening and show that the optimal retention decision is an up–or–out rule.[7] Carmichael (1988) discusses how the institution of tenure can induce academic departments to hire the best available junior researchers.[8] None of these papers, however, consider the optimality of commitment when retention involves an opportunity cost to the firm.

3   Baseline 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. We assume that workers have a strictly increasing and weakly concave Bernoulli utility function $v : \mathbb{R}_+ \to \mathbb{R}$; our analysis

---

[7]O'Flaherty and Siow (1995) use the model of O'Flaherty and Siow (1992) to analyze the use of up–or–out rules in the market for lawyers.

[8]See also Bar–Isaac (2007), who considers the problem of selection and retention in partnerships when the contribution of an individual worker cannot be inferred from the total output of the partnership.

equally applies to risk–neutral and risk–averse workers. Workers live for $T \geq 3$ periods once they enter the market and discount future utility at a rate $\delta_w \in [0,1]$.[9] Each worker can be either of high (H) or low (L) 'ability'. A worker's type is unknown to both him and the firm; the probability that a new entrant to the labor market is of high ability is $\phi_0 \in (0,1)$.

Every period he is employed, a worker can either exert effort ($\bar{e}$), incurring a disutility cost $c > 0$, or not ($\underline{e}$), and can either produce high ($\bar{y}$) or low ($\underline{y}$) output. A worker's choice of effort, which is unobservable, affects his output. We consider two cases. In the first case, our benchmark, the output produced by each type of worker is only affected by his current choice of effort. In the second case, a worker's choice of effort has an impact on both his 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, where $e_- \equiv \underline{e}$ for a worker of age 1. Moreover, let $\Pr\{y|k, e, e_-\}$ be the probability, as a function of e and e_-, that a worker of type $k \in \{L, H\}$ produces output $y \in \{\underline{y}, \bar{y}\}$. In the IID case,

$$\Pr\{\bar{y}|H, e, e_-\} = 1 - \Pr\{\underline{y}|H, e, e_-\} = \alpha + \eta(e) \quad \text{and} \quad \Pr\{\bar{y}|L, e, e_-\} = 0$$

where $\alpha \in (0,1)$, $\eta(\underline{e}) = 0$, and $\eta(\bar{e}) = \eta > 0$. In the non–IID case,

$$\Pr\{\bar{y}|H, e, e_-\} = \alpha + \eta(e_-, e) \quad \text{and} \quad \Pr\{\bar{y}|L, e, e_-\} = 0$$

where $\eta(e, \bar{e}) \geq \eta(e, \underline{e})$, $\eta(\bar{e}, e) \geq \eta(\underline{e}, e)$, and $\eta(\bar{e}, \underline{e}) > 0$.

Note that in both cases a low type worker cannot produce high output. Then, a worker who produces high output reveals that he is talented. Put differently, a worker of high ability can prove his talent if given enough opportunities. Also notice that in both cases a worker's performance is more informative about his ability when he exerts effort: a worker is more likely to produce high output when he exerts effort if he is of high ability. Specifically, $\eta(\bar{e}) > \eta(\underline{e})$ in the IID case and $\eta(e, \bar{e}) \geq \eta(e, \underline{e})$, $\eta(\bar{e}, \underline{e}) > \eta(\underline{e}, \underline{e})$, and $\eta(\bar{e}, \bar{e}) \geq \eta(\underline{e}, \bar{e})$ in the non–IID case.

Workers in the market have available an outside option that pays a wage w_R dependent on their *reputation*, where a worker's reputation is the firm's belief ϕ that he is of the high type. Since a worker's effort choice is private, his belief about his type can differ from his reputation. We assume

<hr/>

[9]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.

that $w_R(1) = \overline{w}$ and $w_R(\phi) = w_R(\phi_0) = \underline{w}$ for all $\phi \leq \phi_0$. Thus, as long as a worker fails to produce high output, his outside option remains constant at $\underline{w}$, but it increases to $\overline{w}$ the first time that he produces high output. We also assume that a worker who collects his outside option can no longer be hired by the firm.[10] Therefore, the value of a worker's outside option is determined by his reputation at the time he first collects it. Finally, workers cannot observe the history of play before they enter the market.[11]

Firm. The firm is infinitely–lived, risk–neutral, and discounts future utility at a rate $\delta_f \in (0,1)$.[12] It can employ at most one worker. The flow payoff to the firm when it does not employ a worker is $\Pi < \underline{y} - \underline{w}$. So, the firm would rather employ a worker it knows is of the low type at his outside option than not employ any worker. We normalize payoffs to the firm by $(1 - \delta_f)$.

Besides an 'incumbent', a worker employed by the firm in the previous period and who is still in the market, the only other worker the firm can employ in a given period is the available age 1 worker. At the beginning of each period the firm can offer a worker to pay him a wage of at least w at the end of the period if he accepts employment. In particular, implicit bonus payments are possible. The firm can also commit to offer a worker of age $k \leq T - 1$ a one–period wage of at least w' in the next $q \in \{1, \ldots, T - k\}$ periods. Thus, an offer to a worker is a list $(w, (q, w'))$ consisting of a one–period wage offer w and the number q of subsequent periods in which the firm is committed to make a one–period wage offer of w' or more to him. If the firm offers $(w, (q, w'))$ with $q \geq 1$ to a worker, then it must propose $(w, 0)$ with $w \geq w'$ to him in the next q periods. Note that w' needs to be specified only if $q \geq 1$. We say the firm offers ($q + 1$ periods of) 'probation' to an age 1 worker when it offers him $(w, (q, w'))$ with $q \geq 1$.[13]

Observe that explicit output–contingent contracts are not possible. We discuss in Section 7 the extent to which our results hold when such contracts are feasible. In order to simplify the

---

[10] The no recall restriction amounts to assuming that once a worker collects his outside option, he no longer generates information about his ability that is valuable to the firm.

[11] Thus, workers cannot coordinate their behavior in response to deviations by the firm. We will see that together with the finite lifetime of workers this prevents the use of implicit bonus payments. In light of our results in the full commitment case, see Section 7, we believe that this assumption is not crucial for our results.

[12] The assumption that the firm and the workers have different discount factors is to emphasize that only the firm's discount factor matters for the results of interest.

[13] It is straightforward to show that our results do not change if the firm can propose contracts of the form $(w, (q, \{w'_s\}_{s=1}^q))$, where w'_s is the smallest one–period wage offer that it is committed to make s periods from now.

exposition, we also assume that the firm cannot offer a worker a wage smaller than $\underline{w}$. We discuss the assumption of limited liability at the end of the section.

The following restriction is a maintained assumption in the rest of the paper:

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

The right–hand side of the above inequality is an upper bound to the flow payoff the firm obtains from an age 1 worker. Thus, (A1) states that the firm's flow profit from employing an age 1 worker is always smaller than its flow payoff from employing a worker of the high type. A necessary condition for (A1) is $\phi_0 < \alpha$.

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 or make an offer, either to the incumbent or to the available age 1 worker. The worker who receives an offer decides whether to accept it or not. In case he accepts the offer, he chooses how much effort to exert, output is realized, and the firm pays him a wage that is not smaller than the wage promised at the beginning of the period. A worker who does not receive an offer or rejects one collects his outside option, and so does the firm if its offer is rejected.

Equilibrium. Let Σ_w be the set of behavior strategies for a worker and Σ_f be the set of behavior strategies for the firm. Since workers do not observe the history of play before they enter the market, the set Σ_w is the same for all workers. A strategy profile for the workers is a map $\sigma_w : \mathbb{N} \to \Sigma$, where $\sigma_w(t)$ is the behavior strategy of the worker who enters the market in period t. A strategy profile $(\sigma_w, \sigma_f) \in \Sigma_w^{\mathbb{N}} \times \Sigma_f$ is *worker–symmetric* if $\sigma_w(t)$ is independent of t. We restrict attention to worker–symmetric perfect Bayesian equilibria.[14]

Discussion. We finish this section with a discussion of some of our assumptions.

Labor Market. As in a standard career concerns model—see Holmström (1999)—we focus on a labor market where talent is valuable but scarce and is revealed over time through performance.

---

[14]One alternative to the assumption that workers follow symmetric strategies is to consider an environment where in each period the identity of the worker who becomes available is random. The approach we follow is simpler.

Thus, a worker is able to command a higher wage once he produces high output. Differently from a standard career concerns model, though, we consider a market where firms possess enough monopsony power to be able to extract more surplus from a match with a talented worker than from a match with a worker of unproven ability. Without imperfect competition, the firm would not benefit from identifying talented workers.

A further departure from standard career concerns models is the assumption we that a worker's outside option does not decrease below its initial level: $w_R(\phi) = \underline{w}$ for all $\phi \leq \phi_0$. This is the case, for instance, if there is a secondary labor market where talent is not valuable for production and the wage is $\underline{w}$.

Limited Liability. We assume the firm cannot offer a one–period wage that is smaller than $\underline{w}$. Notice, however, that a worker of age $k \leq T - 1$ who has yet to prove his talent—this includes age 1 workers—is willing to work for the firm for less than $\underline{w}$. Indeed, by working for the firm, such a worker has the chance to prove that he is of high ability, in which case his outside option increases from $\underline{w}$ to $\overline{w}$. The key observation is that among the workers of unproven talent, the one who is willing to sacrifice the most in terms of present wage payments is the age 1 worker. So, with or without limited liability, the firm still faces an opportunity cost when it retains a worker of uncertain ability after he performs poorly.

Effort. Effort has two roles in our environment. First, it increases the likelihood of high output. Second, it makes performance more informative about ability. We are interested on the second role of effort. As such, we abstract from the incentive problem the firm faces when it employs a worker of high ability by assuming that explicit compensation schemes are not feasible—the only workers who can benefit from effort are the ones who have yet to prove their talent. The focus on the informational role of effort is made possible by (A1), which implies that regardless of his behavior, a talented worker is better for the firm than any other worker.

4 Preliminary Analysis

In this section we present results that will be useful for the analysis to follow. Observe first that both an age 1 worker and an incumbent of age $T - 1$ or less who has never produced high output

always accept an offer by the firm. Indeed, by accepting employment, such a worker receives at least $\underline{w}$ and has the chance to prove himself to be talented before he is of age T.

Lemma 1. *In equilibrium, both an age 1 worker and an incumbent of age $T-1$ or less who has never produced high output accept any offer by the firm.*

It is immediate to see that both a worker known to be of the high type and a worker of age T who has never produced high output always accept an offer where the one–period wage w is greater than their respective outside options. This observation is also true when w is equal to their respective outside options since $\underline{y} - \underline{w} > \Pi$.

Lemma 2. *In equilibrium, both an incumbent known to be of the high type and an incumbent of age T who has never produced high output accept an offer with the one–period wage equal to their respective outside options.*

A consequence of Lemmas 1 and 2 is that the firm never collects its outside option.

Lemma 3. *In equilibrium, the firm makes an offer in every period and its offers are never rejected.*

The next result shows that implicit bonus payments are not possible for workers who are known to be of the high type, and so such workers never exert effort.

Lemma 4. *Suppose the firm has an incumbent it knows is of the high type and let w' be the smallest one–period wage the firm can offer him if it is committed to make him an offer. The following holds in equilibrium: (i) the firm never offers the incumbent a one–period wage greater than $\max\{\overline{w}, w'\}$; (ii) if the firm makes the incumbent an offer, then it never commits to future one–period wage offers greater than $\overline{w}$; (iii) if the incumbent accepts employment, then the firm pays him the one–period wage it offered him; (iv) the incumbent never exerts effort.*

Notice, by (i), that if the firm has an incumbent it knows is of the high type, and is not committed to make him an offer, then it never offers him a one–period wage greater than $\overline{w}$. A sketch of the proof of Lemma 4 is as follows. Consider a worker known to be of the high type who is in his last period of employment. Implicit bonus payments are not possible in this case, for otherwise the firm would have a profitable deviation. Hence, the only incentive for such a worker

11

to exert effort is the variation of his future payoff in his output. This variation, however, is zero even if this period is not the worker's last period of life, for his reputation does not change with his output. Thus, no effort is uniquely optimal for him. This, in turn, implies the firm has no incentive to offer a one–period wage greater than $\max\{\overline{w}, w'\}$ to the worker. The desired result now follows from a backward induction argument and Lemma 2—a consequence of Lemma 2 is that an incumbent known to be of the high type never punishes the firm for a deviation by rejecting an offer with a one–period wage equal to $\overline{w}$.

An immediate corollary to Lemma 4 is that the firm always pays a worker who produces high output for the first time the one–period wage it offered him. Thus, we can only observe implicit bonus payments when the firm either employs an age 1 worker or an incumbent who has never produced high output *and* only after either worker produces low output.

Corollary 1. *In equilibrium, the firm pays a worker who produces high output for the first time the one–period wage it offers him.*

The next result follows from the fact that workers use symmetric strategies. It states that the firm's (expected) continuation value from hiring an age 1 worker is independent of calendar time. This introduces a recursive structure in the firm's problem, which plays a key role in our analysis. In what follows, let $V(h|\sigma)$ denote the firm's expected discounted payoff after a history h when the strategy profile under play is σ.

Lemma 5. *If σ 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.*

By Lemma 4, the firm always obtains the same (flow) payoff when it employs a worker of high ability at wage $\overline{w}$. By (A1), this is also the highest payoff the firm can obtain. Moreover, by Lemma 5, the firm's continuation value when it hires an age 1 worker is independent of calendar time. Thus, the firm retains an incumbent it knows is of the high type both when it is not committed to make him an offer and when it has promised to offer him a one–period wage smaller than $\overline{w}$. For convenience, in what follows we sometimes say that the firm 'offers w' to a worker whenever it makes an offer where the one–period wage is w.

Lemma 6. *Suppose the firm has an incumbent it knows is of the high type and let w' be the smallest one–period wage the firm can offer him if it is committed to make him an offer. In equilibrium, the firm offers $\max\{\overline{w}, w'\}$ to this worker, who accepts such an offer.*

Consider the IID case and suppose the firm has an incumbent who has only produced low output. Then, the firm is better off by hiring the available age 1 worker if the incumbent will never exert effort.

Lemma 7. *Consider an equilibrium in the IID case and suppose the firm has an incumbent who has only produced low output. If the incumbent will never exert effort, then the firm only makes him an offer if it is committed to do so.*

Consider now the non–IID case and suppose the firm has an incumbent who has only produced low output and did not exert effort in the previous period. The same logic of the proof of Lemma 7 proves that if the incumbent will never exert effort, then the firm only makes him an offer if it is committed to do so.

Lemma 8. *Consider an equilibrium in the non–IID case and suppose the firm has an incumbent who has only produced low output and did not exert effort in the previous period. If the incumbent will never exert effort, then the firm only makes him an offer if it is committed to do so.*

5   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. We divide the analysis in two parts. First, we consider the case where

$$\phi_0 \eta \delta_w (1 - \delta_w^{T-1})[v(\overline{w}) - v(\underline{w})] \leq (1 - \delta_w)c. \tag{1}$$

Then, we consider the case where

$$\phi_0 \eta \delta_w (1 - \delta_w^{T-1})[v(\overline{w}) - v(\underline{w})] > (1 - \delta_w)c. \tag{2}$$

Observe that the left–hand side of (1) and (2) is the lifetime payoff gain from effort to an age 1 worker who is dismissed after low output, whereas the right–hand side of each expression is the

discounted one–period cost of effort. Thus, it is only when (2) holds that a worker of age 1 who anticipates he is dismissed after low output has an incentive to exert effort.

When (1) is satisfied we prove that: (*i*) a worker of age 2 or more who has only produced low output never exerts effort when employed; and (*ii*) offering probation to an age 1 worker decreases his incentive to exert effort. Thus, it is uniquely optimal for the firm to dismiss any incumbent who has only produced low output. As a result, there exists no equilibrium where the firm offers probation to a worker of age 1 when (1) holds. Equilibria where the firm offers probation to age 1 workers exist when (2) is satisfied, though. We provide an example in Appendix B. Nevertheless, there exists an equilibrium σ^* where the firm *never* offers probation to age 1 workers and such that the firm's payoff in this equilibrium is greater than its payoff in any equilibrium in which it offers commitment. Hence, commitment to employment has no value in either case.

Proposition 1. *Suppose* (1) *holds. There is no equilibrium where the firm offers probation to a worker of age 1.*

A sketch of the proof of Proposition 1 is as follows. First recall that implicit bonus payments are not possible for an incumbent of age T who has never produced high output, and so such a worker has no incentive to exert effort. Consider then a worker of age $k \in \{1, \ldots, T-1\}$ who has only produced low output, let $\pi \leq \phi_0$ be his (private) belief that he is of the high type, and suppose the firm employs him—notice that $\pi < \phi_0$ if $k \geq 2$. The worker's incentive–compatibility constraint for effort exertion is

$$\pi(\alpha + \eta) \left\{ v(w^{\overline{y}}) + \delta_w R(\overline{y}, \overline{e}|\pi, k) \right\} + [1 - \pi(\alpha + \eta)] \left\{ v(w^{\underline{y}}) + \delta_w R(\underline{y}, \overline{e}|\pi, k) \right\}$$

$$\geq c + \pi\alpha \left\{ v(w^{\overline{y}}) + \delta_w R(\overline{y}, \underline{e}|\pi, k) \right\} + (1 - \pi\alpha) \left\{ v(w^{\underline{y}}) + \delta_w R(\underline{y}, \underline{e}|\pi, k) \right\}, \quad (3)$$

where w^y is the wage the firm pays him if his output is y, and $R(y, e|\pi, k)$ is his continuation payoff if he chooses e and produces y, which depends on π and k. Since the worker reveals himself to be of high ability if he produces high output, $R(\overline{y}, \underline{e}|\pi, k) = R(\overline{y}, \overline{e}|\pi, k)$ by Lemmas 4 and 6. This implies that (3) can be rewritten as

$$\pi\eta\underbrace{[v(w^{\overline{y}}) - v(w^{\underline{y}})]}_{\Delta_0} + \pi\eta\delta_w\underbrace{[R(\overline{y}, \overline{e}|\pi, k) - R(\underline{y}, \overline{e}|\pi, k)]}_{\Delta_1} + \delta_w(1 - \pi\alpha)\underbrace{[R(\underline{y}, \overline{e}|\pi, k) - R(\underline{y}, \underline{e}|\pi, k)]}_{\Delta_2} \geq c.$$

$$(4)$$

14

Let w be the one–period wage the firm offered to the worker. By Corollary 1, $w^{\overline{y}} = w$, and so $\Delta_0 \leq 0$. We prove that Δ_1 is bounded above by $(1 - \delta_w)^{-1}(1 - \delta_w^{T-k})[v(\overline{w}) - v(\underline{w})]$, that Δ_2 is bounded above by zero, and that both upper bounds are achieved only if the worker is dismissed after low output. Thus, (4) cannot be satisfied when $k \geq 2$ and is only satisfied when $k = 1$ if the worker is dismissed after low output and (1) holds with equality.

We can then conclude that an incumbent of age 2 or more who has only produced low output never exerts effort and that an age 1 worker has an incentive to exert effort only if (1) holds with equality and he is dismissed after low output. This implies that in equilibrium the firm always offers $(\underline{w}, 0)$ to an age 1 worker and dismisses him if he produces low output. In particular, the firm never offers probation in equilibrium.

Suppose now that (2) holds and let σ^* be the strategy profile where: (i) the firm offers $(\overline{w}, 0)$ to an incumbent it knows is of high ability; (ii) the firm offers the available age 1 worker $(\underline{w}, 0)$ if it has no incumbent or if its incumbent has always produced low output and the firm is not committed to employ him; (iii) the firm offers an incumbent $(w, 0)$ if it is committed to offer him a one–period wage of at least w; (iv) the firm pays the one–period wage it offers; (v) an incumbent of the high type does not exert effort; (vi) the effort choice of any other worker is sequentially rational given the firm's offer and (i) to (v). Notice that under σ^* a worker of age 1 exerts effort if the firm offers him $(\underline{w}, 0)$. Indeed, by (ii), an age 1 worker knows he is dismissed after low output, in which case his incentive–compatibility constraint for effort exertion is satisfied.

We claim that σ^* is an equilibrium, that the firm's payoff under σ^* is the highest payoff that it can obtain, and that the firm's payoff under any equilibrium σ where it offers probation to an age 1 worker is strictly smaller than the payoff it obtains under σ^*.

Proposition 2. *The strategy profile σ^* is an equilibrium and the payoff V^* to the firm under σ^* is the greatest payoff it can obtain in equilibrium. In particular, the payoff to the firm in any equilibrium where it offers probation to an age 1 worker is strictly smaller than V^*.*

By Lemmas 4 and 6, we only need to show that the firm's decision in (ii) is incentive–compatible in order to prove that σ^* is an equilibrium. The incentive–compatibility of (ii) follows from the fact that an age 1 worker exerts effort when the firm offers him $(\underline{w}, 0)$. The second part of Proposition

2 follows from the fact that an incumbent of age 2 or more who has never produced high output is always less profitable to the firm than an age 1 worker who exerts effort.

There exist other equilibria where the firm obtains the same payoff as it obtains under σ^*. For instance, the strategy profile that differs from σ^* only in that the firm offers $(\overline{w}, (T-2, \overline{w}))$ to an age 2 incumbent who has produced high output in his first period of employment is also an equilibrium. The proof of Proposition 2 shows that any equilibrium where the firm's payoff is V^* can differ from σ^* only in the amount commitment the firm offers to an incumbent it knows is of the high type.

6   The non–IID Case

We now consider the case where effort has an impact on both current and future output. In this case we prove that there exists scope for commitment to probation when effort mostly affects future output. For simplicity, we consider the situation in which $\eta(\underline{e}, \overline{e}) = 0$ and $\eta(\overline{e}, \overline{e}) = \eta(\overline{e}, \underline{e})$. It will become clear from our analysis that the results we obtain also hold when $\eta(\underline{e}, \overline{e})$ and $\eta(\overline{e}, \overline{e}) - \eta(\overline{e}, \underline{e})$ are positive but small. Let $\alpha + \eta(\overline{e}, \underline{e}) = \gamma > \alpha$ and recall that $e_- = \underline{e}$ for an age 1 worker. When $\eta(\underline{e}, \overline{e}) = \eta(\overline{e}, \overline{e}) - \eta(\overline{e}, \underline{e}) = 0$, the non–IID case is summarized by the following two information matrices:

$e_- = \underline{e}$	$\overline{y}$	$\underline{y}$
H	α	$1 - \alpha$
L	0	1

$e_- = \overline{e}$	$\overline{y}$	$\underline{y}$
H	γ	$1 - \gamma$
L	0	1

Notice that an age 1 worker has no incentive to exert effort if he anticipates he is dismissed after low output. In fact, a worker of age 1 only benefits from exerting effort if he is retained after low output, for in this case his output when he is of age 2 is more informative about his ability. In what follows we identify circumstances in which an age 1 worker is willing to exert effort if he expects to be retained after low output. This is in stark contrast to the IID case, where either the threat of dismissal after low output is sufficient to induce an age 1 worker to exert effort, or the promise of retention after low output discourages such a worker to exert effort. Thus, the firm can benefit from retaining an age 1 worker who produces low output. Moreover, we will see that there

exist situations in which the firm retains such a worker only if it explicitly promises to do so, that is, only if to commits to employment. The question we address is whether the gain to the firm from inducing an age 1 worker to exert effort through the use of commitment can compensate it for the lack of flexibility in employment decisions that commitment entails.

We divide our analysis in three parts. First, we derive conditions under which the firm can only induce an age 1 worker to exert effort if it offers probation. Then, we identify situations in which the firm benefits from offering commitment. We are interested in the case where the gain to the firm is not due to the extra output a worker produces when he exerts effort, but to the extra information about ability that this effort generates. We conclude by showing that it is possible for commitment to be both necessary for an age 1 worker to exert effort and beneficial for the firm.

6.1 Commitment is Necessary

Let $\underline{\phi} = (1 - \alpha)\phi_0/(1 - \phi_0\alpha)$ and $\underline{\underline{\phi}}(e) = (1 - \alpha)(1 - \xi(e))\phi_0/[(1 - \alpha)(1 - \xi(e))\phi_0 + 1 - \phi_0]$, where $\xi(\underline{e}) = \alpha$ and $\xi(\overline{e}) = \gamma$. Note that $\underline{\phi}$ is the reputation of an age 2 worker who produces low output in his first period of employment and that $\underline{\underline{\phi}}(e)$ is the reputation of an age 3 worker who chooses $e \in \{\underline{e}, \overline{e}\}$ in his first period of employment and produces low output in his first two periods of employment. We make the following two assumptions:

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

(A3) $\underline{\phi}(1 - \alpha)(\gamma - \alpha)\delta_w^2(1 - \delta_w^{T-3})[v(\overline{w}) - v(\underline{w})] < (1 - \delta_w)c$.

In order to understand (A2), consider an age 1 worker who accepts employment and suppose that: (i) the firm offers him $(\underline{w}, 0)$ in the next period if he produces low output, but dismisses him if he produces low output one more time; (ii) the firm never pays him a bonus; (iii) his flow payoff is $v(\overline{w})$ after he produces high output for the first time. Since $\eta(e, \overline{e}) = \eta(e, \underline{e})$ for all $e \in \{\underline{e}, \overline{e}\}$, the worker has no incentive to exert effort when of age 2 if he produces low output in his first period of employment. The incentive–compatibility constraint for effort exertion in his first period

17

of employment is then

$$-(1 - \delta_w)c + \phi_0\alpha\delta_w(1 - \delta_w^{T-1})v(\overline{w})$$

$$+ (1 - \phi_0\alpha)\delta_w \left[(1 - \delta_w)v(\underline{w}) + \underline{\phi}\gamma\delta_w(1 - \delta_w^{T-2})v(\overline{w}) + (1 - \underline{\phi}\gamma)\delta_w(1 - \delta_w^{T-2})v(\underline{w})\right]$$

$$\geq \phi_0\alpha\delta_w(1 - \delta_w^{T-1})v(\overline{w})$$

$$+ (1 - \phi_0\alpha)\delta_w \left[(1 - \delta_w)v(\underline{w}) + \underline{\phi}\alpha\delta_w(1 - \delta_w^{T-2})v(\overline{w}) + (1 - \underline{\phi}\alpha)\delta_w(1 - \delta_w^{T-2})v(\underline{w})\right],$$

which is satisfied by virtue of (A2).

Consider now (A3). It is easy to see that this condition implies that an age 2 worker who failed to produce high output in his first period of employment does not exert effort if (i), (ii), and (iii) from the previous paragraph hold for him. It turns out that (A3) implies that no worker of age 2 or more who has only produced low output has an incentive to exert effort—we prove this in Appendix A. In other words, (A3) implies that the only worker who can ever exert effort is an age 1 worker. A straightforward consequence of this fact is that the firm always pays a worker the one–period wage it offers him.

Lemma 9. *In equilibrium, a worker of age $k \geq 2$ never exerts effort when employed.*

An implication of Lemmas 8 and 9 is that the firm only hires an incumbent of age 3 or more who has never produced high output if it is committed to do so. We now establish that there exist situations in which the firm retains an incumbent of age 2 who failed to produce high output when of age 1 only if it is committed to do so. Hence, in such circumstances, the firm must offer probation to an age 1 worker if he is to exert effort. In what follows, let $y(\phi, \xi) = \phi\xi\overline{y} + (1 - \phi\xi)\underline{y}$ be the expected output of a worker with reputation ϕ when the probability that he produces high output is ξ, $\Delta_y = (\overline{y} - \underline{y})$, $\Delta_w = \overline{w} - \underline{w}$, and $\Delta = [y(1, \alpha) - \overline{w}] - [y(\phi_0, \alpha) - \underline{w}] = \alpha(1 - \phi_0)\Delta_y - \Delta_w$. Then, $\Delta < \alpha(1 - \phi_0)\Delta_y$. Notice that (A1) implies that $\Delta > \phi_0(1 - \alpha)\Delta_y$.

Proposition 3. *Suppose that*

$$\underline{\phi}\gamma < \phi_0\alpha \left\{1 + \frac{(\gamma - \alpha)\Delta_y + \Delta + \phi_0\alpha(\gamma - \alpha)\Delta_y(T - 2)}{(1 + \phi_0\alpha)\Delta_y + \{\Delta + \phi_0\alpha[1 - (\gamma - \alpha)]\Delta_y\}(T - 2)}\right\}. \tag{5}$$

There exists $\underline{\delta}_f \in (0, 1)$ such that if $\delta_f \geq \underline{\delta}_f$, then in equilibrium an incumbent of age 2 who has failed to produce high output in his first period of employment is dismissed unless the firm is committed to make him an offer.

Note that $\phi_0\alpha$ is the probability that an age 1 worker produces high output, while $\underline{\phi}\gamma$ is an upper bound on the probability that an age 2 worker who produced low output in his first period of employment, produces high output. Proposition 3 is then not surprising when $\underline{\phi}\gamma \leq \phi_0\alpha$—in fact, $\underline{\delta}_f = 0$ in this case. However, as we will argue in the next subsection, we are interested in the case where $\underline{\phi}\gamma > \phi_0\alpha$. In this second case, the firm's flow payoff when it employs an age 2 worker who exerted effort in his first period of employment but failed to produce high output is greater than the firm's flow payoff when it employs an age 1 worker. So, the firm needs to be sufficiently patient for the result to hold.

A sketch of the proof of Proposition 3 helps us understand how condition (5) is derived. The complete proof is in Appendix A. Suppose, by contradiction, that there is an equilibrium σ with a history $\widehat{h}$ for the firm after which, even though not committed to do so, it offers $(w,(q,w'))$ to an age 2 worker who failed to produce high output in his first period of employment. Denote the worker who receives this offer by W. By Lemma 9, a worker of age 2 or more never exerts effort when employed. Hence, σ can be an equilibrium only if $q = 0$ and $w = \underline{w}$. Assume that this is the case and let $V = V(h|\sigma)$ for any history h for the firm after which it makes an offer to an age 1 worker. By (A2), W exerts effort in his first period of employment. So,

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

where $\Delta' = [y(1,\alpha) - \overline{w}] - [y(\underline{\phi},\gamma) - \underline{w}]$. Now observe that

$$
\begin{aligned}
V(\widehat{h}|\sigma) &= (1-\delta_f)[y(\underline{\phi},\gamma) - \underline{w}] + \underline{\phi}\gamma\left\{\delta_f(1-\delta_f^{T-2})[y(1,\alpha) - \overline{w}] + \delta_f^{T-1}V\right\} + (1-\underline{\phi}\gamma)\delta_f V \\
&= \delta_f V + (1-\delta_f)\underbrace{\left\{[y(\underline{\phi},\gamma) - \underline{w}] + \underline{\phi}\gamma\frac{\delta_f(1-\delta_f^{T-2})}{1-\delta_f}[y(1,\alpha) - \overline{w} - V]\right\}}_{\widetilde{V}}.
\end{aligned}
$$

We are done if $V > \widetilde{V}$ for δ_f sufficiently close to one—this implies the firm can profitably deviate after $\widehat{h}$ by replacing W with the available age 1 worker.[15] Straightforward algebra shows that (5) is necessary and sufficient for $V > \widetilde{V}$ when δ_f is sufficiently high.

---

[15]Since $V < y(1,\alpha) - \overline{w}$, $V(\widehat{h}|\sigma) \leq (1-\delta_f)[y(\phi_0,\alpha) - \underline{w}] + \phi_0\alpha\{\delta_f(1-\delta_f^{T-2})[y(1,\alpha) - \overline{w}] + \delta_f^{T-1}V\} + (1-\phi_0\alpha)\delta_f V$ when $\underline{\phi}\gamma \leq \phi_0\alpha$. It is easy to see that $V(\widehat{h}|\sigma) < V$ for all $\delta_f \in (0,1)$ in this case.

It follows from the previous paragraph that an increase in γ has two opposing effects. First, it increases V, the value to the firm from hiring an age 1 worker. Second, it increases $V(\widehat{h}|\sigma)$, the value to the firm from retaining an age 2 worker who exerts effort and produces low output when of age 1. Nevertheless, the second effect dominates the first, that is, increasing γ makes it more difficult for (5) to be satisfied. For this, notice that (5) is equivalent to

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

Now observe that we can rewrite the above inequality as

$$f(\gamma) = \gamma^2 A + \gamma B + C > 0,$$

where $A = (1-\alpha)\phi_0\alpha(T-2)\Delta_y$, $C = \alpha(1-\phi_0\alpha)\{[1-\alpha+\phi_0\alpha(T-1)]\Delta_y + (T-1)\Delta\}$, and

$$B = \alpha(1-\phi_0\alpha)\Delta_y - (1-\alpha)\left\{(1+\phi_0\alpha)\Delta_y + (T-2)[\Delta + \phi_0\alpha(1+\alpha)\Delta_y]\right\}.$$

Since $\Delta > \phi_0(1-\alpha)\Delta_y$ by (A1), we then have that

$$
\begin{aligned}
f'(1) &= \phi_0\alpha(1-\alpha)^2(T-2)\Delta_y - (1-\alpha)(T-2)\Delta + [\alpha(1-\phi_0\alpha) - (1-\alpha)(1+\phi_0\alpha)]\Delta_y \\
&< \left[\phi_0\alpha(1-\alpha)^2 - \phi_0(1-\alpha)^2\right](T-2)\Delta_y + \left[\phi_0\alpha - \underline{\phi}(1+\phi_0\alpha)\right]\frac{(1-\phi_0\alpha)\Delta_y}{\phi_0}.
\end{aligned}
$$

Our interest is in the case where $\underline{\phi}\gamma > \phi_0\alpha$, and a necessary condition for this is $\underline{\phi} > \phi_0\alpha$. Hence, $f'(1) < 0$, and so f is decreasing in the interval $(\alpha, 1)$. This establishes the desired result.

Observe finally that Proposition 3 holds when $\underline{\underline{\phi}}(\underline{e})\gamma < \phi_0\alpha$ even if (A3) is not satisfied. Indeed, $\underline{\underline{\phi}}(\underline{e})$ is the largest reputation possible for a worker of age 3 or more who has only produced low output. Therefore, when $\underline{\underline{\phi}}(\underline{e})\gamma < \phi_0\alpha$, the firm only hires an age $k \geq 3$ worker who has never produced high output if it is committed to do so—a proof of this result follows along the lines of the proof of Lemma 7. The above sketch of the proof of Proposition 3 shows that it is precisely this last fact that implies the result.

6.2 Commitment is Beneficial

Assume that the conditions of Proposition 3 are satisfied, so that commitment to employment is necessary for age 1 workers to exert effort. Recall that the firm benefits in two ways when a worker

of age 1 exerts effort. First, it gains additional output when this worker is of age 2. Second, the output of this worker when he is of age 2 is more informative about his ability. As discussed at the end of Section 3, the focus of our analysis is on the informational role of effort. Given our objective, we now derive conditions under which the use of probation is beneficial to the firm just for the informational gain from inducing an age 1 worker to exert effort.

In order to quantify the informational gain from inducing an age 1 worker to exert effort, consider first the situation in which the firm cannot offer commitment, that is, the firm is constrained to offer $q = 0$ to all workers. Let V_1 be the firm's payoff in this case. In light of the previous subsection, if $\underline{\phi}\gamma$ is not too large, then the firm always dismisses an age 1 worker who produces low output when it is patient enough. In fact, the conditions under which the firm dismisses such a worker are the same as those of Proposition 3—this follows from the proof of Proposition 3. Thus, V_1 satisfies

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

Solving the above equation for V_1 we obtain that

$$V_1 = \frac{(1 - \delta_f)[y(\phi_0, \alpha) - \underline{w}] + \phi_0\alpha\delta_f(1 - \delta_f^{T-1})[y(1, \alpha) - \overline{w}]}{1 - \delta_f + \phi_0\alpha\delta_f(1 - \delta_f^{T-1})}. \tag{6}$$

Notice that

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

where

$$\lambda_1 = \frac{1 - \delta_f}{1 - \delta_f + \phi_0\alpha\delta_f(1 - \delta_f^{T-1})}.$$

Consider now the case where the extra output a worker generates when he exerts effort is lost, so that the only gain from commitment is informational. Since the firm has always the option of offering $(\underline{w}, (1, \underline{w}))$ to an age 1 worker, it can obtain a payoff of at least V_2', where

$$V_2' = (1 - \delta_f)[y(\phi_0, \alpha) - \underline{w}] + \phi_0\alpha \left\{ \delta_f(1 - \delta_f^{T-1})[y(1, \alpha) - \overline{w}] + \delta_f^T V_2' \right\}$$
$$+ (1 - \phi_0\alpha)\left\{ \delta_f(1 - \delta_f)[y(\underline{\phi}, \alpha) - \underline{w}] + \underline{\phi}\gamma \left\{ \delta_f^2(1 - \delta_f^{T-2})[y(1, \alpha) - \overline{w}] + \delta_f^T V_2' \right\} + (1 - \underline{\phi}\gamma)\delta_f^2 V_2' \right\}.$$

In fact, Lemmas 8 and 9 hold in this case as well, so the firm still dismisses an incumbent of age 3 or more who has only produced low output if it is not committed to make an offer to him. Then,

21

by (A2), an age 1 worker exerts effort if the firm offers him $(\underline{w}, (1, \underline{w}))$. Solving the above equation for V_2' we obtain

$$
\begin{aligned}
V_2' &= \frac{(1-\delta_f)[y(\phi_0,\alpha)-\underline{w}]+\delta_f(1-\delta_f)\{\phi_0\alpha[y(1,\alpha)-\overline{w}]+(1-\phi_0\alpha)[y(\underline{\phi},\alpha)-\underline{w}]\}}{1-\delta_f^2+\phi_0[\alpha+\gamma(1-\alpha)]\delta_f^2(1-\delta_f^{T-2})} \\
&\qquad +\frac{\phi_0[\alpha+\gamma(1-\alpha)]\delta_f^2(1-\delta_f^{T-2})[y(1,\alpha)-\overline{w}]}{1-\delta_f^2+\phi_0[\alpha+\gamma(1-\alpha)]\delta_f^2(1-\delta_f^{T-2})} \\
&= \frac{(1-\delta_f)[y(\phi_0,\alpha)-\underline{w}]+[\delta_f(1-\delta_f)+\phi_0[\alpha+\gamma(1-\alpha)]\delta_f^2(1-\delta_f^{T-2})][y(1,\alpha)-\overline{w}]}{1-\delta_f^2+\phi_0[\alpha+\gamma(1-\alpha)]\delta_f^2(1-\delta_f^{T-2})} \\
&\qquad -\frac{(1-\phi_0\alpha)\delta_f(1-\delta_f)\{[y(1,\alpha)-\overline{w}]-[y(\underline{\phi},\alpha)-\underline{w}]\}}{1-\delta_f^2+\phi_0[\alpha+\gamma(1-\alpha)]\delta_f^2(1-\delta_f^{T-2})}. \quad (7)
\end{aligned}
$$

Notice that

$$
V_2' = \lambda_2[y(\phi_0,\alpha)-\underline{w}]+(1-\lambda_2)[y(1,\alpha)-\overline{w}]-\delta_f\lambda_2(1-\phi_0\alpha)[\Delta+y(\phi_0,\alpha)-y(\underline{\phi},\alpha)],
$$

where

$$
\lambda_2 = \frac{1-\delta_f}{1-\delta_f+\delta_f(1-\delta_f)+\phi_0[\alpha+\gamma(1-\alpha)]\delta_f^2(1-\delta_f^{T-2})}.
$$

By construction, a sufficient condition for the informational gain from commitment to be positive is then $V_2' > V_1$. Moreover, by Lemma 9, the firm does not gain from offering more than two periods of probation to an age 1 worker. Hence, $V_2' > V_1$ is also necessary for the informational gain of commitment to be positive. Now observe that

$$
\begin{aligned}
V_2' - V_1 &= [\lambda_1-\lambda_2(1+\delta_f(1-\phi_0\alpha))]\Delta-\delta_f\lambda_2(1-\phi_0\alpha)[y(\phi_0,\alpha)-y(\underline{\phi},\alpha)] \\
&= \{\lambda_1-\lambda_2[1+\delta_f(1-\phi_0\alpha)(1+\psi)]\}\Delta,
\end{aligned}
$$

where $\psi = [y(\phi_0,\alpha)-y(\underline{\phi},\alpha)]/\Delta$, and that

$$
\begin{aligned}
(1-\delta_f)^{-1}\{\lambda_1-\lambda_2[1+\delta_f(1-\phi_0\alpha)(1+\psi)]\} &= \\
\phi_0\gamma(1-\alpha)\delta_f^2(1-\delta_f^{T-2})-\delta_f(1-\delta_f)(1-\phi_0\alpha)\psi &- \phi_0\alpha(1-\phi_0\alpha)(1+\psi)\delta_f^2(1-\delta_f^{T-1}).
\end{aligned}
$$

So, a necessary and sufficient condition for the gain in information to be beneficial for the firm is

$$
\frac{\phi_0(1-\alpha)}{1-\phi_0\alpha}\gamma = \underline{\phi}\gamma > \phi_0\alpha(1+\psi)\frac{1-\delta_f^{T-1}}{1-\delta_f^{T-2}}+\psi\frac{(1-\delta_f)}{\delta_f(1-\delta_f^{T-2})}. \quad (8)
$$

22

The term $(1-\delta_f^{T-1})/(1-\delta_f^{T-2})$ in (8) reflects the fact that even if the output of an age 2 worker is more informative about his ability than the output of an age 1 worker, the former lives for one less period than the latter. This difference in lifetimes is more important the more patient the firm is, but it decreases with the lifetime T of the workers—the ratio $(1 - \delta_f^{T-1})/(1 - \delta_f^{T-2})$ increases with δ_f and decreases with T. The term $y(\phi_0, \alpha) - y(\underline{\phi}, \alpha) = \psi\Delta$ is the loss in flow payoff to the firm when it retains an incumbent of age 2 who failed to produce high output if the extra output a worker generates when he exerts effort is lost. For a fixed δ_f, this loss is less important for the firm the greater T is—both terms on the right–hand side of (8) decrease with T. In fact, the benefit to the firm from being able to identify the ability of the workers increases with their lifetime.

6.3 Probation

Condition (8) reduces to

$$\underline{\phi}\gamma > \phi_0\alpha(1 + \psi)\frac{T - 1}{T - 2} + \frac{\psi}{T - 2} \tag{9}$$

when δ_f converges to one. There is a natural tension between (5) and (9). We need $\underline{\phi}\gamma > \phi_0\alpha$ for the informational gain from commitment to be positive. The difference between $\underline{\phi}\gamma$ and $\phi_0\alpha$ cannot be too large, though, for otherwise the firm would retain an age 2 worker who exerted effort and produced low output in his first period of employment regardless of its discount factor, rendering commitment unnecessary. Nevertheless, it is possible to show that (5) and (9) can both be satisfied when T is large; we know from above that T large helps the informational gain from commitment be positive. The details of the argument are in Appendix B.

Proposition 4. *Suppose that*

$$\phi_0\alpha(1 + \psi) < \underline{\phi}\gamma < \phi_0\alpha\left\{1 + \frac{\phi_0\alpha(\gamma - \alpha)}{\Delta + \phi_0\alpha[1 - (\gamma - \alpha)]\Delta_y}\right\}.$$

There exists $\underline{T} \geq 3$ with the property that if $T \geq \underline{T}$, then there exists $\underline{\delta}_f \in (0, 1)$ such that if $\delta_f \geq \underline{\delta}_f$, then: (i) the firm always offers probation to age 1 workers in equilibrium; (ii) the informational gain from commitment is positive.

Proof: By assumption, there exists $\underline{T}$ such that both (5) and (9) are satisfied when $T \geq \underline{T}$. Suppose then $T \geq \underline{T}$. This implies there exists $\underline{\delta}_f \in (0, 1)$ with the property that if $\delta_f \geq \underline{\delta}_f$, then: (i) in

23

equilibrium, the firm dismisses an incumbent of age $k \geq 2$ who has never produced high output if it is not committed to employ him; (ii) the informational gain from commitment is positive. Suppose now, by contradiction, that there is an equilibrium σ with a history $\widehat{h}$ for the firm after which it offers $(\underline{w}, 0)$ to the available age 1 worker. By Lemma 5, the firm's equilibrium payoff is $V(\widehat{h}|\sigma) = V_1$ given by (6). The firm, however, can always offer $(\underline{w}, (1, \underline{w}))$ to an age 1 worker and obtain a payoff greater than V_2' given by (7), a contradiction. $\qquad\qquad\qquad\square$

Proposition 4 is silent as to how many periods of probation the firm offers to age 1 workers. By Lemma 9, however, a worker of age 2 or more never exerts effort when employed. So, relative to offering two periods of probation to an age 1 worker, the firm cannot gain by offering more than two periods of employment, since this does not affect the behavior of a worker when he is of age 2 or more. This implies, by a straightforward argument, that under the conditions of Proposition 4 the firm offers always offers $(\underline{w}, (1, \underline{w}))$ to age 1 workers.

One limitation of the analysis in the non–IID case is that there is no scope for more than two periods of probation. This result is in part driven by the particular production technology we consider. Specifically, by the assumption that there is a one–period delay in the effect of effort on output. Our analysis, in principle, can allow for more than two periods of probation as an (unique) equilibrium outcome if, for instance, it takes more than one period of effort before effort has an impact on output.

7 Output–Contingent Contracts

Potentially there are several reasons why output–contingent contracts can fail to be feasible. One possibility is that a third party responsible for the enforcement of contracts cannot distinguish between high and low output. Output is of a very specialized nature in many of the environments where probation is used. Another possibility is that the contingencies determining the level of output cannot be described ex–ante.[16] We are nevertheless interested in understanding the extent to which our results are driven by the assumption that explicit output–contingent incentives are not possible. In this section we address this issue by considering the case in which the firm has full commitment power, and so can offer long–term output–contingent contracts.

---

[16]Notice that the results of Maskin and Tirole (1999) on incomplete contracts do not apply to our setting.

Let $Y = \{\underline{y}, \overline{y}\}$ be the output space and denote a typical element of Y^t, with $t \leq T$, by y^t. The set Y^t is the set of possible end–of–period output histories of length t; $y^t = (y_1, \ldots, y_t) \in Y^t$ is the output history of a worker who produces y_s in his sth period of employment, with $s \leq t$. A *wage policy* is a list $\omega = \{\omega_t\}_{t=1}^T$, with $\omega_t : Y^t \to \mathbb{R}$, where $\omega_t(y^t)$ is the wage the firm pays to a worker if his output history is y^t. A *retention policy* is a list $\tau = \{\tau_t\}_{t=2}^T$, with $\tau_t : Y^{t-1} \to [0, 1]$, where $\tau_t(y^{t-1})$ is the probability the firm offers employment to a worker with output history y^{t-1}.[17] Recall that wages are paid at the end of a period, while retention decisions are made at the beginning of a period. For any $k \leq t$, let $\chi_k^t : Y^t \to Y^k$ be such that $\chi_k^t(y^t)$ is the projection of y^t into its first k coordinates. A retention policy is *consistent* if $\tau_t(y^{t-1}) > 0$ is possible only when $\tau_k(\chi_{k-1}^{t-1}(y^{t-1})) > 0$ for all $k \in \{2, \ldots, t-1\}$. A long–term (output–contingent) contract is a pair (ω, τ) where ω is a wage policy and τ is a consistent retention policy. The firm offers long–term contracts to age 1 workers.

The timing of events in a period is as follows. Suppose the firm must offer employment to its incumbent according to the long–term contract (ω, τ) in place—this is the contract the firm offered to the incumbent when he was of age 1. In this case, the incumbent first decides whether to work for the firm or not. If he accepts employment, he then chooses whether to exert effort or not, output is realized, and the firm pays him according to ω. Note that we rule out implicit bonus payments.[18] If he rejects employment, both him and the firm collect their respective outside options. Suppose now the firm either has no incumbent or has an incumbent that it must dismiss according to the long–term contract in place. In these circumstances, the firm first decides whether it offers a long–term contract to the available age 1 worker or not. If it does, then the timing is as in the first case. If it does not, then both the firm and the available age 1 worker collect their respective outside options.

We still consider worker–symmetric perfect Bayesian equilibria. For simplicity, we assume that a worker who is known to be of the high type always accepts an offer by the firm and always exerts

<hr>

[17] Our results do not change if we allow for random wage policies. In fact, since workers are risk–averse, it is possible to show that the firm never offers a random wage policy in equilibrium. We restrict attention to deterministic wage policies for ease of notation.

[18] This assumption is without any loss. Under full commitment, any outcome that involves bonus payments can be replicated without the use of bonus payments by changing the long–term contracts offered in an appropriate way.

effort if indifferent between exerting effort or not.[19] As before, the firm's continuation value when it hires an age 1 worker is independent of calendar time. The proof of this fact is identical to the proof of Lemma 5 in Section 4.

A straightforward consequence of (A1) and the fact that the firm's continuation payoff when it hires an age 1 worker is independent of calendar time is that the firm always employs an age 2 worker who produced high output in his first period of employment.

Lemma 10. *In equilibrium, the firm always offers (ω, τ) with $\tau_2(\overline{y}) = 1$ to an age 1 worker and an age 2 worker who is known to be of the high type always works for the firm.*

Motivated by Lemma 10, we say that the firm offers probation to an age 1 worker if it offers him a long–term contract (ω, τ) with $\tau_2(\underline{y}) > 0$. As discussed in Section 3, our emphasis is on the informational role of effort. So, we assume that choosing no effort is optimal for a worker known to be of the high type in both the IID case and the non–IID case. We provide conditions for this to hold below. Thus, as in the case without full commitment, effort can only be beneficial for the firm if it is exerted by a worker who has yet to prove his talent. The difference is that now the firm can supplement career concerns motives for effort exertion with explicit incentives.

We assume in this section that v is continuously differentiable. Given the supplementary nature of this section, the analysis will be somewhat terse.

7.1   IID Case

Consider the static principal–agent problem involving the firm and a worker of the high type and let $R(e)$ denote the highest payoff possible for the firm when the worker's choice of effort is e. It is straightforward to show that $R(e) = r(\overline{w}, e)$, where $r(w, \underline{e}) = y(1, \alpha) - w$ and

$$r(w, \overline{e}) = y(1, \alpha) + \eta\Delta_y - (\alpha + \eta)v^{-1}\left(v(w) + \frac{(1-\alpha)c}{\eta}\right) - (1 - \alpha - \eta)v^{-1}\left(v(w) - \frac{\alpha c}{\eta}\right).$$

Since v^{-1} is convex, $r(w, \overline{e}) \leq G(w)$, where $G(w) = y(1, \alpha) + \eta\Delta_y - v^{-1}(v(w) + c)$. We then make the following assumption:

(A4) $v^{-1}(v(\overline{w}) + c) - \overline{w} > \eta\Delta_y.$

[19]This is true in equilibrium since the firm can break ties in incentive–compatibility and participation constraints by making infinitesimal changes in the wage policy.

Condition (A4) has a natural interpretation. It implies that the lowest increase in the expected wage payment necessary to compensate a worker of the high type for exerting effort is smaller than the expected increase in output when such a worker exerts effort. If $v(w) = w$, (A4) reduces to $\eta \Delta_y < c$. A consequence of Assumption (A4) is that the firm never offers a long–term contract that induces an incumbent known to be of the high type to exert effort when employed.

Lemma 11. *In equilibrium, the firm pays a worker known to be of the high type his outside option and such a worker never exerts effort.*

An immediate implication of Lemma 11 and (A1) is that the firm always makes an offer to an incumbent it knows is of high ability. Let Y_H be the subset of $\bigcup_{t=1}^{T} Y^t$ such that $y^k = (y_1, \ldots, y_k)$ belongs to Y_H if, and only if, at least one element of $\{y_1, \ldots, y_k\}$ is equal to $\overline{y}$. Moreover, for any $y^k \in Y^k$ and $y^{k'} \in Y^{k'}$, let $y^k y^{k'}$ denote the output history y^k followed by the output history $y^{k'}$.

Lemma 12. *In equilibrium, the firm only offers (ω, τ) such that if $y^{t-1} \in Y_H$ and $\tau_t(y^{t-1}) > 0$, then $\tau_{t+k}(y^{t-1}y^k) = 1$ for all $y^k \in Y^k$ with $1 \leq k \leq T - t$.*

Suppose now that (1) holds, that is,

$$\phi_0 \eta \delta_w (1 - \delta_w^{T-1})[v(\overline{w}) - v(\underline{w})] \leq (1 - \delta_w)c.$$

We know from the analysis of Section 5 that unless (1) holds with equality, the firm needs to resort to explicit incentives if it wishes an age 1 worker to exert effort. The next result shows, as in the case without full commitment, that it is never optimal for the firm to retain an age 1 worker who produces low output.

Proposition 5. *In equilibrium, the firm never offers probation to age 1 workers when (1) holds.*

Suppose instead that (2) holds, that is,

$$\phi_0 \eta \delta_w (1 - \delta_w^{T-1})[v(\overline{w}) - v(\underline{w})] > (1 - \delta_w)c.$$

This implies that concerns for his future reputation are sufficient to induce an age 1 worker who is dismissed after low output to exert effort. With full commitment we are able to establish a stronger result than Proposition 2 in Section 5. For this, let (ω^*, τ^*) be the long–term contract where: (i)

27

$\tau_2^*(\underline{y}) = 0$; (ii) $\tau_{k+1}^*(\overline{y}y^{k-1}) = 1$ for all $y^{k-1} \in Y^{k-1}$ with $1 \le k \le T - 1$, with the convention that Y^0 is the empty set; (iii) $\omega_1(y) \equiv \underline{w}$; (iv) $\omega_{k+1}(\overline{y}y^{k-1}) = \overline{w}$ for all $y^k \in Y^k$ with $1 \le k \le T - 1$. By construction, the firm obtains a lifetime payoff of V^* when it offers (ω^*, τ^*) to age 1 workers, where V^* is the payoff in the equilibrium σ^* of Proposition 2. Now observe, by Lemma 11 and the proof of Proposition 2, that V^* is still the highest payoff the firm can obtain in any equilibrium. Thus, the strategy profile where the firm offers (ω^*, τ^*) to age 1 workers is an equilibrium. Moreover, from the proof of Proposition 2, there can be no equilibrium where firm offers a long–term contract with $\tau_2(\underline{y}) > 0$. We have thus established the following result.

Proposition 6. *In equilibrium, the firm never offers probation to age 1 workers when (2) holds.*

7.2   Non-IID Case

A key element in the analysis of the non–IID case is that it is necessary for the firm to retain an age 1 worker after he produces low output in order to induce him to exert effort. This is no longer the case when long–term contracts are possible, as the firm can now induce an age 1 worker to exert effort even if he is dismissed after low output. This is because the firm can promise to reward the worker if he produces high output when he is of age 2.

We make the following assumption, which is the counterpart of (A4) to the non–IID case and has a similar interpretation. The reason the discount factor of the worker appears in (A5) is that an explicit contract can compensate a worker of the high type for his effort only in the period after he exerts it—in the non–IID case effort only affects future output.

(A5) $v^{-1}(v(\overline{w}) + c/\delta_w) - \overline{w} > (\gamma - \alpha)\Delta_y$.

Assumption (A5) implies that Lemmas 11 and 12 also hold in the non–IID case. The proof of Lemma 12 is identical.

Suppose, by contradiction, that there exists an equilibrium in the non–IID case where the firm does not offer probation to an age 1 worker. By Lemma 12, an upper bound for the equilibrium payoff V to the firm is

$$V^+ = \frac{(1 - \delta_f)[y(\phi_0, \alpha) - \underline{w}] + \phi_0 \alpha \delta_f \left\{ (1 - \delta_f)[y(1, \gamma) - \overline{w}] + \delta_f(1 - \delta_f^{T-2})[y(1, \alpha) - \overline{w}] \right\}}{1 - \delta_f + \phi_0 \alpha \delta_f(1 - \delta_f^{T-1})}. \quad (10)$$

Consider now the deviation for the firm where it offers (ω^*, τ^*) such that: (i) $\omega_1(y) \equiv \underline{w}$; (ii) $\omega_t(\overline{y}y^{t-1}) = \overline{w}$ for all $y^{t-1} \in Y^{t-1}$ with $t \geq 2$; (iii) $\omega_2(yy^1) \equiv \underline{w}$; (iv) $\tau_2(y) = 1$; (v) $\tau_3(yy) = 0$; (vi) $\tau_t(\overline{y}y^{t-2}) = 1$ for all $y^{t-2} \in Y^{t-2}$ with $t \geq 2$; and (vii) $\tau_t(y\overline{y}y^{t-3}) = 1$ for all $y^{t-3} \in Y^{t-3}$ with $t \geq 3$. By (A2), the payoff to the firm from this deviation is

$$V_2 = \frac{(1-\delta_f)[y(\phi_0,\alpha) - \underline{w}] + \phi_0\alpha\delta_f(1-\delta_f)[y(1,\gamma) - \overline{w}] + (1-\phi_0\alpha)\delta_f(1-\delta_f)[y(\underline{\phi},\gamma) - \underline{w}]}{1 - \delta_f^2 + \phi_0(\alpha + \gamma(1-\alpha))\delta_f^2(1-\delta_f^{T-1})}$$
$$+ \frac{\phi_0(\alpha + \gamma(1-\alpha))\delta_f^2(1-\delta_f^{T-2})[y(1,\alpha) - \overline{w}]}{1 - \delta_f^2 + \phi_0(\alpha + \gamma(1-\alpha))\delta_f^2(1-\delta_f^{T-1})}. \quad (11)$$

Now observe that

$$\lim_{\delta_f \to 1} V^+ = [y(1,\alpha) - \overline{w}] - \frac{\Delta - \phi_0\alpha(\gamma-\alpha)\Delta_y}{1 + \phi_0\alpha(T-1)}$$

and that

$$\lim_{\delta_f \to 1} V_2 = [y(1,\alpha) - \overline{w}] - \frac{\Delta - \phi_0\alpha(\gamma-\alpha)\Delta_y + (1-\phi_0\alpha)\Delta'}{2 + \phi_0(\alpha + \gamma(1-\alpha))(T-2)},$$

where, as before, $\Delta = [y(1,\alpha) - \overline{w}] - [y(\phi_0,\alpha) - \underline{w}]$ and $\Delta' = [y(1,\alpha) - \overline{w}] - [y(\underline{\phi},\gamma) - \underline{w}]$. Hence, $V_2 > V^+$ if, and only if,

$$\frac{\Delta - \phi_0\alpha(\gamma-\alpha)\Delta_y}{1 + \phi_0\alpha(T-1)} \geq \frac{\Delta - \phi_0\alpha(\gamma-\alpha)\Delta_y + (1-\phi_0\alpha)\Delta'}{2 + \phi_0(\alpha + \gamma(1-\alpha))(T-2)}. \quad (12)$$

Since $[2 + \phi_0(\alpha + \gamma(1-\alpha))(T-2)] - [1 + \phi_0\alpha(T-1)] = 1 - \phi_0\alpha + \phi_0\gamma(1-\alpha)(T-2)$, straightforward algebra shows that (12) is equivalent to

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

where we used the fact that $\Delta' = \Delta + (\phi_0\alpha - \underline{\phi}\gamma)\Delta_y$. We thus have the following result, which is an analogue to Proposition 4 for the full–commitment case.

Proposition 7. *Suppose that $\underline{\phi}\gamma - \phi_0\alpha > \phi_0\alpha(\gamma-\alpha)$. There exists $\underline{T}$ such that if $T \geq \underline{T}$, then in equilibrium the firm always offers probation to age 1 workers if it is patient enough.*

8 Conclusion

This paper provides a rationale for the use of short–term commitment to employment in markets where individual talent is uncertain. A firm can gain from offering probation to workers of un-known ability if this commitment encourages them to invest in their reputation, thus increasing

the informativeness of their performance and so helping the firm identify their talent. We show that probation can only be beneficial if the effect of effort on performance is persistent, otherwise probation never increases, and may actually decrease, the incentives of workers to invest in their reputation. More precisely, we show that the firm can gain from offering probation when the impact of effort on output is persistent *but* delayed. The reason for this is that probation solves a time–consistency problem when it takes time for effort to affect output. In the absence of commitment, the firm cannot credibly promise to retain a worker of uncertain talent whose initial performance is poor, and this undermines his incentives to exert effort. We also show that the use of explicit output–contingent contracts cannot substitute for the use of probation.

References

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

[2] Bar–Isaac, H. (2007): "Something to Prove: Reputation in Teams", *RAND Journal of Economics*, 38(2), 495-511.

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

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

[5] Che, Y., and J. Sakovics (2004): "Contractual Remedies to the Holdup Problem: A Dynamic Perspective", mimeo.

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

[7] Eeckhout, J. (2006): "Employer Learning and General Human Capital", mimeo.

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

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

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

[11] Jarque, A. (2005):"Repeated Moral Hazard with Effort Persistence", mimeo.

[12] Jeitschko, T.D., and L.J. Mirman (2002a): "Information and Experimentation in Short-Term Contracting", *Economic Theory*, 19(2), 311-331.

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

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

[15] Klein, B., R. Crawford, and A. Alchian (1978): "Vertical Integration, Appropriable Rents, and the Competitive Contracting Process", *Journal of Law and Economics*, 21(2), 297-326.

[16] Maskin, E., and J. Tirole (1999): "Unforseen Contingencies and Incomplet Contracts", *Review of Economic Studies*, 66(1), 83-114.

[17] O'Flaherty, B., and A. Siow (1995): "Up or Out Rules in the Market for Lawyers", *Journal of Labor Economics*, 13(4), 709-735.

[18] 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.

[19] Osband, K. (1989): "Optimal Forecasting Incentives", *Journal of Political Economy*, 97(5), 1091-1112.

[20] Laffont, J.J., and J. Tirole (1988): "The Dynamics of Incentive Contracts", *Econometrica*, 56(5), 1153-1175.

[21] Levin, J. (2003): "Relational Incentive Contracts", *American Economic Review*, 93(3), 835-857.

[22] Manso, G. (2007): "Motivating Innovation", mimeo.

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

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

Appendix A: Omitted Proofs

Proof of Lemma 2: Suppose, by contradiction, than an age T incumbent who has always produced low output rejects an offer of $(\underline{w}, 0)$. The lifetime payoff to the firm after this is $(1 - \delta_f)\Pi + \delta_f V$, where V is its continuation payoff from next period on. Consider then the following deviation for the firm: (i) offer $(w, 0)$, with $w > \underline{w}$, to the age T worker and pay him w regardless of his performance; (ii) behave from next period on as if no deviation has occurred. The firm's continuation payoff after this deviation is at least $(1 - \delta_f)[\underline{y} - w] + \delta_f V$. Since $\underline{y} - \underline{w} > \Pi$ by assumption, this deviation is profitable as long as w is sufficiently close to $\underline{w}$, a contradiction. The other part of this lemma follows from a similar argument and the fact that $\alpha \overline{y} + (1 - \alpha)\underline{y} - \overline{w} > \underline{y} - \underline{w}$. $\qquad \square$

Proof of Lemma 3: Suppose the firm is not committed to make an offer to its incumbent or has no incumbent. Since, by Lemma 1, an age 1 worker accepts $(\underline{w}, 0)$, the same argument used in the proof of Lemma 2 shows that the firm makes an offer that is not rejected. Suppose now the firm is committed to make an offer to its incumbent. By Lemmas 1 and 2, we only need to consider the case in which the incumbent is known to be of the high type and the lowest one–period wage the firm can offer is less than $\overline{w}$. Once more, the same argument used in the proof of Lemma 2 shows that the firm makes an offer that is accepted. $\qquad \square$

Proof of Lemma 4: Denote the incumbent by W, let $k \in \{2, \ldots, T\}$ be his age, and let $\ell \in \{0, \ldots, T - k\}$ be the maximum number of future periods that the firm employs W if it makes him an offer that he accepts. The proof is by induction in ℓ.

We know from the main text that if $\ell = 0$, then: (i) the firm never W offers a one–period wage greater than $\max\{\overline{w}, w'\}$; (ii) if the firm makes W an offer, then it never commits to future one–period wage offers greater than $\overline{w}$ (trivially satisfied); (iii) if W accepts employment, then the firm pays him the one–period wage it offers; (iv) W does not exert effort. We also know that if the firm offers W a one–period wage of $\overline{w}$, then he accepts the offer (regardless of k and ℓ). In particular, W will never punish the firm for a deviation by rejecting a one–period wage offer of $\overline{w}$. Denote this last fact by (v).

Suppose, by induction, that there exists $\ell' \in \{0, \ldots, T-k\}$ such that (i) to (iv) hold if $\ell \leq \ell'$ and let $\ell = \ell' + 1$. We claim that (iii) is true. Suppose not and let w be the one–period wage the firm offers W. Consider the following deviation for the firm: pay w to W regardless of his output and behave from next period on as if no deviation has occurred. This is a profitable deviation for the firm since its continuation payoff after the deviation is the same by (v) and the induction hypothesis. Thus, (iv) is also true, since W'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 could profitably deviate either by lowering its one–period wage offer to W or by lowering the future one–period wage offers that it promises to W. $\qquad\square$

Proof of Lemma 5: 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' happened. Since workers follow symmetric strategies, this deviation increases the firm's payoff after h by $V(h'|\sigma) - V(h|\sigma)$, a contradiction. $\qquad\square$

Proof of Lemma 6: Let σ be an equilibrium. By Lemma 4, an incumbent known to be of high type never exerts effort when employed. Moreover, such a worker always rejects an offer $(w, (q, w'))$ with $w < \underline{w}$. Hence, by (A1), $V(h|\sigma) < \alpha \overline{y} + (1-\alpha)\underline{y} - \overline{w}$ if h is the initial history of the game. Lemma 5 then implies that $V(h'|\sigma) < \alpha \overline{y} + (1-\alpha)\underline{y} - \overline{w}$ for every history h' for the firm after which it hires the available age 1 worker. The desired result now follows from Lemma 2. $\qquad\square$

Proof of Lemma 7: Suppose, by contradiction, that there is an equilibrium σ with a history $\widehat{h}$ for the firm after which, despite not committed to do so, it offers $(w, (q, w'))$, with $q \geq 0$, to an incumbent of age $k \geq 2$ who has only produced low output and will never exert effort. Denote this worker by W and assume, without loss, that he is dismissed when he is of age $k + q + 1$ if he has not revealed himself to be of the high type by then. Note that if σ is to be an equilibrium, then it must be that: (i) $w = w' = \underline{w}$; (ii) the firm always offers the lowest one–period wage possible to W; (iii) the firm never pays W a bonus. Let $\phi < \phi_0$ be W's reputation, $y(1, \alpha) = \alpha \overline{y} + (1-\alpha)\underline{y}$, and $V < y(1, \alpha) - \overline{w}$ be the continuation payoff to the firm after it hires an age 1 worker. Then,

by Lemma 6, $V(\widehat{h}|\sigma) = V(\phi)$, where

$$V(\phi) \;=\; \phi\sum_{j=0}^{q}(1-\alpha)^j\alpha\left\{(1-\delta_f^j)[\underline{y}-\underline{w}]+\delta_f^j(1-\delta_f)[\overline{y}-\underline{w}]+\delta_f^{j+1}(1-\delta_f^{T-k-j})[y(1,\alpha)-\overline{w}]+\right.$$
$$\left.\delta_f^{T-k+1}V\right\} + \left[\phi(1-\alpha)^{q+1}+1-\phi\right]\left\{(1-\delta_f^{q+1})[\underline{y}-\underline{w}]+\delta_f^{q+1}V\right\}.$$

Since $\phi\sum_{j=0}^q(1-\alpha)^j\alpha = 1-[\phi(1-\alpha)^{q+1}+1-\phi]$, we can rewrite $V(\phi)$ as

$$V(\phi) \;=\; \phi\sum_{j=0}^{q}(1-\alpha)^j\alpha\left\{\delta_f^j(1-\delta_f)[\overline{y}-\underline{y}]+\delta_f^{j+1}(1-\delta_f^{q-j})\left\{[y(1,\alpha)-\overline{w}]-[\underline{y}-\underline{w}]\right\}\right.$$
$$\left.\delta_f^{q+1}(1-\delta^{T-k-q})[y(1,\alpha)-\overline{w}-V]\right\} + (1-\delta_f^{q+1})[\underline{y}-\underline{w}]+\delta_f^{q+1}V,$$

from which it follows that $V(\phi)$ is strictly increasing in ϕ. Consider now the following deviation for the firm after $\widehat{h}$: offer $(\underline{w},(q,\underline{w}))$ to the available age 1 worker and then behave as if no deviation has occurred—in particular, the firm treats this new worker as if he were W. The firm's continuation payoff after this deviation is at least $V(\phi_0)$—at worst for the firm, the new worker never exerts effort. This implies the firm has a profitable deviation, a contradiction. $\qquad\square$

Proof of Proposition 1: Consider first a worker of age $k \in \{2,\dots,T-1\}$ who has only produced low output. Denote this worker by W, let $\pi < \phi_0$ be his private belief that he is of the high type, and suppose the firm employs him. We claim that W does not exert effort and that the firm pays W the one–period wage it offers him. For this, let $\ell \le T-k+1$ be the maximum number of periods (including the present one) that W is employed if he never produces high output. We proceed by induction in ℓ.

(1) Suppose $\ell = 1$, that is, W is dismissed after low output. In particular, the firm is not committed to make W an offer in the next period and does not offer W commitment in the present period. Then, by Lemmas 4 and 6, W's incentive–compatibility constraint for effort exertion is

$$-c+\pi(\alpha+\eta)\left\{v(w^{\overline{y}})+\delta_w(1-\delta_w)^{-1}(1-\delta_w^{T-k})v(\overline{w})\right\}$$
$$+[1-\pi(\alpha+\eta)]\left\{v(w^{\underline{y}})+\delta_w(1-\delta_w)^{-1}(1-\delta_w^{T-k})v(\underline{w})\right\}$$
$$\ge \pi\alpha\left\{v(w^{\overline{y}})+\delta_w(1-\delta_w)^{-1}(1-\delta_w^{T-k})v(\overline{w})\right\}$$
$$+(1-\pi\alpha)\left\{v(w^{\underline{y}})+\delta_w(1-\delta_w)^{-1}(1-\delta_w^{T-k})v(\underline{w})\right\}, \qquad (13)$$

where w^y is the wage the firm pays this worker after he produces y. By Corollary 1, $w^{\overline{y}} = w$, the one–period wage the firm offers W. Since $\ell = 1$, $w^{\underline{y}} = w$ as well. Hence, we can rewrite (13) as

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

which is not satisfied by assumption. Thus, W does not exert effort.

(2) Suppose, by induction, that there exists $\ell' \leq T - k + 1$ such that if $\ell \leq \ell'$, then W does not exert effort and the firm pays W the one–period wage it offers him. Now let $\ell = \ell' + 1$. We know from the main text that W's incentive–compatibility constraint for effort exertion is given by

$$\pi\eta\underbrace{[v(w^{\overline{y}}) - v(w^{\underline{y}})]}_{\Delta_0} + \pi\eta\delta_w[R(\overline{y}, \overline{e}|\pi, k) - R(\underline{y}, \overline{e}|\pi, k)] + \delta_w(1 - \pi\alpha)[R(\underline{y}, \overline{e}|\pi, k) - R(\underline{y}, \underline{e}|\pi, k)] \geq c,$$

(14)

where w^y has the same interpretation as in Step 1 and $R(y, e|\pi, k)$ is W's continuation payoff after he chooses e and produces y, which depends on π and k.[20]

Lemma 7 and the induction hypothesis imply that the firm must be committed to employ W for the next ℓ' periods. Let $\widetilde{w} \geq \underline{w}$ be the lowest one–period wage the firm must offer W during this period of time. Then, by Lemmas 4 and 6,

$$R(\overline{y}, e|\pi, k) = (1 - \delta_w)^{-1}(1 - \delta_w^{\ell'})v(\max\{\overline{w}, \widetilde{w}\}) + \delta_w^{\ell'}(1 - \delta_w)^{-1}(1 - \delta_w^{T-k-\ell'})v(\overline{w}).$$

Now let $\underline{\pi}(e) = [1 - \xi(e)]\pi/[1 - \pi\xi(e)]$, where $\xi(\underline{e}) = \alpha$ and $\xi(\overline{e}) = \alpha + \eta$. By the induction hypothesis, W never exerts effort after producing low output. So, the firm has no incentive to offer him more than $\widetilde{w}$ as long as he does not produce high output. Using Lemmas 4 and 6 one more time, we have that

$$R(s) = (1 - \delta_w)^{-1}(1 - \delta_w^{s+1})v(\widetilde{w}) + \delta_w^{s+1}(1 - \delta_w)^{-1}(1 - \delta_w^{\ell'-s})v(\max\{\overline{w}, \widetilde{w}\})$$
$$+ \delta_w^{\ell'}(1 - \delta_w)^{-1}(1 - \delta_w^{T-k-\ell'})v(\overline{w})$$

is W's continuation payoff after producing low output (and receiving $w^{\underline{y}}$) if he produces high output for the first time after $s \in \{0, \ldots, \ell' - 1\}$ periods, and

$$R(\ell') = (1 - \delta_w)^{-1}(1 - \delta_w^{\ell'})v(\widetilde{w}) + \delta_w^{\ell'}(1 - \delta_w)^{-1}(1 - \delta_w^{T-k-\ell'})v(\underline{w})$$

---

[20]We show below that $R(y, e)$ does not depend on W's reputation.

36

is W's continuation payoff after low output if he never produces high output afterwards. Thus,

$$
\begin{aligned}
R(\underline{y}, e | \pi, k) &= \underline{\pi}(e) \sum_{s=0}^{\ell'-1} (1-\alpha)^s \alpha R(s) + \left[\underline{\pi}(e)(1-\alpha)^{\ell'} + 1 - \underline{\pi}(e)\right] R(\ell') \\
&= \underline{\pi}(e) \sum_{s=0}^{\ell'-1} (1-\alpha)^s \alpha [R(s) - R(\ell')] + R(\ell').
\end{aligned}
$$

Notice that $R(s) > R(\ell')$ for all $s \leq \ell' - 1$. Since $\underline{\pi}(\bar{e}) > \underline{\pi}(\underline{e})$, we then have that $R(\underline{y}, \underline{e} | \pi, k) > R(\underline{y}, \bar{e} | \pi, k)$. Moreover, $\Delta_0 \leq 0$ by Corollary 1. So, a necessary condition for (14) is

$$
\pi \eta \delta_w [R(\overline{y}, \bar{e} | \pi, k) - R(\underline{y}, \bar{e} | \pi, k)] \geq c. \tag{15}
$$

Now observe that $v(\max\{\overline{w}, \widetilde{w}\}) - v(\widetilde{w}) \leq v(\overline{w}) - v(\underline{w})$, and so

$$
R(\overline{y}, \bar{e} | \pi, k) - R(\underline{y}, \bar{e} | \pi, k) < R(\overline{y}, \bar{e} | \pi, k) - R(\ell') \leq (1 - \delta_w)^{-1}(1 - \delta_w^{T-k})[v(\overline{w}) - v(\underline{w})].
$$

Thus, (15) cannot be satisfied by assumption. This implies that W does not exert effort, and so the firm has no incentive to set $w^{\underline{y}}$ greater than the one–period wage it offers W.

We can then conclude that an incumbent of age 2 or more who has only produced low output never exerts effort when employed and that if the firm employs such a worker, then it always pays him the one–period wage it offers. To finish, consider an age 1 worker. The same argument used in Step 2 shows that he does not exert effort if the firm offers him probation. The desired result now follows. □

Proof of Proposition 2: In order to show that σ^* is an equilibrium, we need to prove that: (a) it is optimal for the firm to dismiss an incumbent who has never produced high output if it is not committed to employ him; (b) if the firm is to make an offer to the available age 1 worker, then it is optimal for it to offer $(\underline{w}, 0)$. The following facts will be useful. First, V^* satisfies

$$
\begin{aligned}
V^* &= (1 - \delta_f)[y(\phi_0, \alpha + \eta) - \underline{w}] + \phi_0(\alpha + \eta) \left\{ \delta_f(1 - \delta_f^{T-1})[y(1, \alpha) - \overline{w}] + \delta_f^T V^* \right\} \\
&\qquad + [1 - \phi_0(\alpha + \eta)]\delta_f V^*, \tag{16}
\end{aligned}
$$

where $y(\phi, \xi) = \phi \xi \overline{y} + (1 - \phi \xi)\underline{y}$. Second, $y(\phi_0, \alpha + \eta) - \underline{w} < V^* < y(1, \alpha) - \overline{w}$. Third, the reputation of an incumbent who has never produced high output is at most $\underline{\phi} = (1 - \alpha)\phi_0 / [1 - \phi_0 \alpha] < \phi_0$.

We start with (a). Consider an incumbent of age T who has never produced high output. If the firm employs him, then its continuation payoff is smaller than $(1-\delta_f)[y(\underline{\phi}, \alpha+\eta) - \underline{w}] + \delta_f V^* < V^*$, so that it is optimal for the firm to dismiss this worker if it can do so. Suppose now, by induction, that the firm's continuation payoff when it employs an incumbent of age $k+1$, with $k \geq 2$, who has never produced high output is smaller than V^*. Consider then an incumbent of age k who has never produced high output. By the induction hypothesis, if the firm employs him, then its continuation payoff is smaller than

$$\overline{V} = (1-\delta_f)[y(\underline{\phi}, \alpha+\eta) - \underline{w}] + \underline{\phi}(\alpha+\eta)\left\{\delta_f(1-\delta_f^{T-k})[y(1,\alpha) - \overline{w}] + \delta_f^{T+1-k}V^*\right\}$$
$$+[1 - \underline{\phi}(\alpha+\eta)]\delta_f V^*.$$

Since $V^* < (1-\delta_f^q)[y(1,\alpha) - \overline{w}] + \delta_f^q V^* < (1-\delta_f^{T-1})[y(1,\alpha) - \overline{w}] + \delta_f^{T-1}V^*$ for all $q < T-1$, we can then conclude that $\overline{V} < V^*$ by (16). Thus, by induction, the continuation payoff to the firm from employing an incumbent who has only produced low output is always smaller than V^*, which implies (a). It is easy to see that (b) follows immediately from this.

For the second part of the proof, consider an equilibrium σ where, with positive probability, the firm retains an age 2 incumbent after he produces low output. We are done if we show that the firm's payoff in any such equilibrium is strictly smaller than V^*. By construction, there exists $q \geq 2$ such that with positive probability the firm retains an incumbent of age $k \geq 2$ who has never produced high output if, and only if, $k \leq q$. Let $V(q)$ be the firm's payoff in this case. At best for the firm, a worker in σ exerts effort as long as he does not reveal himself to be of the high type. Suppose that this is the case. Then, $V(q)$ satisfies the following recursion:

$$V(q) = \phi_0 \sum_{j=0}^{q-1}(1-\alpha-\eta)^j(\alpha+\eta)\left\{(1-\delta_f^j)[\underline{y} - \underline{w}] + \delta_f^j(1-\delta_f)[\overline{y} - \underline{w}]+\right.$$
$$\left.\delta_f^{j+1}(1-\delta_f^{T-1-j})[y(1,\alpha) - \overline{w}] + \delta_f^T V(q)\right\}$$
$$+ [\phi_0(1-\alpha-\eta)^q + 1 - \phi_0]\left\{(1-\delta_f^q)[\underline{y} - \underline{w}] + \delta_f^q V(q)\right\}. \qquad (17)$$

Notice that (17) also makes sense when $q = 1$, in which case it reduces to (16). Also notice that (17) can be rewritten as $V(q) = T_q V(q)$, where T_q is a contraction (from $\mathbb{R}$ into $\mathbb{R}$).

Observe that for any $q \le T - 1$,

$$(1 - \delta_f^q)[\underline{y} - \underline{w}] + \delta_f^q V^*$$
$$> (1 - \delta_f^q)[\underline{y} - \underline{w}] + \delta_f^q \phi_0(\alpha + \eta) \left\{ (1 - \delta_f)[\overline{y} - \underline{w}] + \delta_f(1 - \delta_f^{T-q-1})[y(1, \alpha) - \overline{w}] + \delta_f^{T-q} V^* \right\}$$
$$+ \delta_f^q [1 - \phi_0(\alpha + \eta)] \left\{ (1 - \delta_f)[\underline{y} - \underline{w}] + \delta_f V^* \right\}$$
$$= \phi_0(\alpha + \eta) \left\{ (1 - \delta_f^q)[\underline{y} - \underline{w}] + \delta_f^q(1 - \delta_f)[\overline{y} - \underline{w}] + \delta_f^{q+1}(1 - \delta_f^{T-q-1})[y(1, \alpha) - \overline{w}] + \delta_f^T V^* \right\}$$
$$+ [1 - \phi_0(\alpha + \eta)] \left\{ (1 - \delta_f^{q+1})[\underline{y} - \underline{w}] + \delta_f^{q+1} V^* \right\}. \tag{18}$$

Now observe that $\phi_0(1 - \alpha - \eta)^q + 1 - \phi_0 = (1 - \alpha - \eta)^q + (1 - \phi_0)[1 - (1 - \alpha - \eta)^q]$. Hence,

$$[\phi_0(1 - \alpha - \eta)^q + 1 - \phi_0] \left\{ (1 - \delta_f^q)[\underline{y} - \underline{w}] + \delta_f^q V^* \right\}$$
$$> (1 - \phi_0)[1 - (1 - \alpha - \eta)^q] \left\{ (1 - \delta_f^q)[\underline{y} - \underline{w}] + \delta_f^q V^* \right\}$$
$$+ (1 - \alpha - \eta)^q \phi_0(\alpha + \eta) \left\{ (1 - \delta_f^q)[\underline{y} - \underline{w}] + \delta_f^q(1 - \delta_f)[\overline{y} - \underline{w}] \right.$$
$$\left. + \delta_f^{q+1}(1 - \delta_f^{T-q-1})[y(1, \alpha) - \overline{w}] + \delta_f^T V^* \right\}$$
$$+ (1 - \alpha - \eta)^q [1 - \phi_0(\alpha + \eta)] \left\{ (1 - \delta_f^{q+1})[\underline{y} - \underline{w}] + \delta_f^{q+1} V^* \right\}$$
$$> [1 - \phi_0 + \phi_0(1 - \alpha - \eta)^{q+1}] \left\{ (1 - \delta_f^{q+1})[\underline{y} - \underline{w}] + \delta_f^{q+1} V^* \right\}$$
$$+ (1 - \alpha - \eta)^q \phi_0(\alpha + \eta) \left\{ (1 - \delta_f^q)[\underline{y} - \underline{w}] + \delta_f^q(1 - \delta_f)[\overline{y} - \underline{w}] \right.$$
$$\left. + \delta_f^{q+1}(1 - \delta_f^{T-q-1})[y(1, \alpha) - \overline{w}] + \delta_f^T V^* \right\},$$

where the first inequality follows from (18). Therefore, $T_q V^* > T_{q+1} V^*$ for all $q \le T - 1$. Since $V^* = T_1 V^*$, we then have that $V^* > T_q V^*$ for all $q \in \{2, \ldots, T\}$. Since $(T_q)^n V^*$ converges to $V(q)$ by the contraction mapping theorem, we can then conclude that $V^* > V(q)$ for all $q \in \{2, \ldots, T\}$, the desired result. $\qquad \square$

Proof of Lemma 9: Consider an incumbent of age $k \in \{2, \ldots, T\}$ who has only produced low output and denote this worker by W. We proceed by induction in k.

(1) Suppose that $k = T$. It is immediate to see that W never exerts effort when employed. This implies that: (i) if the firm makes W an offer, then it offers him the lowest one–period wage possible; (ii) the firm never pays W a bonus if it employs him.

(2) Suppose, by induction, that there exists $k' \ge 3$ such that if $k \ge k'$, then (i) and (ii) hold and

W does not exert effort when employed. Let $k = k' - 1$ and consider W's incentive–compatibility constraint for effort exertion. It is given by

$$-c + \pi\xi \left\{v(w^{\overline{y}}) + \delta_w R(\overline{y}, \overline{e}|\pi, k)\right\} + (1 - \pi\xi)\left\{v(w^{\underline{y}}) + \delta_w R(\underline{y}, \overline{e}|\pi, k)\right\}$$
$$\geq \pi\xi \left\{v(w^{\overline{y}}) + \delta_w R(\overline{y}, \underline{e}|\pi, k)\right\} + (1 - \pi\xi)\left\{v(w^{\underline{y}}) + \delta_w R(\underline{y}, \underline{e}|\pi, k)\right\}, \quad (19)$$

where $\pi \leq \underline{\phi}$ is W's private belief that he is of the high type, ξ is the probability that W produces high output, w^y is the wage the firm pays W if he produces y, and $R(y, e|\pi, k)$ is W's continuation payoff after he chooses e and produces y, expressed as a function of π and k. Notice that ξ is independent of W's (current) choice of effort.

Since $R(\overline{y}, \overline{e}|\pi, k) = R(\overline{y}, \underline{e}|\pi, k)$, a necessary and sufficient condition for (19) is that

$$(1 - \pi\xi)\delta_w \left[R(\underline{y}, \overline{e}|\pi, k) - R(\underline{y}, \underline{e}|\pi, k)\right] \geq (1 - \delta_w)c. \quad (20)$$

Notice that $R(\underline{y}, \overline{e}|\pi, k) = R(\underline{y}, \underline{e}|\pi, k)$ if W is dismissed after low output. So, suppose the firm makes W an offer after he produces low output; let $\widetilde{w}$ be the one–period wage the firm offers. By Lemma 8 and the induction hypothesis, the firm makes W an offer if he produces low output one more time only if it is committed to do so. Let $q \in \{0, \ldots, T - k - 1\}$ be the number of periods the firm is committed to make W an offer after he produces low output two times in a row and let $\widetilde{w}'$ be the smallest one–period wage the firm is committed to offer W during this period of time. Moreover, let $\pi' = (1 - \xi)\pi/(1 - \xi\pi)$, $\xi(\underline{e}) = \alpha$, and $\xi(\overline{e}) = \gamma$. Then, by Lemmas 4 and 6,

$$R(\underline{y}, e|\pi, k) = v(\widetilde{w})$$
$$+\pi'\xi(e)\left\{\delta_w(1 - \delta_w)^{-1}(1 - \delta_w^q)v(\max\{\widetilde{w}', \overline{w}\}) + \delta_w^{q+1}(1 - \delta_w)^{-1}(1 - \delta_w^{T-k-q-1})v(\overline{w})\right\}$$
$$+\pi'(1 - \xi(e))R(q) + \left[1 - \pi' + \pi'(1 - \xi(e))(1 - \alpha)^q\right]\left\{\delta_w(1 - \delta_w)^{-1}(1 - \delta_w^q)v(\widetilde{w}')\right.$$
$$\left. + \delta_w^{q+1}(1 - \delta_w)^{-1}(1 - \delta_w^{T-k-q-1})v(\underline{w})\right\},$$

where $R(0) = \delta_w(1 - \delta_w)^{-1}(1 - \delta_w^{T-k-1})v(\underline{w})$ and

$$R(q) = \sum_{j=0}^{q-1}(1 - \alpha)^j\alpha\left\{\delta_w(1 - \delta_w)^{-1}(1 - \delta_w^{j+1})v(\widetilde{w}') + \delta_w^{j+1}(1 - \delta_w)^{-1}(1 - \delta_w^q)v(\max\{\widetilde{w}', \overline{w}\})\right.$$
$$\left. + \delta_w^{q+1}(1 - \delta_w)^{-1}(1 - \delta_w^{T-k-q-1})v(\overline{w})\right\}$$

if $q \geq 1$. Therefore,

$$R(\underline{y}, \overline{e}|\pi, k) - R(\underline{y}, \underline{e}|\pi, k) =$$
$$\pi'(\gamma - \alpha) \left\{ \delta_w (1 - \delta_w)^{-1} (1 - \delta_w^q) v(\max\{\widetilde{w}', \overline{w}\}) + \delta_w^{q+1} (1 - \delta_w)^{-1} (1 - \delta_w^{T-k-q-1}) v(\overline{w}) \right\}$$
$$+ \pi'(\alpha - \gamma) R(q)$$
$$+ \pi'(\alpha - \gamma)(1 - \alpha)^q \left\{ \delta_w (1 - \delta_w)^{-1} (1 - \delta_w^q) v(\widetilde{w}') + \delta_w^{q+1} (1 - \delta_w)^{-1} (1 - \delta_w^{T-k-q-1}) v(\underline{w}) \right\}.$$

Since $v(\max\{\widetilde{w}', \overline{w}\}) - v(\widetilde{w}') \leq v(\overline{w}) - v(\underline{w})$ and

$$R(q) \geq [1 - (1 - \alpha)^q] \left\{ \delta_w (1 - \delta_w)^{-1} (1 - \delta_w^q) v(\widetilde{w}') + \delta_w^{q+1} (1 - \delta_w)^{-1} (1 - \delta_w^{T-k-q-1}) v(\underline{w}) \right\}$$

for all $q \in \{0, \ldots, T - k - 1\}$, we can then conclude that

$$R(\underline{y}, \overline{e}|\pi, k) - R(\underline{y}, \underline{e}|\pi, k) \leq \pi'(\gamma - \alpha)\delta_w (1 - \delta_w)^{-1} (1 - \delta_w^{T-k-1})[v(\overline{w}) - v(\underline{w})].$$

Thus, (20) cannot be satisfied by (A3), which implies that W does not exert effort when $k = k' - 1$. Consequently, (i) and (ii) also hold when $k = k' - 1$. □

Proof of Proposition 3: We know from the main text that a profile σ with a history $\widehat{h}$ for the firm after which, even though not committed to do so, it offers $(w, (q, w'))$ to an age 2 worker who did not produce high output when of age 1 can be an equilibrium only if $q = 0$ and $w = \underline{w}$. In fact, if $q > 0$, then the following deviation by the firm is profitable by Lemma 8: (i) offer the age 2 worker $(\underline{w}, 0)$; (ii) behave as if no deviation has occurred if the age 2 worker produces high output; (iii) hire the available age 1 worker if the age 2 worker produces low output. It is immediate to see that there is a profitable deviation for the firm if $q = 0$ and $w > \underline{w}$.

Let $V = V(h|\sigma)$ for any history h for the firm after which it makes an offer to an age 1 worker. By (A2), V satisfies the following recursion,

$$V = (1 - \delta_f)[y(\phi_0, \alpha) - \underline{w}] + \phi_0 \alpha \left\{ \delta_f (1 - \delta_f)[y(1, \gamma) - \overline{w}] + \delta_f^2 (1 - \delta_f^{T-2})[y(1, \alpha) - \overline{w}] + \delta_f^T V \right\} +$$
$$(1 - \phi_0 \alpha) \left\{ \delta_f (1 - \delta_f)[y(\underline{\phi}, \gamma) - \underline{w}] + \underline{\phi}\gamma \left\{ \delta_f^2 (1 - \delta_f^{T-2})[y(1, \alpha) - \overline{w}] + \delta_f^T V \right\} + (1 - \underline{\phi}\gamma)\delta_f^2 V \right\}$$
$$= (1 - \delta_f)[y(\phi_0, \alpha) - \underline{w}] + \delta_f (1 - \delta_f) \left\{ \phi_0 \alpha[y(1, \gamma) - \overline{w}] + (1 - \phi_0 \alpha)[y(\underline{\phi}, \gamma) - \underline{w}] \right\} +$$
$$\phi_0 [\alpha + \gamma(1 - \alpha)] \left\{ \delta_f^2 (1 - \delta_f^{T-2})[y(1, \alpha) - \overline{w}] + \delta_f^T V \right\} + \{1 - \phi_0 [\alpha + \gamma(1 - \alpha)]\}\delta_f^2 V,$$

from which we obtain

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

Now recall from the main text that the desired result holds if

$$V > \widetilde{V} = [y(\underline{\phi}, \gamma) - \underline{w}] + \underline{\phi}\gamma \frac{\delta_f(1 - \delta_f^{T-2})}{1 - \delta_f}[y(1,\alpha) - \overline{w} - V]$$

for δ_f sufficiently close to one. Since

$$\lim_{\delta_f \to 1} V = A = [y(1,\alpha) - \overline{w}] - \frac{\Delta + \phi_0\alpha[y(1,\alpha) - y(1,\gamma)] + (1 - \phi_0\alpha)\Delta'}{2 + \phi_0[\alpha + \gamma(1 - \alpha)](T - 2)}$$

and

$$\lim_{\delta_f \to 1} \widetilde{V} = B = [y(\underline{\phi}, \gamma) - \underline{w}] + \frac{\Delta + \phi_0\alpha[y(1,\alpha) - y(1,\gamma)] + (1 - \phi_0\alpha)\Delta'}{2 + \phi_0[\alpha + \gamma(1 - \alpha)](T - 2)}\underline{\phi}\gamma(T - 2),$$

we are done if we show that

$$A - B = \Delta' - \frac{1 + \underline{\phi}\gamma(T - 2)}{2 + \phi_0[\alpha + \gamma(1 - \alpha)](T - 2)}\left[\Delta - \phi_0\alpha(\gamma - \alpha)\Delta_y + (1 - \phi_0\alpha)\Delta'\right] > 0.$$

For this, notice that $A > B$ if, and only if,

$$\{2 + \phi_0[\alpha + \gamma(1 - \alpha)](T - 2)\}\Delta' > [1 + \underline{\phi}\gamma(T - 2)]\left[\Delta - \phi_0\alpha(\gamma - \alpha)\Delta_y + (1 - \phi_0\alpha)\Delta'\right],$$

which reduces to

$$[1 + \phi_0\alpha(T - 1)]\Delta' - [1 + \underline{\phi}\gamma(T - 2)]\left[\Delta - \phi_0\alpha(\gamma - \alpha)\Delta_y\right] > 0.$$

Since $\Delta' = \Delta + y(\phi_0, \alpha) - y(\underline{\phi}, \gamma) = \Delta + (\phi_0\alpha - \underline{\phi}\gamma)\Delta_y$, we then have that $A > B$ if, and only if,

$$\underline{\phi}\gamma\left\{[1 + \phi_0\alpha(T - 1)]\Delta_y + [\Delta - \phi_0\alpha(\gamma - \alpha)\Delta_y](T - 2)\right\}$$
$$< \phi_0\alpha\left\{(T - 1)\Delta + [1 + \phi_0\alpha(T - 1)]\Delta_y + (\gamma - \alpha)\Delta_y\right\};$$

that is, if and only if

$$\underline{\phi}\gamma < \phi_0\alpha\left\{1 + \frac{\Delta + (\gamma - \alpha)\Delta_y + \phi_0\alpha(\gamma - \alpha)(T - 2)\Delta_y}{(T - 2)\Delta + [1 + \phi_0\alpha(T - 1)]\Delta_y - \phi_0\alpha(\gamma - \alpha)\Delta_y(T - 2)}\right\},$$

which reduces to condition (5). $\qquad\qquad\square$

Proof of Lemma 11: Suppose not. So, there is a history for the firm after which it employs a worker of age $k \geq 2$ with an output history $\widetilde{y}^k$ that includes at least one high output and the

worker, whom we denote by W, exerts effort. Suppose the long–term contract the firm offers W is (ω, τ) and consider the deviation where the firm offers him (ω', τ) instead of (ω, τ), where the only difference between ω and ω' is that $\omega'_{k+1}(\widetilde{y}^k, y) \equiv w'$, with

$$v(w') = (\alpha + \eta)v(w_{k+1}(\widetilde{y}^k, \overline{y})) + (1 - \alpha - \eta)v(w_{k+1}(\widetilde{y}^k, \underline{y})) - c.$$

Notice that $w' > \overline{w}$ since W's participation constraint after $\widetilde{y}^k$ must be satisfied. By construction, a worker behaves under (ω', τ) in the same as he behaves under (ω, τ), except that he does not exert effort after $\widetilde{y}^k$—the fact that a worker knows he is of the high type after $\widetilde{y}^k$ is key for this. So, this deviation changes flow payoffs for the firm only after $\widetilde{y}^k$. Before the deviation, the flow payoff to the firm after $\widetilde{y}^k$ is

$$
\begin{aligned}
R &= (1 - \delta_f)\left\{ y(1, \alpha) + \eta\Delta_y - (\alpha + \eta)w_{k+1}(\widetilde{y}^k, \overline{y}) - (1 - \alpha - \eta)w_{k+1}(\widetilde{y}^k, \underline{y}) \right\} \\
&= (1 - \delta_f)\left\{ y(1, \alpha) + \eta\Delta_y - (\alpha + \eta)v^{-1}(v(w_{k+1}(\widetilde{y}^k, \overline{y}))) - (1 - \alpha - \eta)v^{-1}(v(w_{k+1}(\widetilde{y}^k, \underline{y}))) \right\} \\
&\leq (1 - \delta_f)\left\{ y(1, \alpha) + \eta\Delta_y - v^{-1}\left[(\alpha + \eta)v(w_{k+1}(\widetilde{y}^k, \overline{y})) + (1 - \alpha - \eta)v(w_{k+1}(\widetilde{y}^k, \underline{y})) \right] \right\} \\
&= (1 - \delta_f)\left\{ y(1, \alpha) + \eta\Delta_y - v^{-1}(v(w') + c) \right\} = (1 - \delta_f)G(w'),
\end{aligned}
$$

where the inequality follows from the fact that v^{-1} is convex. After the deviation, the flow payoff to the firm after $\widetilde{y}^k$ is $(1 - \delta_f)r(w', \underline{e})$. Now observe, by the inverse function theorem, that G is differentiable and $G'(w) = -v'(w)/v'(v^{-1}(v(w) + c))$.[21] Since v^{-1} is strictly increasing and v' is weakly decreasing, we then have that $G'(w) \leq -1$. So,

$$G(w') - G(\overline{w}) = \int_{\overline{w}}^{w'} G'(w)dw \leq -(w' - \overline{w}) = r(w', \underline{e}) - r(\overline{w}, \underline{e}).$$

We can then conclude by (A4) that

$$R \leq (1 - \delta_f)G(w') \leq (1 - \delta_f)\{G(\overline{w}) + r(w', \underline{e}) - r(\overline{w}, \underline{e})\} < (1 - \delta_f)r(w', \underline{e}),$$

which implies that the deviation under consideration is profitable. $\qquad\square$

Proof of Proposition 5: We first show that a worker who has only produced low output has the greatest incentive to exert effort when, all else the same, he is dismissed after low output.

---

[21]It is possible to show that the concavity of v implies that G is absolutely continuous on $[\overline{w}, w']$ with $G'(w)$, which exists almost everywhere, less than or equal to -1 for almost all $w \in [\overline{w}, w']$. So, the result of Lemma 11 does not depend on the assumption that v is continuously differentiable. Details are available upon request.

Consider such a worker, let k be his age, $\pi \leq \phi_0$ be his private belief that he is of the high type, and q be the number of consecutive times he can produce low output before he is dismissed; $q = 0$ implies he is dismissed after low output. The worker's incentive–compatibility constraint for effort exertion is

$$\pi(\alpha + \eta)\left\{v(w^{\overline{y}}) + \delta_w R(\overline{y}, \overline{e}|\pi, k, q)\right\} + [1 - \pi(\alpha + \eta)]\left\{v(w^{\underline{y}}) + \delta_w R(\underline{y}, \overline{e}|\pi, k, q)\right\}$$
$$\geq c + \pi\alpha\left\{v(w^{\overline{y}}) + \delta_w R(\overline{y}, \underline{e}|\pi, k, q)\right\} + (1 - \pi\alpha)\left\{v(w^{\underline{y}}) + \delta_w R(\underline{y}, \underline{e}|\pi, k, q)\right\}, \quad (21)$$

where w^y is the wage the firm pays him if his output is y, and $R(y, e|\pi, k, q)$ is his continuation payoff if he chooses e and produces y, which depends on π, k, and q. We want to show, holding $w^{\underline{y}}$, $w^{\overline{y}}$, and the worker's continuation payoff after he produces high output constant, that if (21) is satisfied for some $q > 0$, then it is also satisfied when $q = 0$.

First notice that $R(\overline{y}, e|\pi, k, q)$ does not depend on q and, as in the case without full commitment, that $R(\overline{y}, \overline{e}|\pi, k) = R(\overline{y}, \underline{e}|\pi, k)$. So, (21) reduces to

$$\pi\eta[v(w^{\overline{y}}) - v(w^{\underline{y}})] + \pi\eta\delta_w[R(\overline{y}, \overline{e}|\pi, k) - R(\underline{y}, \overline{e}|\pi, k, q)]$$
$$+ \delta_w(1 - \pi\alpha)[R(\underline{y}, \overline{e}|\pi, k, q) - R(\underline{y}, \underline{e}|\pi, k, q)] \geq c. \quad (22)$$

Now notice that (22) reduces to

$$\pi\eta[v(w^{\overline{y}}) - v(w^{\underline{y}})] + \pi\eta\delta_w[R(\overline{y}, \overline{e}|\pi, k) - (1 - \delta_w)^{-1}(1 - \delta_w^{T-k-1})v(\underline{w})] \geq c \quad (23)$$

when $q = 0$. Hence, we are done if we show that the left–hand side of (23) is greater than the left–hand side of (22) when $q > 0$.

Let $q \geq 1$ and suppose the worker exerts effort and produces low output. Now let ξ_s and $R(s)$, with $s \in \{0, \ldots, q-1\}$, be the probability that he produces high output for the first time after $s+1$ periods and his continuation payoff in this event, respectively. Moreover, let ξ_q and $R(q)$ be the probability that the worker never produces high output in the next q periods and his continuation payoff in this event, respectively. To finish, recall that $\underline{\pi}(e) = (1 - \xi(e))\pi/[1 - \pi\xi(e)]$, where $\xi(\underline{e}) = \alpha$ and $\xi(\overline{e}) = \alpha + \eta$. Then,

$$R(\underline{y}, \overline{e}|\pi, k, q) = \underline{\pi}(\overline{e})\sum_{s=0}^{q-1}\xi_s R(s) + [1 - \underline{\pi}(\overline{e}) + \underline{\pi}(\overline{e})\xi_q] R(q) = \underline{\pi}(\overline{e})\sum_{s=0}^{q-1}\xi_s[R(s) - R(q)] + R(q).$$

44

Now suppose the worker exerts no effort and produces low output. Since after this he has the option of behaving as if he exerted effort and produced low output, we have that $R(\underline{y}, \underline{e}|\pi, k, q) \geq \underline{\pi}(\underline{e}) \sum_{s=0}^{q-1} \xi_s[R(s) - R(q)] + R(q)$. Therefore,

$$\pi\eta[R(\overline{y}, \overline{e}|\pi, k) - R(\underline{y}, \overline{e}|\pi, k, q)] + (1 - \pi\alpha)[R(\underline{y}, \overline{e}|\pi, k, q) - R(\underline{y}, \underline{e}|\pi, k, q)]$$

$$\leq \pi\eta \left\{ R(\overline{y}, \overline{e}|\pi, k) - \underline{\pi}(\overline{e}) \sum_{s=0}^{q-1} \xi_s[R(s) - R(q)] - R(q) \right\}$$

$$+ (1 - \pi\alpha) \left\{ (\underline{\pi}(\overline{e}) - \underline{\pi}(\underline{e})) \sum_{s=0}^{q-1} \xi_s[R(s) - R(q)] \right\}$$

$$= \pi\eta\left[R(\overline{y}, \overline{e}|\pi, k) - R(q)\right] + \left\{[1 - \pi(\alpha + \eta)]\underline{\pi}(\overline{e}) - (1 - \pi\alpha)\underline{\pi}(\underline{e})\right\} \sum_{s=0}^{q-1} \xi_s[R(s) - R(q)]$$

$$= \pi\eta\left[R(\overline{y}, \overline{e}|\pi, k) - R(q) - \sum_{s=0}^{q-1} \xi_s[R(s) - R(q)]\right] = \pi\eta\left[R(\overline{y}, \overline{e}|\pi, k) - \sum_{s=0}^{q} \xi_s R(s)\right].$$

The desired result follows from the fact that $\sum_{s=0}^{q} \xi_s R(s) > (1 - \delta_w)^{-1}(1 - \delta_w^{T-k-1})v(\underline{w})$.

Suppose, by contradiction, that there is an equilibrium σ with a history $\widehat{h}$ for the firm after which it offers (ω, τ) with $\tau_2(\underline{y}) > 0$ to the available age 1 worker. Denote this worker by W. We consider the case where $\tau_3(\underline{y}, \underline{y}) = 0$. It is straightforward, but notationally cumbersome, to modify the following argument to cover the case where there exists $k > 3$ such that $\tau_s(\underline{y}, \ldots, \underline{y}) > 0$ for all $3 \leq s \leq k - 1$ and $\tau_k(\underline{y}, \ldots, \underline{y}) = 0$. Let V_{y^t} be the firm's continuation payoff after W produces y^t and is paid $\omega_t(y^t)$. Moreover, let e_1 be W's age 1 effort choice and e_2 be W's age 2 effort choice if he produces low output when of age 1 and receives an offer by the firm. Then,

$$V(\widehat{h}|\sigma) = \xi(e_1)\phi_0 \left\{(1 - \delta_f)[\overline{y} - \omega_1(\overline{y})] + \delta_f V_{\overline{y}}\right\}$$

$$+ [1 - \phi_0\xi(e_1)]\left\{(1 - \delta_f)[\underline{y} - \omega_1(\underline{y})] + \tau_2(\underline{y})\underline{\phi}\xi(e_2)\delta_f\left\{(1 - \delta_f)[\overline{y} - \omega_2(\underline{y}, \overline{y})] + \delta_f V_{(\underline{y}, \overline{y})}\right\}\right.$$

$$+ \tau_2(\underline{y})[1 - \underline{\phi}\xi(e_2)]\delta_f\left\{(1 - \delta_f)[\underline{y} - \omega_2(\underline{y}, \underline{y})] + \delta_f V(\widehat{h}|\sigma)\right\}$$

$$\left. + [1 - \tau_2(\underline{y})]V(\widehat{h}|\sigma)\right\},$$

where $\underline{\phi} = [1 - \xi(e_1)]\phi_0/[1 - \phi_0\xi(e_1)]$, $\xi(\overline{e}) = \gamma$, and $\xi(\underline{e}) = \alpha$.

Consider now the following deviation by the firm, where we only describe the instances in which the firm does not behave according to σ. First, offer (ω', τ') to W such that: (i) $\tau_2'(\underline{y}) = 0$; (ii) $\tau_t'(\overline{y}y^{t-2}) = 1$ for all $y^{t-2} \in Y^{t-2}$ with $2 \leq t \leq T$; (iii) $\omega_1'(\underline{y}) = \underline{w}$, $\omega_1'(\overline{y}) = \omega_1(\overline{y})$, and $\omega_t' \equiv \omega_t$

for all $2 \leq t \leq T$ if $e_1 = \bar{e}$; (iv) $\omega'_1 \equiv \underline{w}$ and $\omega'_t(\bar{y}y^{t-1}) = \bar{w}$ for all $y^{t-1} \in Y^{t-1}$ with $2 \leq t \leq T$ if $e_1 = \underline{e}$. Second, if W produces low output in his first period of employment, then behave as if no deviation has occurred with probability $1 - \tau_2(\underline{y})$ and offer (ω^*, τ^*) to the available age 1 worker with probability $\tau_2(\underline{y})$, where: (i) $\tau_2^*(\underline{y}) = 0$ and $\tau_2^*(\bar{y}) = \underline{\phi}/\phi_0$; (ii) $\tau_t^*(\bar{y}y^{t-2}) = 1$ for all $y^{t-2} \in Y^{t-2}$ with $3 \leq t < T - 1$ and $\tau_T^* \equiv 0$; (iii) $\omega_1^*(\underline{y}) = \underline{w}$, $\omega_1^*(\bar{y}) = \omega_2(\underline{y}, \bar{y})$, and $\omega_t^*(y^t) = \omega_{t+1}(\underline{y}y^t)$ for all $y^t \in Y^t$ with $2 \leq t \leq T - 1$ if $e_2 = \bar{e}$—ω_T^* is irrelevant since the worker who is offered (ω^*, τ^*) is never employed when he is of age T; (iv) $\omega_1^* \equiv \underline{w}$ and $\omega_t^*(\bar{y}y^{t-1}) = \bar{w}$ for all $y^{t-1} \in Y^{t-1}$ with $1 \leq t \leq T - 1$ if $e_2 = \underline{e}$.

Notice that W's behavior when he is of age 1 is the same before and after the deviation. This follows from the first result we established in the proof if $e_1 = \bar{e}$ and from (1) if $e_1 = \underline{e}$. Likewise, an age 1 worker who is offered (ω^*, τ^*) behaves in the same way as W behaves under σ after he produces low output in his first period of employment. So, if V^* denotes the firm's continuation payoff after $\widehat{h}$ when it follows the deviation described above, then

$$
\begin{aligned}
V^* = {} & \xi(e_1)\phi_0 \left\{ (1 - \delta_f)[\bar{y} - \omega'_1(\bar{y})] + \delta_f V'_{\bar{y}} \right\} \\
& + [1 - \phi_0\xi(e_1)] \Big\{ (1 - \delta_f)[\underline{y} - \omega'_1(\underline{y})] + \tau_2(\underline{y})\phi_0\xi(e_2)\delta_f \Big\{ (1 - \delta_f)[\bar{y} - \omega_1^*(\bar{y})] + \delta_f\tau_2^*(\bar{y})V^*_{(\underline{y},\bar{y})} + \\
& \hspace{6cm} + \delta_f[1 - \tau_2^*(\bar{y})]V(\widehat{h}|\sigma) \Big\} \\
& \hspace{2cm} + \tau_2(\underline{y})[1 - \phi_0\xi(e_2)]\delta_f \Big\{ (1 - \delta_f)[\underline{y} - \omega_1^*(\underline{y})] + \delta_f V(\widehat{h}|\sigma) \Big\} \\
& \hspace{2cm} + [1 - \tau_2(\underline{y})]V(\widehat{h}|\sigma) \Big\},
\end{aligned}
$$

where $V'_{\bar{y}} \geq V_{\bar{y}}$ and $V^*_{(\underline{y},\bar{y})} \geq V_{(\underline{y},\bar{y})}$ by construction. Since $\tau_2(\underline{y})\phi_0[1 - \tau_2^*(\bar{y})] = \tau_2(\underline{y})(\phi_0 - \underline{\phi})$ and $\phi_0\xi(e_2)\bar{y} + (1 - \phi_0\xi(e_2))\underline{y} > \underline{\phi}\xi(e_2)\bar{y} + (1 - \underline{\phi}\xi(e_2))\underline{y}$, we then have $V^* > V(\widehat{h}|\sigma)$, and so the deviation described above is profitable for the firm. $\qquad\square$

Proof of Lemma 11 in the non–IID case: Suppose not. So, there is a history for the firm after which it employs a worker of age $k \in \{2, \ldots, T - 1\}$ with an output history $\widetilde{y}^k$ that includes at least one high output and the worker, that we denote by W, exerts effort. Let t be the period in which this happens and suppose the long–term contract the firm offers W is (ω, τ). Moreover, let y_1 denote W's output in t, y_2 denote W's output in $t + 1$ (in case he is employed by the firm), and ξ be the probability that $y_1 = \bar{y}$. Consider now the deviation where the firm offers W the contract

(ω', τ), where the only difference between ω and ω' is that $\omega'_{k+2}(\widetilde{y}^k, y_1, y_2) = w^{y^1 y^2} \equiv w'$, with

$$\xi \delta_w \left\{ \tau_{k+2}(\widetilde{y}^k, \overline{y})[\gamma v(w^{\overline{y}\overline{y}}) + (1-\gamma)v(w^{\overline{y}\underline{y}})] + [1 - \tau_{k+2}(\widetilde{y}^k, \overline{y})]v(\overline{w}) \right\}$$
$$+ (1-\xi)\delta_w \left\{ \tau_{k+2}(\widetilde{y}^k, \underline{y})[\gamma v(w^{\underline{y}\overline{y}}) + (1-\gamma)v(w^{\underline{y}\underline{y}})] + [1 - \tau_2(\widetilde{y}^k, \underline{y})]v(\overline{w}) \right\} - c = \delta_w v(w').$$

Notice that we must have $\tau_{k+2}(\widetilde{y}^k, y) > 0$ for at least one $y \in Y$ in order for W to exert effort after $\widetilde{y}^k$. Also notice, since W's incentive–compatibility constraint for effort is satisfied after $\widetilde{y}^k$ under (ω, τ), that the left–hand side of the above equation is at least equal to

$$\xi \delta_w \left\{ \tau_{k+2}(\widetilde{y}^k, \overline{y})[\alpha v(w^{\overline{y}\overline{y}}) + (1-\alpha)v(w^{\overline{y}\underline{y}})] + [1 - \tau_{k+2}(\widetilde{y}^k, \overline{y})]v(\overline{w}) \right\}$$
$$+ (1-\xi)\delta_w \left\{ \tau_{k+2}(\widetilde{y}^k, \underline{y})[\alpha v(w^{\underline{y}\overline{y}}) + (1-\alpha)v(w^{\underline{y}\underline{y}})] + [1 - \tau_2(\widetilde{y}^k, \underline{y})]v(\overline{w}) \right\}.$$

So, limited liability implies that $w' > \overline{w}$. By construction, W's behavior after the deviation only changes after $\widetilde{y}^k$, when he does not exert effort. So, the (flow) payoffs to the firm stay the same except after $\widetilde{y}^k y_1$, when they increase from $(1 - \delta_f)\{y(1, \alpha) + (\gamma - \alpha)\Delta_y - \gamma w^{y^1 \overline{y}} - (1 - \gamma)w^{y^1 \underline{y}}\}$ to $(1 - \delta_f)r(\overline{w}, \underline{e})$ with probability $\tau_{k+2}(\widetilde{y}^k, y^1)$ by condition (A5).

Appendix B: Omitted Details

1. IID Case

Here we give an example of an equilibrium in the IID case where the firm offers probation to age 1 workers. For this, let $\underline{\phi}(e) = [1 - \xi(e)]\phi_0 / [1 - \phi_0 \xi(e)]$, where $\xi(\underline{e}) = \alpha$ and $\xi(\overline{e}) = \alpha + \eta$. Then, $\underline{\phi}(e)$ is the reputation of an age 2 worker if he chooses e and produces low output in his first period of employment. Moreover, let $\underline{\underline{\phi}} = (1 - \alpha)^2 \phi_0 / [(1 - \alpha)^2 \phi_0 + 1 - \phi_0]$ be the highest reputation possible for an incumbent of age 3 or more who has never produced high output. Suppose then that

$$\underline{\phi}(\underline{e})(\alpha + \eta) > 2\phi_0 \alpha, \tag{24}$$

$$\underline{\phi}(\underline{e})\eta \delta_w (1 - \delta_w^{T-2})[v(\overline{w}) - v(\underline{w})] > (1 - \delta_w)c, \tag{25}$$

$$\max\{\underline{\underline{\phi}}, \underline{\phi}(\overline{e})\}\eta \delta_w (1 - \delta_w^{T-2})[v(\overline{w}) - v(\underline{w})] < (1 - \delta_w)c, \tag{26}$$

$$\phi_0 \eta \delta_w [v(\overline{w}) - v(\underline{w})] < c. \tag{27}$$

Notice that (24) is satisfied for all $\phi_0 \in (0,1)$ if $\eta(1-\alpha) > \alpha(1+\alpha)$—this condition reduces to $\alpha < \eta/2$ when $\alpha + \eta = 1$—and that $\underline{\underline{\phi}}(\bar{e}) < \underline{\underline{\phi}}$ when $\alpha + \eta$ is close to one. Now observe that we can choose T and $v(\overline{w}) - v(\underline{w})$ to be such that (25) to (27) are satisfied as long as the workers are patient enough. For instance, let T be such that $\phi_0 < \underline{\underline{\phi}}(T-2)$ and then choose $v(\overline{w}) - v(\underline{w})$ to be such that $\underline{\underline{\phi}}\eta(T-2)[v(\overline{w}) - v(\underline{w})] = c$.

Consider the strategy profile σ^{**} where: (i) the firm offers $(\overline{w}, 0)$ to an incumbent it knows is of high ability; (ii) the firm offers $(\underline{w}, (1, \underline{w}))$ 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 $(\underline{w}, 0)$ 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 $(w, 0)$ to an incumbent if it is committed to offer him a one–period wage of at least w; (v) the firm always pays the one–period wage it offers; and (vi) the only worker who exerts effort is an age 2 worker who, after not exerting effort and producing low output in his first period of employment, receives an offer of $(w, 0)$ by the firm—observe that if the firm offers $(\underline{w}, (1, \underline{w}))$ to an age 1 worker, then it offers this worker $(\underline{w}, 0)$ after he produces low output. We claim that σ^{**} is an equilibrium if the difference between the left–hand side and the right–hand side of (25) is not too large and the firm is sufficiently patient.

Proof: A straightforward modification of the proof of Proposition 1 shows that: (a) if (26) is satisfied, then a worker of age 3 or more who has always produced low output and an age 2 worker who exerts effort and produces low output when of age 1 do not exert effort when employed; (b) if the difference between the left–hand side and the right–hand side of (25) is small enough, then an age 2 worker who does not exert effort and produces low output when of age 1 only exerts effort if offered $(w, 0)$ by the firm. So, in order to prove that (vi) is incentive–compatible, we only need to show that an age 1 worker never exerts effort. By (i) to (vi), an age 1 worker who is offered $(w, 0)$ and an age 1 worker who is offered $(w, (1, \underline{w}))$ have the same incentive–compatibility constraint for

effort exertion, which is given by

$$-(1 - \delta_w)c + \phi_0(\alpha + \eta)\delta_w(1 - \delta_w^{T-1})v(\overline{w})$$
$$+[1 - \phi_0(\alpha + \eta)]\delta_w\left\{(1 - \delta_w)v(\underline{w}) + \underline{\phi}(\overline{e})\alpha\delta_w(1 - \delta_w^{T-2})v(\overline{w}) + [1 - \underline{\phi}(\overline{e})\alpha]\delta_w(1 - \delta_w^{T-2})v(\underline{w})\right\}$$
$$\geq \phi_0\alpha\delta_w(1 - \delta_w^{T-1})v(\overline{w}) + (1 - \phi_0\alpha)\delta_w\left\{(1 - \delta_w)v(\underline{w}) + \underline{\phi}(\underline{e})(\alpha + \eta)\delta_w(1 - \delta_w^{T-2})v(\overline{w})\right.$$
$$\left. +[1 - \underline{\phi}(\underline{e})(\alpha + \eta)]\delta_w(1 - \delta_w^{T-2})v(\underline{w})\right\}. \tag{28}$$

It is possible, but tedious, to show that if (28) is not satisfied, then an age 1 worker has no incentive to exert effort regardless of the firm's offer. Intuitively, when instead of $(w, (1, \underline{w}))$ the firm offers an age 1 $(w, (q, w'))$ with either $q > 1$ or $w' > \underline{w}$, it increases the worker's continuation payoff after low output by more than it increases his continuation payoff after high output. Now notice that (28) can be rewritten as

$$-(1 - \delta_w)c + \phi_0(\alpha + \eta)\delta_w(1 - \delta_w)v(\overline{w}) + [1 - \phi_0(\alpha + \eta)]\delta_w(1 - \delta_w)v(\underline{w})$$
$$+\phi_0[\alpha + (1 - \alpha)(\alpha + \eta)]\delta_w^2(1 - \delta_w^{T-2})v(\overline{w}) + \{1 - \phi_0[\alpha + (1 - \alpha)(\alpha + \eta)]\}\delta_w^2(1 - \delta_w^{T-2})v(\underline{w})$$
$$\geq \phi_0\alpha\delta_w(1 - \delta_w)v(\overline{w}) + (1 - \phi_0\alpha)\delta_w(1 - \delta_w)v(\underline{w})$$
$$+\phi_0[\alpha + (1 - \alpha)(\alpha + \eta)]\delta_w^2(1 - \delta_w^{T-2})v(\overline{w}) + \{1 - \phi_0[\alpha + (1 - \alpha)(\alpha + \eta)]\}\delta_w^2(1 - \delta_w^{T-2})v(\underline{w}),$$

which cannot be satisfied by (27). Thus, (vi) is indeed incentive–compatible.

Suppose the firm has an incumbent of age 3 or more who has always produced low output and is not committed to make him an offer. We know by Lemma 7 that σ^{**} can only be an equilibrium if the firm makes an offer to the available age 1 worker. By construction, the firm is indifferent between offering $(\underline{w}, 0)$ and $(\underline{w}, (1, \underline{w}))$ to this worker. By (vi), any other offer by the firm does not lead to a higher continuation payoff. So, the only thing left to prove is that the decision in (iii) is incentive–compatible for the firm. For this, let V denote the firm's continuation payoff after it offers an age 1 worker $(\underline{w}, (1, \underline{w}))$ and let V' denote the firm's continuation payoff after is offers $(\underline{w}, 0)$ to an age 2 worker who produces low output when of age 1. We are done if we show that $V' > V$. The same argument used in the proof of Proposition 3 shows that

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

where $\Delta = y(1, \alpha) - \overline{w} - [y(\phi_0, \alpha) - \underline{w}]$ and $\Delta' = y(1, \alpha) - \overline{w} - [y(\underline{\phi}(\underline{e}), \alpha + \eta) - \underline{w}]$. Notice that (24) implies that $\Delta > \Delta'$. Now observe, again mimicking the steps of Proposition 3, that

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

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

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

When $\delta_f = 1$, the above equation reduces to

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

which, after some rearrangements, can be rewritten as

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

Because $\Delta > \Delta'$, the last inequality is satisfied if

$$\underline{\phi}(\underline{e})(\alpha + \eta)(T - 2) > \phi_0\alpha + \phi_0\alpha(T - 2) \iff (T - 2)\{\underline{\phi}(\underline{e})(\alpha + \eta) - \phi_0\alpha\} > \phi_0\alpha,$$

which is true by (24) and the fact that $T \geq 3$ by assumption. $\qquad\square$

2. Non–IID Case

First notice that since $\psi = (\phi_0\alpha - \underline{\phi}\gamma)\Delta_y/\Delta = \alpha^2\phi_0(1 - \phi_0)\Delta_y/(1 - \phi_0\alpha)\Delta$, condition (9) can be rewritten as

$$(1 - \alpha)\gamma > \alpha\left\{ 1 - \phi_0\alpha + \alpha^2\phi_0(1 - \phi_0)\frac{\Delta_y}{\Delta} \right\}\frac{T - 1}{T - 2} + \alpha^2\phi_0(1 - \phi_0)\frac{\Delta_y}{(T - 2)\Delta}. \qquad (29)$$

Since $\Delta \leq \alpha(1 - \phi_0)\Delta_y$, a necessary condition for (29) to be satisfied is that $\alpha < 1/2$. When T is large, (29) reduces to

$$(1 - \alpha)\gamma > \alpha\left\{ 1 - \phi_0\alpha + \alpha^2\phi_0(1 - \phi_0)\frac{\Delta_y}{\Delta} \right\}, \qquad (30)$$

while (5) becomes

$$(1 - \alpha)\gamma < \alpha\left\{ 1 - \phi_0\alpha + \frac{(1 - \phi_0\alpha)\phi_0\alpha(\gamma - \alpha)\Delta_y}{\Delta + \phi_0\alpha[1 - (\gamma - \alpha)]\Delta_y} \right\}. \qquad (31)$$

50

It is immediate to see that the largest γ is, the easiest is for (30) to be satisfied—an increase in γ increases the value to the firm from inducing an age 1 worker to exert effort. We know from Subsection 6.1, however, that increasing γ makes it more difficult for (31) to be satisfied. In fact, when $\gamma = 1$, (31) can only be satisfied if $\alpha > 1/2$. This follows from the fact that $\Delta < \alpha(1 - \phi_0)\Delta_y$, and so, when $\gamma = 1$, a necessary condition for (31) is that

$$(1 - \alpha) < \frac{\alpha(1 - \phi_0\alpha)}{1 - \phi_0 + \phi_0\alpha} \iff (1 - \alpha)(1 - \phi_0) < \alpha(1 - \phi_0).$$

Thus, we need $\gamma < 1$ if (30) and (31) are to be jointly satisfied. Let $\gamma = \kappa\alpha$ and consider the case where Δ_w is small, so that Δ is close to $\alpha(1 - \phi_0)\Delta_y$. When $\Delta = \alpha(1 - \phi_0)\Delta_y$, the condition that (30) and (31) are jointly satisfied reduces to

$$1 < (1 - \alpha)\kappa < \underbrace{\frac{(1 - \phi_0\alpha)}{1 - \phi_0(\kappa - 1)\alpha}}_{g(\phi_0)}. \tag{32}$$

Hence, we need $\kappa \in (1/(1 - \alpha), 2)$ for (32) to hold—notice that $1/(1 - \alpha) < 2$ is only possible if $\alpha < 1/2$ and that $\kappa < 2$ is equivalent to $\gamma < 1$ when $\alpha < 1/2$. Now observe that $g'(\phi_0) \propto (\kappa - 2)\alpha$. So, a necessary condition for (32) to be satisfied is that

$$(1 - \alpha)\kappa < \frac{1 - \alpha^2}{1 - \alpha^2(\kappa - 1)} \iff q_\kappa(\alpha) = \alpha^2\kappa(\kappa - 1) + \alpha + \kappa + 1 > 0.$$

Since the discriminant of q_κ is $\sqrt{1 - 4\kappa(\kappa + 1)(\kappa - 1)}$, either q_κ has no real roots, in which case q_κ is always positive (since $q_\kappa(0) > 0$), or q_κ has two negative roots, in which case q_κ is positive in the interval $(0, 1/2)$. So, for every $\kappa \in (1/(1 - \alpha), 2)$, there exists $\phi^* < \alpha$ such that if $\phi_0 \in (\phi^*, \alpha)$, then (32) is satisfied. Notice that ϕ_0 close to α is only possible when Δ_w is small, for otherwise (A1) is violated.