Nonstationary Relational Contracts

Huanxing Yang*†
Ohio State University

July 22, 2009

Abstract

We develop a model of nonstationary relational contracts in order to study internal wage dynamics. Workers are heterogenous and each workers' ability is both private information and fixed for all time. Learning therefore occurs within employment relationships. The inferences, however, are confounded by moral hazard: the distribution of output is determined by both the worker's type and by his unobservable effort. Incentive provision is restricted by an inability to commit to long-term contracts. Relational contracts, which must be self-enforcing, must therefore be used. The wage dynamics in the optimal contract, which is pinned down by the tension between incentive provision and contractual enforcment, is intimately related to the learning effect.

JEL: C73, D82, J41, L14
Key Words: Relational Contracts; Learning; Tenure; Nonstationary; Wage Dynamics.

1 Introduction

Moral hazard pervades employment relationships. One way to alleviate the moral hazard problem is to use contingent contracts. However, the non-verifiability of workers' performance practically limits the usage of court-enforced contingent contracts. Nevertheless, if an employment relationship is repeated indefinitely, parties may rely on *relational contracts* that include both formal court-enforced and informal provisions. Since the informal provisions are not legally enforceable, it has to be self-enforcing – that is, each party should have no incentive to deviate from the informal provisions. This self-enforcing requirement imposes a contractual enforcement constraint on relational contracts.

There is a growing literature on relational contracts (Bull, 1987; Baker et.al, 1994; MacLeod and Malcomson, 1989, 1998; Levin, 2003). However, all of these papers focus on *stationary contracts* with contractual terms invariant to the length of relationships. In reality, contractual terms often

*Email: yang.1041@osu.edu. Mailing address: 405 Arps Hall, 1945 N. High St., Columbus, OH 43210.
†I am grateful to Jan Eeckhout, Dirk Kruger, Volker Nocke, George Mailath, Nicola Persico, Rafael Rob, Petra Todd, and especially to Steven Matthews and two anonymous referees for valuable comments. I also would like to thank the seminar participants at Iowa, OSU, LSE and SMU for their useful comments.

vary with the length of relationships. The wage-tenure effect – wage increases with tenure – has been a well established stylized fact (Mincer, 1974; Becker, 1975; Jovanovic and Mincer, 1981; Topel, 1991).

The main purpose of this paper is to develop a model of *nonstationary* relational contracts to account for wage dynamics. We do so by incorporating adverse selection (heterogenous workers), which creates a learning effect: firms learn the characteristics of workers as the relationships continue. The main message of the paper is that it is the interaction between incentive provision and contractual enforcement that ties wage dynamics to the learning effect, thus making wage increasing with tenure.

More specifically, we construct a repeated principal-agent model with the following key features. First, we model a labor market as a repeated matching market, with matches constantly reshuffled. This reshuffle is partly exogenous and partly endogenous, i.e., induced by workers' or firms' decision whether to continue the current relationship. Second, workers are heterogenous. Low type workers are inherently inept, while high type workers are potentially productive but have a moral hazard problem: they choose an unobservable effort based on incentives. A worker's type is *persistent* and is his own private information. Third, a worker's output is only observable to his current employer, not to the court nor to other potential employers. Finally, following the relational contract literature, we assume that firms cannot commit to long-term contracts; the only legally binding contracts are spot non-contingent contracts.

We focus on high-effort equilibria with high type workers exerting effort in every period. In each relationship, a relational contract specifies the conditions under which the relationship continues and wage as a function of tenure; and if either party is found to have deviated, the employment relationship is endogenously terminated. We study two kinds of contracts. With pooling contracts, firms offer both types of workers the same contract. With separating contracts, firms offer two contracts which are targeted at the two types of workers respectively, and let workers self-select at the beginning of employment. The important difference between pooling and separating contracts is that learning is completed in the first period of employment with separating contracts, while learning occurs gradually with pooling contracts.

With both kinds of contracts, wage must increase with tenure (at least across some periods) to provide an incentive to high type workers. However, the contractual enforcement constraint entails that wage cannot increase too fast with tenure; since otherwise senior workers would be less profitable than new workers, and firms will renege by terminating the current employment and hiring new workers. This tension between incentive provision and contractual enforcement drives the main results of the paper.

With pooling contracts, we study the conditions under which high-effort equilibria exist, and the wage dynamics under the optimal contract(s) that maximize firms' expected profits. We establish that high-effort equilibria with pooling contracts exist only if the proportion of low type workers is not too small nor too large. This implies that the presence of adverse selection might help alleviate moral hazard when firms are not able to commit to long-term contracts in a repeated matching market setting. Intuitively, the learning effect created by the presence of low type workers can alleviate the tension between incentive provision and contractual enforcement: the expected productivity of a worker increases with tenure due to the learning effect, so wage can increase with tenure without violating firms' no-reneging conditions.

If a high-effort equilibrium with pooling contracts exist, then there is a unique optimal pooling contract, under which the wage dynamics exhibits two salient features. First, wage is low and remains constant in earlier tenure periods, which can be interpreted as probation periods. Second, when wage begins to increase in later tenure periods, the wage increases are intimately related to the learning effect: the wage increase between two tenure periods exactly equals to the increase in the worker's expected productivity. Intuitively, since low type workers are more likely to have a short tenure, in order to minimize the informational rent to low type workers, firms try to "backload" wages: pay low wages in earlier tenure periods and use wage increases in later tenure periods to provide an incentive for high type workers. However, the contractual enforcement constraint limits firms' ability to backload wages: the wage increases cannot exceed the learning effect. As a result, in the optimal pooling contract the wage increases in later tenure periods are tied to the learning effect. One interesting point is that although learning is completely confined to current matches in our model, the wage increases are tied to the learning effect. This implies that even without market competition, wages being tied to workers' expected productivities can be generated by *internal wage dynamics*.

With separating contracts, high-effort equilibria exist only if there are enough L type workers. The optimal separating contract is again driven by the constraint imposed by contractual enforcement on firms' ability to backload wages. The wage dynamics in the optimal separating contract has a similar feature to that under the optimal pooling contract: wage is low and remains constant in earlier tenure periods. The difference is that in the optimal separating contract, wage increases at most in two tenure periods, and then wage remains constant afterwards. This difference comes from the fact that learning is completed in the first tenure period under separating contracts.

We then compare the optimal pooling contract and the optimal separating contract. When there is no unemployment, the optimal separating contract yields a higher discounted profit to firms than the optimal pooling contract does. The intuition for this result is as follows. Under both types of contracts, firms' ability to backload wages are more or less the same. Now what is left to compare is the benefit of separating and the cost of separating. The benefit of separating comes from faster learning of new workers' type with separating contracts. As a result, on average it takes less time for a firm to match with a high type worker with separating contracts. This fast screening effect clearly favors separating contracts. The cost of separating comes from the self-selection constraint: to induce the immediate revelation of types, in the first tenure period a low type worker has to be paid a high enough wage. In typical repeated adverse selection models (e.g., Laffont and Tirole, 1990), inducing immediate separation is very costly, since the discounted sum of informational rents in all future periods has to be paid in the first period. Surprisingly, in our model setup, the cost of separating turns out to be zero. The difference is that in previous models there is a single relationship, thus an agent gets zero rent after revealing his type. In contrast, in our model if a low type worker leaves the current relationship, next period he can match with another firm and get informational rents as well. Therefore, to induce a low type worker to reveal his type, a firm does not need to pay the discounted sum of informational rents in the current relationship. This leads to a zero cost of separating.

However, when there is unemployment, the cost of separating is positive and increasing with the unemployment rate. This is because with unemployment, it becomes costly for a low type worker to reveal his type and leave the current relationship, since it takes him several periods to find a new firm to match with. As a result, to induce immediate type revelation, firms have to pay more which leads to a positive separating cost. This implies that when the unemployment rate is high enough, the optimal pooling contract dominates the optimal separating contract.

The rest of the paper is organized as follows. The next subsection discusses the related literature. Section 2 sets up the model. Some preliminary analysis is offered in Section 3. Section 4 studies pooling contracts and Section 5 studies separating contracts. Section 6 compares pooling contracts to separating contracts. The last Section concludes. All the long proofs can be found in the Appendix.

Related literature As mentioned earlier, this paper differs from previous papers on relational contracts in that we study nonstationary contracts.[1] Actually, none of the previous papers incorporates adverse selection. Though Levin (2003) studies hidden information on the part of the worker, the worker's type is not persistent. Fuchs (2007) considers relational contracts when the principal has private information about the worker's performance, but again there is no persistent hidden information.

There are two existing non-contractual approaches to explain the wage-tenure effect. Neoclassical human capital theory (Becker, 1962; Hashimoto, 1981) argues that wage increases with tenure because individual workers' productivities increase with firm-specific human capital accumulation. The second one is Jovanovic's (1979) matching model with learning. In his model, within each individual match a firm and a worker symmetrically learn the quality of the match. Moreover, low quality matches endogenously break-up and only high quality matches remain. This learning effect combining with endogenous separation leads to the wage-tenure effect.

Our paper differs from Jovanovic (1979) in two aspects. First, our paper models the dynamic contracting problem explicitly. Second, in Jovanovic's model, wages being tied to the learning

[1] Except for MacLeod and Malcomson (1989, 1998), other papers on relational contracts restrict attention to one-pricincipal-one-agent settings, thus both parties' outside options are exogenously given.

effect is due to the market's competition for workers, as workers' past performance is commonly observed. In our model, learning is confined within the current matches, and the wage-tenure effect results from *internal* wage dynamics. Felli and Harris (1996) endogenize the wage determination in Jovanovic's model, but they confine to a setting in which two firms are competing for the service of a worker over time. In their model, the wage-tenure effect exists only if there is a learning externality: learning in the current match also provides information about the workers' productivity in the alternative match.[2]

In a pure moral hazard model, Lazear (1979) considers the increasing wage profile as a contractual device to prevent workers from shirking. However, he assumes that firms are able to commit to long-term contracts. Moreover, his model cannot pin down the wage dynamics, as there are many increasing wage-tenure profiles that can prevent workers from shirking. Harris and Holmstrom (1982) develop a model of wage dynamics based on symmetric learning and insurance concerns. Their model is more relevant in accounting for the relationship between wages and general working experience. Moreover, they also assume that firms can commit to long-term contracts.

In studying community games, Rob and Yang (2005) show that the presence of "bad" type agents can discipline opportunists to adopt cooperative behavior.[3] In their model, however, there are no contracts, hence has no implications about contract dynamics. Moreover, in their model players perfectly learn their partners' type after the first period of interaction. In our model monitoring is imperfect, so learning is gradual unless separating contracts are offered.[4]

2 The Model

There is a continuum of firms with measure 1 and each firm has exactly one job vacancy.[5] Correspondingly, there is a continuum of workers with measure 1. All workers and firms are risk-neutral, live forever and share the same discount factor δ. Time is discrete, indexed by $PT = 1, 2, ...$ In each period, workers and firms are matched to engage in production. Each existing match will continue in the next period with probability $\rho \in (0, 1)$, and break up with probability $1 - \rho$ for exogenous reasons. A match can also be dissolved endogenously if either party in the current match decides to leave the match. All the agents in dissolved matches enter into the unmatched pool, and they are randomly paired at the beginning of the next period. The time line will be specified

[2] Burdett and Coles (2003) study the wage-tenure effect in a job-search framwork. Their main foucs is to separate the wage-tenure effect from the wage growth due to searching for better jobs. And they also assume that firms can commit to long-term contracts.

[3] Mailath and Samuelson (2001) establish that reputational concerns can also be generated by a high type firm's incentive to differentiate itself from low types. But their model focuses on reputation and only studies the one-firm case.

[4] This paper is also loosely related to the following papers. For adverse selection in labor markets , see Greenwald (1986). For information asymmetry between the current employer and alternative firms, see Waldman (1984b) and Bernhardt (1995). For symmetric and public learning in labor market contexts, see Holmstrom (1999) and Farber and Gibbons (1996).

[5] The main results of the paper still go through as long as each firm has a finite number of job vacancies.

shortly. Note that workers and firms are of equal measure, so each agent is guaranteed a match at the beginning of each period.[6]

The stage output y for a match is either 0 or 1, and the value of output y is vy, with v being normalized to 1. Workers are of two types: high type H and low type L. The measure of L type workers is $\beta \in [0, 1]$, and H type is of measure $1 - \beta$. A worker's type is fixed for all time and is his own private information. Two types of workers differ in productivity: H type workers have an option to choose a high effort $\overline{e} > 0$ or a low effort 0; L type workers are inept and can only exert low effort 0.[7] The cost of effort $\overline{e}$ is c and the cost of effort 0 is 0. A worker's effort is not observable.

Output y only depends on the effort level. Specifically,

$$Pr\{y = 1|e\} = \begin{cases} 1 & \text{if } e = \overline{e} \\ p \in (0, 1) & \text{if } e = 0 \end{cases} .$$

This assumption implies that monitoring is imperfect, in the sense that output does not perfectly reveal a worker's effort.[8] We assume $1 - p > c$, so the efficient action for H workers is $\overline{e}$. A worker's output y is observable to the worker and his current employer, but not to the court or to other market participants. Thus, court-enforced contracts that are contingent on y are not feasible, and there is no information flow between matches. We assume that a worker's previous employment history is not observable to firms,[9] and a firm's previous employment history is not observable to workers either. If a worker is not employed in one period, he gets a reservation utility 0 in that period regardless of his type. Since $p > 0$, it is efficient for both types of workers to get employed in each period. Similarly, if a firm does not employ a worker, its profit in that period is 0.

Firms are not able to commit to long-term contracts. The only legally binding contracts are spot contracts, which specify a fixed wage payment w_t. Here t denotes tenure period (starting from 1), which is the periods that a worker has been matched with the current firm. A firm may also offer its worker a discretionary bonus b_t in tenure period t, for which the firm promised to pay if and only if $y_t = 1$. At the beginning of employment, a firm also propose to its worker how the payments are going to evolve as the relationship continues. We name the proposed payment plans $\{w_t, b_t\}$ as contracts. There are two kinds of payment plans: *pooling contracts* C^p and *separating contracts* C^s. We denote a pooling contract as $\{w_t, b_t\}$, under which both types of workers have the same payment plan. In separating contracts, a firm offers two contracts and let the worker self-select in tenure period 1. Specifically, in the contract designed for L type workers, a fixed wage w_L is offered in tenure period 1, and the worker is fired after tenure period 1 regardless of the output. In the

[6] If the measures of workers and firms are not equal, then the long side of the market will have matching friction. For this direction of research, see MacLeod and Malcomson (1989, 1998).

[7] An interpretation is that even if a L type worker exerts the high effort $\overline{e}$, the distribution of output is the same as those when he exerts effort 0.

[8] This assumption also implies that a H type worker who exerts 0 effort is the same as an L type worker in terms of productivity.

[9] This is a simplifying assumption, which makes workers in the unmatched pool homogenous in appearence.

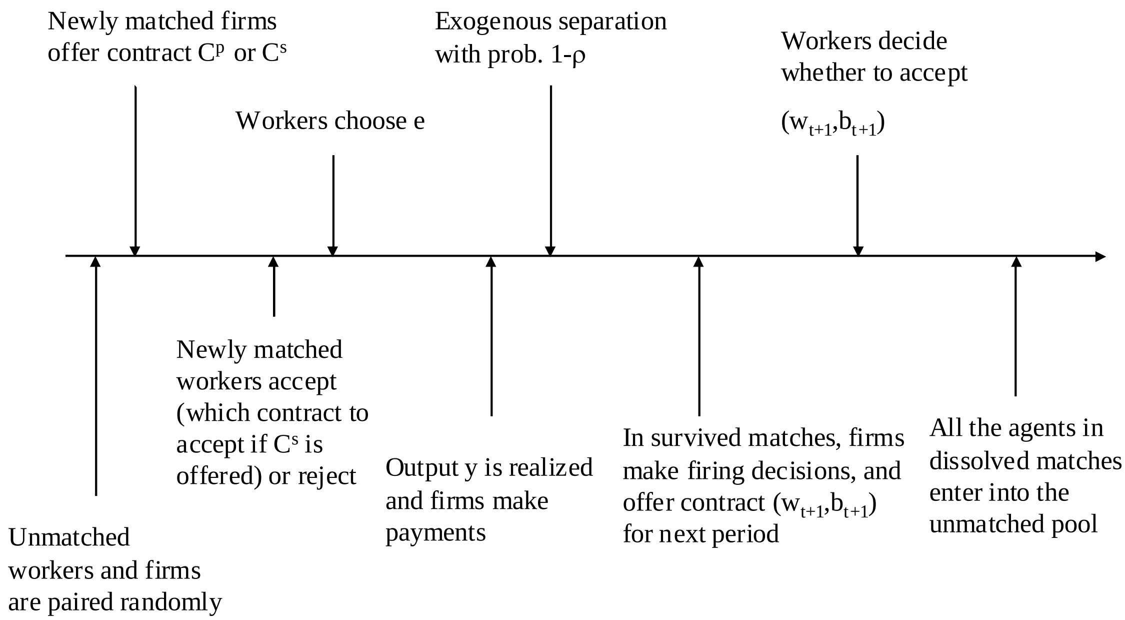

Figure 1: Time Line of a Typical Period

contract designed for H type workers, the payment plan evolves according to $\{w_t^s, b_t^s\}$. In short, we denote a separating contract as $(w_L, \{w_t^s, b_t^s\})$. Let W_t be the total wage payment actually made in tenure period t. Finally, workers are subject to limited liability, that is, $w_t \geq 0$ for all t.

Figure 1 specifies the time line within a period. At the beginning of a period, unmatched workers and unmatched firms are paired randomly. In each newly formed match, either a pooling contract C^p or a separating contract C^s is offered. For a pooling contract, a spot contract (w_1, b_1) is offered and $\{w_t, b_t\}$ is proposed; and the worker decides whether to accept the offer. For a separating contract, w_L is offered targeting at type L workers, and a spot contract (w_1^s, b_1^s) is offered and $\{w_t^s, b_t^s\}$ is proposed targeting at H type workers; and the worker decides which contract to accept or to reject both. If the worker rejects both offers, he leaves the match and collects reservation utility 0 in that period. Then among all employed workers, H type workers choose their effort level. Afterwards, output y (in each match) is realized and workers are paid. Then exogenous separation occurs to existing matches with probability $1 - \rho$. In each survived match, firms make firing decisions. If a firm wants to retain the worker, it offers a spot contract (w_{t+1}, b_{t+1}) for next period; and the worker decides whether to accept the offer. A match is dissolved endogenously if the firm fires the worker or the worker rejects the firm's offer. All the agents in dissolved matches enter into the unmatched pool. And then the next period begins.

3 Preliminary Analysis

We adopt symmetric perfect public equilibrium (SPPE) as our solution concept. By symmetry we mean that all firms adopt the same strategy and each type of workers also adopt the same strategy. Public strategies require that each agent's strategy only depend on the public history within the current relationship, since the previous employment history is not observable.[10] Denote $d_t \in \{0,1\}$ as a tenure-period t worker's decision regarding whether to accept its current firm's offer. Denote $d_1 \in \{0, L, H\}$ as a tenure-period 1 worker's decision regarding whether and which contract to accept if the firm offers a separating contract. Denote $k_t \in \{0,1\}$ as a firm's decision regarding whether to fire its current worker who is in tenure period t. The public history of a relationship that has lasted for t tenure periods can be denoted as $h^t = (\{w_t, b_t\}, w_1, b_1, y_1, W_1, ..., w_t, b_t, y_t, W_t)$ with pooling contracts, and $h^t = (w_L, \{w_t^s, b_t^s\}, w_1^s, b_1^s, d_1, y_1, W_1, ..., w_t^s, b_t^s, y_t, W_t)$ with separating contracts.

A (behavior) strategy σ^f for a firm in tenure period 1 consists of: a spot contract offer $\{(w_1, b_1), (w_L, (w_1^s, b_1^s)\}$, a proposed payment plan $\{w_t, b_t\}$ or $\{w_t^s, b_t^s\}$, decision of payment, decision of firing, and next period's spot contract offer.[11] σ^f in tenure period $t \geq 2$ consists of $(h^{t-1}, w_t, b_t, y_t) \to W_t$ $(\geq w_t)$ (payment decision), $(h^{t-1}, w_t, b_t, y_t, W_t) \to k_t$ (firing decision) and $(h^{t-1}, w_t, b_t, y_t, W_t) \to (w_{t+1}, b_{t+1})$ (spot contract offer for tenure period $t+1$). A strategy σ^H for a H type worker in tenure period 1 consists of: $((w_1, b_1), \{w_t, b_t\}) \to d_1$ or $(w_L, (w_1^s, b_1^s), \{w_t^s, b_t^s\}) \to d_1$, effort choice, and the decision whether to accept the contract offer in the next period. σ^H in tenure period $t \geq 2$ consists of: $(h^{t-1}, w_t, b_t) \to e_t$ (effort choice) and $(h^t, w_{t+1}, b_{t+1}) \to d_t$ (quit decision). A strategy σ^L for a L type worker is similarly defined as σ^H except that L type workers have no effort choice.

A relational contract, which is a complete plan for a relationship, consists of a strategy profile $\sigma = (\sigma^H, \sigma^L, \sigma^f)$. Denote $\phi(h^{t-1})$ as a firm's belief that its worker is of H type, given history h_{t-1}.

Definition 1 *A relational contract σ and $\phi(h^{t-1})$ consists of a SPPE of the repeated matching game if: (i) σ^H and σ^L are best responses to σ^f after any history h^{t-1} and σ^f is a best response to σ^H and σ^L given $\phi(h^{t-1})$ after any history, (ii) $\phi(h^{t-1})$ are consistent with σ^H and σ^L and updated using Bayes' rule, whenever possible.*

There is always a trivial equilibria in which H type workers always exert 0 effort, firms always offer 0 wage, workers always accept nonnegative wage offers, and firms always fire their current workers. Given that H type workers always exert 0 effort, firms have no incentive to offer positive wages and firing decisions become irrelevant. In this equilibrium of a zero-wage contract, each firm gets a per-period profit p. Such non-reputational equilibrium is not the focus of this paper. Recall

[10] This implies that each relationship is played out in the same way in equilibrium.

[11] The proposed payment plan $\{w_t, b_t\}$ has the role of pinning down the worker's expectation about future wages within the relationship. This allows firms to potentially offer different contracts.

that the efficient outcome is for all workers to get employed and H type workers to exert high effort in each period. We call equilibria with this outcome as *high-effort equilibria*, which are the primary focus of this paper.

A necessary condition for a high-effort equilibrium is that H type workers should have an incentive to exert effort $\bar{e}$ in each period (*no-shirking conditions*). To effectively prevent shirking, we restrict attention to the following trigger strategy: a firm retains its worker if and only if the worker produces $y_t = 1$ in each previous tenure period and fires the worker immediately if $y_t = 0$.

Given that only fixed-wage spot contracts can be legally enforced, another necessary condition for a high-effort equilibrium is that firms have no incentive to renege (*no-reneging conditions*). Specifically, there are three kinds of reneging. First, firms can renege on bonus b_t, by not paying b_t when $y_t = 1$. Second, a firm's spot contract offer in tenure period $t \geq 2$ can be different from the proposed payment plan $\{w_t, b_t\}$, which was agreed upon by both parties at the beginning of employment. Third, a firm can fire a worker even if the worker always produces $y_t = 1$ in the relationship (recall that the firms' trigger strategy specifies that a worker is retained if he always produces $y_t = 1$ in the relationship). Intuitively, if senior workers are less profitable than new workers, then firms may fire senior workers regardless of their performance. To effectively deter firms' reneging, we restrict attention to the following trigger strategy: a worker stays in his current firm if and only if the firm always pay bonus b_t and the spot contracts have always followed the proposed payment plan $\{w_t, b_t\}$, otherwise he quits immediately.

We focus on trigger strategies because they provide the severest punishment for the deviating party, thus making high-effort equilibria easier to be sustained. A trigger strategy is clearly a best response for firms since only L type workers produce $y = 0$ on the equilibrium path. A trigger strategy is also a best response for workers (both types), if we assume that workers hold the most pessimistic belief off the equilibrium path. More specifically, once the actual payments $\{W_t\}$ deviate from the proposed payment plan $\{w_t, b_t\}$, the worker holds the belief that the firm will offer the lowest possible wage (0) in all future spot contracts if the relationship continues. Given this belief, it is a best response for the worker, regardless of his type, to quit immediately after the firm deviates.

The next simplifying step is that, without loss of generality, we can restrict attention to fixed wage contracts $\{w_t\}$. That is, it is without loss to set $b_t = 0$ for all t. The underlying reason is that for any contract that includes bonus payments, $\{w_t, b_t\}$, there is always a corresponding payoff-equivalent fixed wage contract $\{w'_t\}$.[12] Therefore, we will only consider fixed wage contracts $\{w_t\}$ hereafter. Specifically, a pooling contract is denoted as $\{w_t\}$, and a separating contract is denoted as $(w_L, \{w^s_t\})$.

[12] A formal proof of this claim can be found in an earlier version of the paper. The idea is that any bonus b_t can be incorporated into the fixed wage payment of the next tenure period w_{t+1} without affecting the expected payoff for each party. Levin (2003) establishes that focusing on stationary bonus contracts is without loss of generality. The difference is that in his model there is no persistent type, which essentially yields a stationary environment in terms of contracting.

With fixed wage contracts, we do not need to worry about firms' reneging on bonus payments. Note that the workers' trigger strategy effectively deters firms' reneging of the second category: a firm's spot contract offers will always follow the proposed $\{w_t\}$ if it wants to retain the worker. Therefore, we only need to worry about firms' reneging of the third category. As a result, firms' no-reneging conditions boil down to the condition that firms always have an incentive to retain a worker who always produces $y_t = 1$ in a relationship.

3.1 Pooling Contracts

We first study pooling contracts. Under trigger strategies, tenure period t is a sufficient statistic of the previous public history. A worker in tenure period t means that $y_j = 1$ for all $j \leq t-1$ in the current firm, and the wage offers have followed $\{w_t\}$ so far. At any physical time PT, H type workers will be in different tenure periods because of exogenous separation. Type L workers are also in different tenure periods because of imperfect monitoring. Define x_t (β_t) as the population of H (L) type workers who are in tenure period t. We restrict attention to a *stationary state*, that is, the distributions of the types of workers in different tenure periods $\{x_t\}$ and $\{\beta_t\}$ are invariant to physical time PT.[13] In the stationary state, $\beta_t = (\rho p)^{t-1}\beta_1$. Summing up β_t and using the fact that the total population of L type workers is β, we get $\beta_1 = (1-\rho p)\beta$. Similarly, one can get $x_t = \rho^{t-1}x_1$ and $x_1 = (1-\rho)(1-\beta)$ in the stationary state.[14]

Definition 2 *A high-effort (trigger strategy) equilibrium with pooling contract $\{w_t\}$ exists if: (i) all the workers accept offers in tenure period* 1*, and all firms have incentives to employ new workers (participation constraints), (ii) H type workers will exert high effort $\overline{e}$ in each period (no-shirking conditions), (iii) firms always retain a worker who always produces $y_t = 1$ in the relationship (no-reneging conditions), (iv) no firm has an incentive to unilaterally deviate to offer another pooling contract different from $\{w_t\}$ or a zero-wage contract, (v) no firm has an incentive to unilaterally deviate to offer a separating contract.*

The equilibrium requirements (iv) and (v) will be discussed at the end of this subsection. To proceed, we first ignore requirement (v) (which will be considered in Section 6), and we call a relational pooling contract that satisfies requirement (i)-(iv) as a *quasi high-effort equilibrium with pooling contracts.*

One important observation is that a firm learns its worker's type as tenure period t increases. Under trigger strategies, a firm's initial belief in tenure period t, $\phi(h^{t-1})$, can be simply denoted as ϕ_t. Recall the assumption that workers' previous employment history is not observable. This leads to a common initial belief ϕ_1 about all workers in the unmatched pool. Specifically,

$$\phi_1 = \frac{x_1}{x_1 + \beta_1} = \frac{(1-\rho)(1-\beta)}{(1-\rho)(1-\beta) + (1-\rho p)\beta}. \qquad (1)$$

[13] We assume that the economy settles into the stationary state in the first physical time period.
[14] On the equilibrium path, H type workers turn over only because of exogenous separation.

Note that ϕ_1 is decreasing in β. Firms update their beliefs according to the following Bayes' rule

$$\phi_t = \frac{\phi_1}{\phi_1 + p^{t-1}(1-\phi_1)}; \text{ and } \phi_t + p(1-\phi_t) = \frac{\phi_t}{\phi_{t+1}}. \tag{2}$$

Note that ϕ_t only depends on tenure period t and it is updated gradually since $p \in (0,1)$.

Define U_t (U_t^L) as the equilibrium discounted payoff of a type H (L) worker who is in tenure period t. The recursive value functions are:

$$U_t = (w_t - c) + \delta[\rho U_{t+1} + (1-\rho)U_1]; \tag{3}$$

$$U_t^L = w_t + \delta[p\rho U_{t+1}^L + (1-p\rho)U_1^L]. \tag{4}$$

Define U_t^d as the discounted payoff of a type H worker who is in tenure period t and shirks in that period,

$$U_t^d = w_t + \delta[p\rho U_{t+1} + (1-p\rho)U_1]. \tag{5}$$

Similarly, define V_t as a firm's equilibrium discounted profit who currently matches with a tenure-period t worker:

$$V_t = (\frac{\phi_t}{\phi_{t+1}} - w_t) + \delta[\rho\frac{\phi_t}{\phi_{t+1}}V_{t+1} + (1-\rho\frac{\phi_t}{\phi_{t+1}})V_1]. \tag{6}$$

Note that a worker's expected output in tenure period t is $\phi_t + p(1-\phi_t)$. By (2), $\phi_t + p(1-\phi_t) = \frac{\phi_t}{\phi_{t+1}}$. Let V_t^d be a firm's discounted profit who currently matches with a tenure-period t worker, and reneges in that period. As discussed earlier, the only reneging we need to consider is that the firm fires its worker who has produced $y_j = 1$ for any $j \leq t$. Thus,

$$V_t^d = (\frac{\phi_t}{\phi_{t+1}} - w_t) + \delta V_1. \tag{7}$$

Note that all the value functions are nonstationary due to the gradual learning effect.

Now the no-shirking conditions can be explicitly written as:

$$\begin{aligned} & U_t - U_t^d \geq 0 \;\Leftrightarrow \delta\rho(1-p)[U_{t+1} - U_1] \geq c \text{ for any } t \geq 1 \\ \Leftrightarrow\; & U_t - U_1 \geq \widehat{c} \;\text{ for all } t \geq 2, \text{ where } \widehat{c} \equiv \frac{c}{\delta\rho(1-p)}. \end{aligned} \tag{8}$$

Inequality (8) says that to prevent H type workers from shirking, the equilibrium discounted payoff in later tenure periods relative to that in the first tenure period has to be big enough. This implies that in general w_t has to be increasing in t. Similarly, firm's no-reneging conditions can be written as:

$$\begin{aligned} & V_t - V_t^d \geq 0 \Leftrightarrow V_{t+1} - V_1 \geq 0 \text{ for all } t \\ \Leftrightarrow\; & V_t - V_1 \geq 0 \text{ for all } t \geq 2. \end{aligned} \tag{9}$$

According to (9), to prevent a firm from reneging, its equilibrium discounted payoff when matched with a senior worker should be bigger than that when matched with a new worker. That is, senior workers cannot be less profitable than new workers. Note that both no-shirking conditions and no-reneging conditions consist of an infinite number of constraints. To ease exposition, we call a contract $\{w_t\}$ that satisfies the no-shirking conditions (8) and no reneging conditions (9) as a self-enforcing pooling contract.

Workers' participation constraints require $U_t \geq 0$ and $U_t^L \geq 0$ for all t. However, given that H type workers can always mimic L type workers, the no-shirking conditions (8) imply $U_t \geq U_t^L$ for all t. Thus, workers' participation constraints boil down to $U_t^L \geq 0$ for all t. But given that $w_t \geq 0$ due to limited liability, by (4) $U_t^L \geq 0$ for all t is always satisfied. Firms' participation constraints require $V_t \geq 0$ for all t. But given the no-reneging conditions (9), $V_1 \geq 0$ is sufficient.

Here we discuss the equilibrium requirements (iv) and (v) in Definition 2 in more detail. Recall that firms are free to propose contracts for new relationships. Thus a (symmetric) high-effort equilibrium with pooling contract $\{w_t\}$ requires that: given that all the other firms offer contract $\{w_t\}$, no firm has an incentive to unilaterally deviate to offer another contract different from $\{w_t\}$. A firm can deviate to three kinds of contracts. First, a firm can deviate to offer the zero-wage contract and get a stage profit p in each period. To prevent this deviation, the discounted payoff V_1 under pooling contract $\{w_t\}$ must be bigger than $p/(1-\delta)$. Second, a firm can deviate to another self-enforcing pooling contract $\{w_t'\}$, for which both types of workers accept and H type workers always exert high effort. Finally, a firm can deviate to offer a separating contract.

To prevent unilateral deviation to another self-enforcing pooling contract $\{w_t'\}$, the discounted payoff V_1 under $\{w_t\}$ must be bigger than the discounted payoff V_1' under $\{w_t'\}$. This means that if a quasi high-effort equilibrium with pooling contracts exists, the associated equilibrium pooling contract(s) must be a solution to the following programming problem: maximize V_1 subject to the no-shirking conditions (8) and no reneging conditions (9). We call such contracts as *optimal pooling contracts*. Based on the above analysis, we have the following lemma.

Lemma 1 *A quasi high-effort equilibrium with pooling contracts exists if and only if: (i) the programming problem maximizing V_1 subject to (8) and (9) has a solution, (ii) V_1 under the solution is bigger than $p/(1-\delta)$. If the equilibrium exists, the equilibrium contract must be an optimal pooling contract.*

3.2 Separating Contracts

With separating contracts $(w_L, \{w_t^s\})$, L type workers are always in the unmatched pool, since they choose contract w_L in tenure period 1 and are fired immediately. In the stationary state, $1-\rho$ proportion of H type workers are in the unmatched pool due to exogenous separation. Thus the percentage of H type workers in the unmatched pool, λ, is:

$$\lambda = \frac{(1-\rho)(1-\beta)}{\beta + (1-\rho)(1-\beta)}. \tag{10}$$

Note that for the same β, in the stationary state the percentage of H type workers in the unmatched pool is lower under separating contracts than that under pooling contracts.

Definition 3 *A high-effort (trigger strategy) equilibrium with separating contract* $(w_L, \{w_t^s\})$ *exists if: (i) all the workers accept offers in tenure period* 1*, and all firms have incentives to employ new workers (participation constraints), (ii) in tenure period* 1*,* L *type workers choose contract* w_L *and* H *type workers choose contract* $\{w_t^s\}$ *(self-selection conditions), (iii)* H *type workers will exert high effort* $\overline{e}$ *in each period (no-shirking conditions), (iv) firms have an incentive to retain workers who chooses contract* $\{w_t^s\}$ *and always produces* $y_t = 1$ *in the relationship. (no-reneging conditions), (v) no firm has an incentive to deviate to offer another separating contract different from* $(w_L, \{w_t^s\})$ *or the zero-wage contract, (vi) no firm has an incentive to deviate to offer a pooling contract.*

Again, we will ignore requirement (vi) for the moment (which will be considered in Section 6), and we call a relational separating contract that satisfies requirement (i)-(v) as a *quasi high-effort equilibrium with separating contracts.*

Unlike pooling contracts, with separating contracts firms learn the type of new workers in the first tenure period. But now two self-selection conditions are added. Define U_t^s (U_t^L) as a H type's (L type chooses contract $\{w_t^s\}$) expected discounted payoff who is currently in tenure period t, U_L as a type L's equilibrium discounted payoff, and U_t^{sd} as the discounted payoff of a type H worker who is in tenure period t and shirks in that period. Define V_t^s as a firm's expected discounted profit who is currently matched with a tenure period t H type worker, V_L as a firm's expected discounted profit who currently matches with a type L worker in tenure period 1, and V_N as a firm's expected discounted profit who is in the unmatched pool (before it matches with a new worker). The value functions are as follows:

$$\begin{aligned}
U_t^s &= (w_t^s - c) + \delta[\rho U_{t+1}^s + (1-\rho)U_1^s], \\
U_t^{sd} &= w_t^s + \delta[\rho p U_{t+1}^s + (1-\rho p)U_1^s], \\
U_t^L &= w_t^s + \delta[\rho p U_{t+1}^L + (1-\rho p)U_L], \\
U_L &= w_L + \delta U_L, \\
V_t^s &= (1 - w_t^s) + \delta[\rho V_{t+1}^s + (1-\rho)V_N], \\
V_L &= (p - w_L) + \delta V_N, \\
V_N &= \lambda V_1^s + (1-\lambda)V_L.
\end{aligned}$$

Again, as long as firms have no incentive to deviate to the zero-wage contract, $V_N \geq p/(1-\delta)$, we do not need to worry about firms' and workers' participation constraints. The self-selection constraints are written as: $U_1^s \geq w_L + \delta U_1^s$ (H type has no incentive to choose the L contract) and $U_L \geq U_1^L$ (L type has no incentive to choose the H contract). The no-shirking conditions become $U_t^s \geq U_t^{sd}$ for any t, and the no-reneging conditions are $V_t^s \geq V_N$ for any t. After some manipulation, the last four constraints become:

$$(1-\delta\rho)\sum_{t=1}^{\infty}(\delta\rho)^{t-1}w_t^s \geq w_L + c; \tag{11}$$

$$(1-\delta\rho p)\sum_{t=1}^{\infty}(\delta\rho p)^{t-1}w_t^s \leq w_L; \tag{12}$$

$$\text{For any } j \geq 2,\ \sum_{t=j}^{\infty}(\delta\rho)^{t-j}w_t^s - \sum_{t=1}^{\infty}(\delta\rho)^{t-1}w_t^s \geq \widehat{c}\ ; \tag{13}$$

$$\text{For any } j \geq 2,\ \sum_{t=j}^{\infty}(\delta\rho)^{t-j}(1-w_t^s) - \frac{1}{1-\delta\rho(1-\lambda)} \tag{14}$$

$$\times\{\lambda\sum_{t=1}^{\infty}(\delta\rho)^{t-1}(1-w_t^s) + (1-\lambda)(p-w_L)\} \geq 0.$$

To ease exposition, we call a contract $(w_L, \{w_t^s\})$ that satisfies conditions (11)-(14) as a self-enforcing separating contract. Similar to optimal pooling contracts, we call separating contracts $(w_L, \{w_t^s\})$ that maximize V_N subject to (11)-(14) as *optimal separating contracts.* Based on the above analysis, we have the following Lemma, which is similar to Lemma 1 with pooling contracts.

Lemma 2 *A quasi high-effort equilibrium with separating contracts exists if and only if: (i) the programming problem maximizing V_N subject to (11)-(14) (11)-(14) has a solution, (ii) under the solution(s) $V_N \geq p/(1-\delta)$. If the equilirbium exists, the equilibrium contract must be an optimal separating contract.*

The rest of the paper will focus on the following issues. The first one is to identify the conditions under which quasi high-effort equilibria exist under pooling contracts and separating pooling contracts respectively. The second one is to characterize optimal pooling contracts and optimal separating contracts. Finally, we will compare optimal pooling contracts and optimal separating contracts, and identify what type of contracts will be adopted in high-effort equilibria.

4 Equilibrium with Pooling Contracts

We have two major difficulties in our analysis. First, equilibrium conditions (8) and (9) are involved with two sets of an infinite number of constraints. Second, there is too much freedom in the design of contracts, which consist of a whole (infinite) sequence of wages. In the next subsection, we first show that without loss we can focus on some certain class of contracts.

4.1 A Certain Class of Contracts

Inspecting (8), we observe that workers' no-shirking conditions require $\{w_t\}$ be strictly increasing at least across some periods. This observation leads us to nondecreasing contracts.

Definition 4 *A contract* $\{w_t\}$ *is said to be* ***nondecreasing*** *if* w_t *is nondecreasing in* t.

Theorem 1 *If a pooling contract* $\{w_t\}$ *is self-enforcing, then there is another self-enforcing pooling contract* $\{w'_t\}$ *such that: (i)* w'_t *is nondecreasing in* t*, (ii) the no-shirking conditions (8) do not bind for any* $t \geq 2$*, (iii) firms' expected (discounted) profits are the same under two contracts,* $V_1 = V'_1$.

The proof is by construction. If $\{w_t\}$ is strictly decreasing somewhere, we can redesign $\{w_t\}$ by making wages constant in that range (decrease wages in earlier tenure periods and increase wages in later tenure periods) without affecting V_1. By Theorem 1, without loss of generality we can focus on nondecreasing contracts. By the no-shirking conditions, any self-enforcing and nondecreasing contracts must have the following form: there is a $T \geq 1$ such that $w_t = 0$ for $t < T$ and $w_t > 0$ for $t \geq T$. That is, T is the first tenure period such that wage is strictly positive. Observing firms' no-reneging conditions (9), we see that $\{w_t\}$ cannot increase too fast, otherwise firms will find new workers more profitable than old workers. This leads us to the following class of contracts.

Definition 5 *Consider a nondecreasing contract* $\{w_t\}$*, and let* T *be the first tenure period such that wage is strictly positive. Let* π_t *be a firm's expected profit in tenure period* t*:* $\pi_t \equiv \frac{\phi_t}{\phi_{t+1}} - w_t$*. This contract is said to be* ***quasi-monotonic*** *if either (i) for any* $t \geq T$*,* $\pi_{t+1} \geq \pi_t$*, or (ii) for any* $t > T$*,* $\pi_{t+1} \geq \pi_t$*, and* $\pi_T > \pi_{T+1}$ *and* $\pi_T \geq (1-\delta)V_1$.

Theorem 2 *Suppose there is a nondecreasing and self-enforcing contract* $\{w_t\}$*, then there is another self-enforcing contract* $\{w'_t\}$ *such that: (i)* $\{w'_t\}$ *is quasi-monotonic, (ii) the no-shirking conditions do not bind for any* $t \geq 2$*, (iii) firms' expected (discounted) profits are the same under two contracts,* $V_1 = V'_1$.

The results of Theorem 1 and 2 are intuitive. The no-shirking conditions are easier to satisfy when wages are nondecreasing in tenure, and firms' no-reneging conditions are easier to satisfy when firms' stage profits are nondecreasing in tenure (except from tenure period $T-1$ to $T+1$). Combining Theorem 1 and Theorem 2, we have the following corollary.

Corollary 1 *If there is a self-enforcing pooling contract* $\{w_t\}$*, then there is a self-enforcing quasi-monotonic contract* $\{w'_t\}$*. Moreover, firms' expected (discounted) profits are the same under two contracts,* $V_1 = V'_1$*, and the no-shirking conditions do not bind for any* $t \geq 2$ *under* $\{w'_t\}$.

The following lemma specifies self-enforcing quasi-monotonic contracts.

Lemma 3 *Under a quasi-monotonic contract $\{w_t\}$, the no-shirking conditions (8) become*

$$U_2 - U_1 > \widehat{c} \Leftrightarrow \sum_{j=1}^{\infty} (\delta\rho)^{j-1}(w_{j+1} - w_j) \geq \widehat{c}, \tag{15}$$

and the no-reneging conditions (9) become: $V_T \geq V_1$ ($V_{T+1} \geq V_1$) if $\pi_{t+1} \geq \pi_t$ for any $t \geq T$ (if $\pi_{t+1} \geq \pi_t$ for any $t > T$ and $\pi_T > \pi_{T+1}$).

Actually, we can go one step further showing that optimal pooling contracts must be quasi-monotonic.

Theorem 3 *(i) If a quasi-monotonic contract is self-enforcing, but the no-shirking condition (15) is not binding, then it cannot be optimal, (ii) If a self-enforcing contract $\{w_t\}$ is not quasi-monotonic, then it cannot be optimal.*

Proof. (i) Suppose a quasi-monotonic contract $\{w_t\}$ satisfies the no-shirking and no-reneging conditions. Moreover, $U_2 - U_1 > \widehat{c}$. Let j be a tenure period such that $w_{j+1} > w_j$ (such a j must exists, otherwise (15) is violated, and $j \geq T$). The idea is to find another self-enforcing quasi-monotonic contract which yields a strictly larger V_1. Specifically, construct another contract $\{w'_t\}$ as follows: $w'_t = w_t$ for any $t \leq j$, $w'_t = w_t - \varepsilon$ for any $t > j$, where $\varepsilon > 0$ is very small. By construction, $\{w'_t\}$ is also quasi-monotonic. By the construction, it is easy to see that $V'_j > V_j$ and $V'_1 > V_1$.

Now what is left to be shown is that $\{w'_t\}$ is self-enforcing. $\{w'_t\}$ clearly satisfies the no-shirking condition (15). To see this, note that compared with $\{w_t\}$, under $\{w'_t\}$ only the wage increase from j to $j+1$ is reduced by ε. From (15), we can see that that $U_2 - U_1 > \widehat{c}$ implies that $U'_2 - U'_1 \geq \widehat{c}$. The next step is to show that $\{w'_t\}$ satisfies the no-reneging conditions. Since $\{w'_t\}$ is quasi-monotonic, we only need to show $V'_T \geq V'_1$. Note that V'_1 increases because V'_T increases. Given that

$$\{1 - \delta \sum_{t=1}^{T-1} (\delta\rho)^{t-1} \frac{\phi_1}{\phi_t} [1 - \rho \frac{\phi_t}{\phi_{t+1}}]\} V_1 = \sum_{t=1}^{T-1} (\delta\rho)^{t-1} \frac{\phi_1}{\phi_t} \pi_t + (\delta\rho)^{T-1} \frac{\phi_1}{\phi_T} V_T$$

and the same relationship holds between V'_T and V'_1, we have $V'_1 - V_1 < V'_T - V_T$. Since $V_T \geq V_1$ ($\{w_t\}$ satisfies the no-reneging conditions), we must have $V'_T \geq V'_1$. This proves part (i).

Part (ii) is directly implied by part (i) and Corollary 1. ■

By Lemma 1 and 3 and Corollary 1, we have the following proposition.

Proposition 1 *A quasi high-effort equilibrium with pooling contracts $\{w_t\}$ exists if and only if the following programming problem [PP] has a solution: maximize V_1 subject to: (i) $\{w_t\}$ is quasi-monotonic, (ii) the no-shirking condition (15) holds, (iii) $V_T \geq V_1$ and $V_{T+1} \geq V_1$ (ICF), (iv) $V_1 \geq p/(1-\delta)$. Moreover, optimal pooling contracts must be quasi-monotonic.*

4.2 Optimal Pooling Contracts

By Proposition 1, we can safely focus on quasi-monotonic contracts in searching for optimal pooling contracts. Inspecting the programming problem in Proposition 1, two observations are in order. First, $w_1 = 0$ in optimal contracts, since what matters for the no-shirking and the no-reneging conditions is the wage increases Δw_t. Second, in optimal contracts the no-shirking condition (15) must be binding (see the proof of Theorem 3).

Lemma 4 *If the programming problem [PP] has a solution, it also has a solution of the following form: (i) $\pi_t = \pi_{T+1}$ for any $t > T+1$, (ii) $V_{T+1} = V_1$ and*

$$\pi_{T+1} = (1-\delta)V_1 = \frac{\sum_{t=1}^{T}(\delta\rho)^{t-1}\frac{\phi_1}{\phi_t}\pi_t}{\sum_{t=1}^{T}(\delta\rho)^{t-1}\frac{\phi_1}{\phi_t}}. \tag{16}$$

Moreover, optimal pooling contracts must have the above form.

Proof. Part (i). Suppose there is a self-enforcing quasi-monotonic contract $\{w_t\}$ in which $\pi_{T+2} > \pi_{T+1}$. We want to show that there is another self-enforcing quasi-monotonic contract $\{w'_t\}$ which yields a higher expected profit for firms (a similar argument can be applied for later tenure periods). From the original contract $\{w_t\}$, which satisfies (15) and (ICF), we construct another $\{w'_t\}$ as follows: increase w_t by ε (ε is small) for any $t \geq T+2$, and decrease w_{T+1} by $\Delta = \sum_{t=T+2}^{\infty}(\delta\rho)^{t-(T+1)}\frac{\phi_{T+1}}{\phi_t}\varepsilon$. Note that by construction $\{w'_t\}$ is also quasi-monotonic. Moreover, $V'_{T+1} = V_{T+1}$, $V'_T = V_T$ and $V'_1 = V_1$. Therefore, the no-reneging conditions (ICF) hold under $\{w'_t\}$. Now consider the no-shirking condition (15). The change of the LHS of (15) is

$$\begin{aligned} & (\delta\rho)^{T-1}[(1-\delta\rho)(w'_{T+1} - w_{T+1}) + \delta\rho(w'_{T+2} - w_{T+2})] \\ = \; & (\delta\rho)^{T-1}\varepsilon[\delta\rho - (1-\delta\rho)\sum_{t=T+2}^{\infty}(\delta\rho)^{t-(T+1)}\frac{\phi_{T+1}}{\phi_t}] > (\delta\rho)^{T-1}\varepsilon[\delta\rho - (1-\delta\rho)\frac{\delta\rho}{1-\delta\rho}] = 0. \end{aligned}$$

Therefore, under $\{w'_t\}$ (15) is satisfied and not binding. By part (i) of Theorem 3, both $\{w'_t\}$ and the original contract $\{w_t\}$ cannot be optimal. Therefore, we must have $\pi_{T+2} = \pi_{T+1}$ in optimal contracts.

Part (ii). Suppose $V_{T+1} > V_1$ in the original quasi-monotonic contract $\{w_t\}$, which satisfies (15) and (ICF). We construct another contract $\{w'_t\}$ as follows: increase w_t by ε for all $t \geq T+1$ and reduce w_T by $\Delta = \sum_{t=T+1}^{\infty}(\delta\rho)^{t-T}\frac{\phi_1}{\phi_t}\varepsilon$. The new contract $\{w'_t\}$ is still quasi-monotonic. $V'_{T+1} < V_{T+1}$ and $V'_1 = V_1$. But for ε small enough $V'_{T+1} \geq V'_1$ still holds, since $V_{T+1} > V_1$. Therefore, the no-reneging conditions (ICF) are satisfied under $\{w'_t\}$. As in the proof of part (i), it can be verified that (15) is satisfied and not binding under $\{w'_t\}$. But a nonbinding (15) implies that both $\{w'_t\}$ and the original contract $\{w_t\}$ are not optimal. Therefore, we must have $V_{T+1} = V_1$ in optimal contracts.

In optimal contracts, given that π_t is constant after tenure period $T+1$, and $V_{T+1} = V_1$, we must have $\pi_{T+1} = (1-\delta)V_1$. Moreover, V_t is constant after tenure period $T+1$ as well. Writing V_1 recursively and using $V_{T+1} = V_1$, we have

$$(1-\delta)V_1 \sum_{t=1}^{T} (\delta\rho)^{t-1} \frac{\phi_1}{\phi_t} = \sum_{t=1}^{T} (\delta\rho)^{t-1} \frac{\phi_1}{\phi_t} \pi_t,$$

which gives rise to (16). ■

Lemma 4 results from firms' incentive to backload wages. To minimize informational rents to low type workers, it is always better for firms to minimize wages in earlier tenure periods and maximize wage increases in later tenure periods to provide incentives. This is because low type workers are more likely to be in earlier tenure periods. Subject to the constraint that stage-profits are nondecreasing after tenure period $T+1$, the firm's stage profits are constant in all later tenure periods.

Notice that π_t is increasing from tenure period 1 to $T-1$. By Lemma 4, π_T is greater than the weighted average of π_t $(1 \le t \le T-1)$, and π_t $(t \ge T+1)$ is equal to the weighted average of the stage profits from tenure period 1 to T. Therefore, $\pi_t \ge \pi_1$ for any t. As a result, $V_1 \ge \pi_1/(1-\delta) = (\phi_1 + p(1-\phi_1))/(1-\delta) > p(1-\delta)$. This implies that, if an optimal pooling contract exists, firms have no incentive to deviate to the zero-wage contract (we can ignore requirement (iv) for the programming problem [PP]).

By Lemma 4, optimal pooling contracts are characterized by (T, w_T), $T \ge 2$. Define the LHS of (15) with (T, w_T) as

$$\begin{aligned} G(T, w_T) &= \sum_{t=T+1}^{\infty} (\delta\rho)^{t-1} \left(\frac{\phi_{t+1}}{\phi_{t+2}} - \frac{\phi_t}{\phi_{t+1}}\right) + (\delta\rho)^{T-2} w_T + (\delta\rho)^{T-1}(w_{T+1} - w_T), \qquad (17) \\ \text{where } w_{T+1} &= \frac{\phi_{T+1}}{\phi_{T+2}} - \frac{\sum_{t=1}^{T-1} (\delta\rho)^{t-1} \frac{\phi_1}{\phi_{t+1}} + (\delta\rho)^{T-1} \frac{\phi_1}{\phi_T} \left(\frac{\phi_T}{\phi_{T+1}} - w_T\right)}{\sum_{t=1}^{T} (\delta\rho)^{t-1} \frac{\phi_1}{\phi_t}}. \end{aligned}$$

The expression for w_{T+1} follows (16).

Lemma 5 *Fixing T, $G(T, w_T)$ is increasing in W_T. Define $g(T) \equiv \max_{w_T} G(T, w_T)$ subject to $w_T \le w_{T+1}$. $g(T)$ is decreasing in T.*

Proof. Inspecting (17), we see that w_{T+1} is increasing in w_T. Since $G(T, w_T)$ is increasing in both w_T and w_{T+1}, $G(T, w_T)$ is increasing in w_T. However, the restriction of $\pi_T \ge (1-\delta)V_1$ places an upper bound on W_T. Substituting in this upper bound, we have

$$g(T) = \sum_{t=T}^{\infty} (\delta\rho)^{t-1} \left(\frac{\phi_{t+1}}{\phi_{t+2}} - \frac{\phi_t}{\phi_{t+1}}\right) + (\delta\rho)^{T-2} \left[\frac{\phi_T}{\phi_{T+1}} - \frac{\sum_{t=1}^{T-1} (\delta\rho)^{t-1} \frac{\phi_1}{\phi_{t+1}}}{\sum_{t=1}^{T-1} (\delta\rho)^{t-1} \frac{\phi_1}{\phi_t}}\right]. \qquad (18)$$

Inspecting (18), $g(T)$ is decreasing in T. ■

By Lemma 4 and 5, the programming problem [PP] has a solution if and only if $g(2) \geq \widehat{c}$. Actually, the contract that corresponds to $g(2)$ is a constant stage profit contract (π_t is constant in t), in which the wage increases exactly match the learning effect. More explicitly, the condition $g(2) \geq \widehat{c}$ can be written as

$$f(\phi_1) \equiv \sum_{j=1}^{\infty} (\delta\rho)^{j-1}(\phi_{j+1} - \phi_j) \geq \frac{\widehat{c}}{(1-p)}. \tag{19}$$

Note that ϕ_t (for all t) is a function of ϕ_1. Therefore, the left hand side of (19) is a function of ϕ_1, which we define as $f(\phi_1)$.

Lemma 6 *$f(0) = f(1) = 0$; $f(\phi_1)$ is concave; $\frac{df}{d\phi_1}(0) > 0$ and $\frac{df}{d\phi_1}(1) < 0$.*

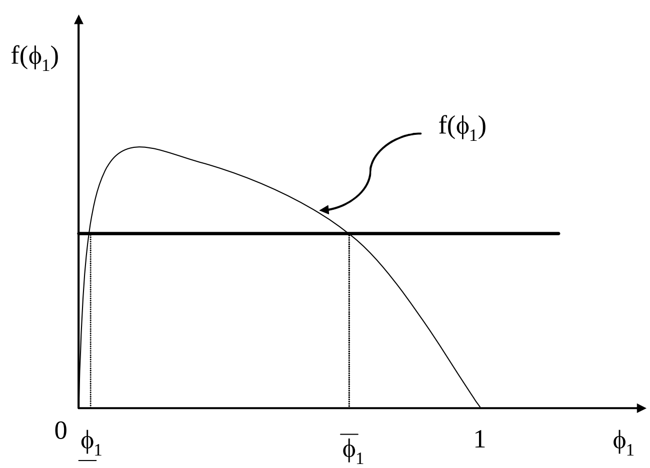

Figure 2: The Shape of $f(\phi_1)$

Figure 2 plots a typical $f(\phi_1)$. Following Lemma 6, $f(\phi_1)$ is concave. Let $\phi_1^* \in (0,1)$ be the maximizer of $f(\phi_1)$, where ϕ_1^* is defined by $\frac{df}{d\phi_1}(\phi_1^*) = 0$. Thus a necessary condition for condition (19) to hold is

$$f(\phi_1^*) \geq \frac{\widehat{c}}{(1-p)} = \frac{c}{\delta\rho(1-p)^2}. \tag{20}$$

Note that $f(\phi_1)$ is independent of c. Thus (20) is satisfied if c is small enough. If (20) is satisfied, then there is an interval $[\underline{\phi_1}, \overline{\phi_1}]$ such that the programming problem [PP] has a solution for any $\phi_1 \in [\underline{\phi_1}, \overline{\phi_1}]$ (see figure 2), where $\underline{\phi_1}$ and $\overline{\phi_1}$ are the two roots of the equation $f(\phi_1) = \frac{\widehat{c}}{(1-p)}$. Also recall that in the stationary state, ϕ_1 is a function of β with the following properties: $\phi_1(0) = 1$, $\phi_1(1) = 0$ and $\frac{d\phi_1}{d\beta} < 0$. It follows that the programming problem [PP] has a solution if and only if $\beta \in [\underline{\beta}, \overline{\beta}]$, where $\overline{\beta}$ is given by $\underline{\phi_1} = \phi_1(\overline{\beta})$ and $\underline{\beta}$ is given by $\overline{\phi_1} = \phi_1(\underline{\beta})$. The following proposition summarizes the previous analysis.

Proposition 2 *A quasi high-effort equilibrium with pooling contracts exists if only if (20) holds and* $\beta \in [\underline{\beta}, \overline{\beta}]$*, with both* $\underline{\beta}$ *and* $\overline{\beta}$ *interior to* $[0, 1]$*.*

Proposition 2 implies that adverse selection helps alleviate moral hazard when firms are not able to commit to long-term contracts, and all the agents are able to change their partners freely in a market. The result is driven by the tension between incentive provision and contractual enforcement. Incentive provision requires that the discounted sum of wage increases be big enough. However, contractual enforcement requires that the wage increases cannot exceed the increases of workers' expected output (the speed of learning), since otherwise longer-tenured workers are less profitable than new workers. If there are no low type workers, then the learning effect is absent, and contractual enforcement requires wage be constant. Thus no incentive can be provided and high-effort equilibria cannot be sustained. On the other hand, the learning effect created by the presence of lower type workers can alleviate the tension between incentive provision and contractual enforcement. Since with the learning effect, workers' expected output is increasing with tenure, so an increasing wage contract can still satisfy the contractual enforcement constraint as long as wage increases more slowly than expected output does. Moreover, how fast wages can increase without violating the contractual enforcement constraint depends on the belief updating process. If the proportion of low type workers is too small, then the magnitude of belief updating is too small, thus not enough incentives can be provided. On the other end of spectrum, if the proportion of low type workers is too big, belief updates slowly in earlier tenure periods. Because of discounting, not enough incentives can be provided either.

The assumption that both workers' past performance (in previous firms) and previous employment history are not observable (no record-keeping) is essential in sustaining high-effort equilibria. If a worker's past track-record were observable, then high-effort equilibria cannot be sustained, similar to a result in Mailath and Samuelson (2001). Intuitively, perfect record-keeping would destroy the learning effect in individual matches, which leads to a constant wage by the contract enforcement constraint and therefore no incentive can be provided. Thus in some sense, the absense of information flows among matches is beneficial in overcoming moral hazard.

Proposition 2 is derived under the assumption that there is no matching friction in markets. Matching friction, as shown by MacLeod and Malcomson (1998) and Yang (2008), also can alleviate moral hazard, since it generates a positive surplus in current employment relationships.[15] Broadly speaking, Proposition 2 can be understood as follows: to overcome moral hazard in markets there must be some friction, and the friction can come from either matching friction or adverse selection.[16]

[15] In MacLeod and Malcomson (1998), the matching friction comes from an unequal number of jobs and workers, while in Yang (2008), it comes from some exogenous turnover costs.

[16] In a repeated matching market with no contracts, Dutta (1993) and Ghosh and Ray (1996) show that high effort can still be *partially* sustained in equilibrium even if there is no adverse selection or matching friction. One may wonder whether a similar result holds in our setting. Specifically, in the first N tenure periods, wages are zero, workers exert zero effort, and no endogenous separation occurs. After tenure period N, wages start to rise, high type

However, if there is only matching friction, then stationary bonus contracts are optimal since the surplus of current employment relationships is independent of tenure. Given the fact that wage increases with tenure (thus not stationary) in reality, we believe that adverse selection plays a role in determining wage dynamics.

Now we characterize the optimal pooling contracts, assuming the programming problem [PP] has a solution, $g(2) \geq \widehat{c}$.

Proposition 3 *Suppose $\beta \in [\underline{\beta}, \overline{\beta}]$. There is a unique optimal pooling contract, which is quasi-monotonic and characterized by T^* and $w^*_{T^*}$. The unique T^* satisfies $g(T^*) \geq \widehat{c}$ and $g(T^*+1) < \widehat{c}$, and $w^*_{T^*}$ satisfies $G(T^*, w^*_{T^*}) = \widehat{c}$. In the optimal pooling contract, $w_t = 0$ if $t < T^*$, $w^*_{T^*+1}$ is determined by (17), and $w_t = w^*_{T^*+1} + (\frac{\phi_t}{\phi_{t+1}} - \frac{\phi_{T^*+1}}{\phi_{T^*+2}})$ if $t > T^* + 1$.*

Proof. By Proposition 1 and Lemma 4, optimal pooling contracts must be quasi-monotonic and are characterized by T and w_T. Given T and w_T, w_{T+1} can be computed according to (17). For all $t < T$, $w_t = 0$, and for $t > T+1$, $w_t = w_{T+1} + (\frac{\phi_t}{\phi_{t+1}} - \frac{\phi_{T+1}}{\phi_{T+2}})$ by the constant-stage-profit requirement. The binding (15) pins down optimal contracts: $G(T, w_T) = \widehat{c}$. We first determine T^*. By Lemma 5, $g(T)$ is decreasing in T. Moreover, $\lim_{T\to\infty} g(T) = 0$ and $g(2) \geq \widehat{c}$. Therefore, there is a unique T^* such that $g(T^*) \geq \widehat{c}$ and $g(T^*+1) < \widehat{c}$. Given T^*, there is a unique $w^*_{T^*} \in (0, \frac{\phi_T}{\phi_{T+1}} - \frac{\sum_{t=1}^{T-1} (\delta\rho)^{t-1} \frac{\phi_1}{\phi_{t+1}}}{\sum_{t=1}^{T-1} (\delta\rho)^{t-1} \frac{\phi_1}{\phi_t}})$ such that $G(T^*, w^*_{T^*}) = \widehat{c}$. Therefore, the optimal contract is unique. ■

The optimal pooling contract is determined by three forces. First, to provide incentives to high type workers, the discounted sum of wage increases must be equal to a given level. Second, to reduce informational rent to low type workers, firms try to backload wages as much as possible since low type workers are more likely to be in earlier tenure periods. Third, firms' ability to backload wages is limited by firms' no-reneging conditions: senior workers have to be more profitable than new workers. The last two forces lead to constant stage profits in later tenure periods and constant (zero) wage in early tenure periods.

By Proposition 3, the wage dynamics in the optimal pooling contract exhibits two salient features. First, wage is low and remains constant in earlier tenure periods. These earlier tenure periods can be interpreted as probation periods, which are very common in labor markets. Second, when wage starts to increase in later tenure periods, it is intimately related to the learning effect: the wage increase in each tenure period (after tenure period T^*) is exactly equal to the increase in a worker's expected output. This is because the stage profit is constant after wage starts to increase.

The wage dynamics exhibited by the optimal pooling contract in our model is different from that in Jovanovic (1979) in several aspects. First, in his model there are no probation periods. Second,

workers exert high effort, and endogenous separation occurs. In the above strategy, the inefficiency endogenously created in early tenure periods might help to prevent parties from deviating. However, the above strategy cannot be supported as an equilibrium in our setting. This is because it is the inefficiency created by other relationships that prevent parties in the current relationship from deviating. An individual firm always has an incentive to deviate to a contract that induces high effort in every period.

in his model wage in each tenure period is exactly equal to a worker's expected productivity, while in our model only the *wage increase* after the probation periods is equal to the increase in a worker's expected productivity. Finally, in his model the fact that wage always matches a worker's expected productivity is due to market competition. On the other hand, in our model workers' performance is not observed by the market, yet in the optimal pooling contract the wage increases are intrinsically tied to workers' performance. This implies that, even without market competition, wages being tied to workers' expected productivities can be generated by *internal wage dynamics*. In our model, it is the interaction between incentive provision and contractual enforcement that leads to the internal wage dynamics being tied to the learning effect. Note that both incentive provision and contractual enforcement are indispensable. Suppose there is no moral hazard, then wage needs not increase to provide incentives. On the other hand, suppose contractual enforcement is not an issue (say firms are able to commit to long-term contracts), then the wage increases can be arbitrary, as firms can backload wages as much as possible.

Our model generates several empirical implications about the wage dynamics. The first implication is that wage increases with tenure, which is a well-established empirical fact (Mincer,1974; Becker; 1975, Mincer and Jovanovic, 1981). In particular, in our model wage increases with seniority because workers' *expected* productivity increases with seniority, but all the remaining workers in a firm have the same *realized* productivity. This is consistent with Medoff and Abraham (1980), who found that though a worker's wage increases with seniority, his performance rating relative to others does not increase with seniority. Our model also implies that after controlling for the learning effect (unobserved heterogeneity in worker quality), tenure should have little effect on wages. This implication is consistent with a number of empirical studies (Abraham and Farber, 1987; Altonji and Shakotko, 1987; Brown, 1989). The second implication is that the variance of workers' productivity (quality) among the workers with the same tenure should decrease with tenure, since low type workers are gradually screened out. This is consistent with the empirical findings in Baker, Gibbs and Holmstrom (1994).

4.3 Discussion

Here we discuss how the various frictions in our model contribute to the features of the optimal pooling contract. The presence of low type workers (adverse selection) creates the incentive for firms to backload wages, as they try to reduce the informational rent to low type workers. On the other hand, a lack of intertemporal commitment limits firms' ability to backload wages. The output being not verifiable is a key assumption that leads to a lack of intertemporal commitment. Typically the length of an employment relationship (tenure period t) is verifiable. If output were verifiable, then the following contract can be enforced by the court: the firm will continue the relationship if and only if the output is 1, and if the relationship continues, the wage follows $\{w_t\}$. Now firms cannot renege and they can backload wages as much as they want. With output not being verifiable, and firing being necessary for workers producing zero output, firms always have flexibility in terminating a relationship. This means that firms cannot commit to long-term contracts even if tenure period t is verifiable. The assumption of limited liability pins down the wages in early tenure periods (the low bound, zero). Without limited liability the wages in early tenure periods will be negative. But since a lack of intertemporal commitment limits firms' ability to backload wages, even without limited liability wages in early tenure periods cannot be too negative, since otherwise new workers are more profitable than old workers and firms would have incentives to renege.

5 Separating Contracts

Without loss of generality, w_L should be set such that (12) is binding (L type workers are indifferent between choosing the L contract w_L and H contract $\{w_t^s\}$): $w_L = (1-\delta\rho p)\sum_{t=1}^{\infty}(\delta\rho p)^{t-1}w_t^s$. That is, w_L equals to the average per-period payoff if an L type chooses the H contract $\{w_t^s\}$. Now it can be verified that (11) is redundant given that (13) holds. Intuitively, if a H type worker chooses contract w_L, he gets the same payoff when he chooses contract $\{w_t^s\}$ and shirks in every period. Therefore, no-shirking conditions imply that H type workers have no incentive to choose the L contract. As in the case of pooling contracts, for separating contracts (without loss of generality) we can focus on nondecreasing contracts, that is, w_t^s is nondecreasing in t. Let T_s be the first tenure period that w_t^s is strictly positive. For a nondecreasing contract $\{w_t^s\}$, U_t^s is nondecreasing in t, hence the no-shirking conditions of (13) hold if and only if the no-shirking condition holds for $t=2$. Now conditions (11)-(14) boil down to the following two conditions:

$$U_2^s - U_1^s \geq \widehat{c} \Leftrightarrow \sum_{t=1}^{\infty}(\delta\rho)^{t-1}(w_{t+1}^s - w_t^s) \geq \widehat{c}, \tag{21}$$

$$\sum_{t=j}^{\infty}(\delta\rho)^{t-j}(1-w_t^s) - \frac{1}{1-\delta\rho(1-\lambda)}\{\lambda\sum_{t=1}^{\infty}(\delta\rho)^{t-1}(1-w_t^s) \tag{22}$$

$$+(1-\lambda)[p-(1-\delta\rho p)\sum_{t=1}^{\infty}(\delta\rho p)^{t-1}w_t^s]\} \geq 0, \text{ for } j > T_s.$$

The self-selection condition for L types deserves more comments. In typical repeated adverse selection models (e.g., Laffont and Tirole, 1990), inducing separation in the first period is very costly, as the discounted sum of informational rents in all future periods has to be paid in the first period. Translating into our setting, w_L would have been $\sum_{t=1}^{\infty}(\delta\rho p)^{t-1}w_t^s$, the discounted informational rents in a relationship. However, in our model w_L is less than the discounted informational rents in a relationship. The difference is that in repeated adverse selection models, there is only a single relationship and thus an agent gets zero rent after revealing his type. In contrast, in our model this is not the case: after leaving the current relationship, next period a L type worker can match with another firm and get informational rents as well. Therefore, there is an opportunity cost for a L type worker to mimic a H type in the current relationship. As a result, to induce a L type worker to reveal his type, a firm does not need to pay the discounted sum of informational rents in the current relationship. In other words, the cost of separating is relatively low. Define the cost of separating as w_L minus the per-period wage that firms pay on average to a L type worker who always chooses the H contract. Note that the latter equals to $(1-\delta\rho p)\sum_{t=1}^{\infty}(\delta\rho p)^{t-1}w_t$. Therefore, in the current setup of the model, the cost of separating is zero.

The following lemma shows that in searching for optimal separating contracts, we can focus on a class of contracts with a particular form.

Lemma 7 *If a separating contract $(w_L, \{w_t^s\})$ is self-enforcing (satisfies (21) and (22)), then there is another self-enforcing separating contract of the following form: w_t^s is constant after tenure period T_s+1 and $1-w_{T_s+1}^s=(1-\delta)V_N$. Moreover, optimal separating contracts must have the above form.*

Proof. First note that (21) must be binding in optimal contracts. Now we show that in optimal separating contracts w_t^s must be constant after tenure period T_s+1. Specifically, we show that $w_{T_s+2}^s$ must equal to $w_{T_s+1}^s$(a similar argument can show that in optimal contracts wage must be constant in later tenure periods). Now suppose there is a self-enforcing and nondecresing contract such that $w_{T_s+2}^s > w_{T_s+1}^s$. Then design another contract $\{w_t^{s\prime}\}$ as follows: $w_{T_s+1}^{s\prime}=w_{T_s+1}^s+\varepsilon$, and $w_{T_s}^{s\prime}=w_{T_s+1}^s-\Delta$, where $\Delta=\frac{\lambda+(1-\lambda)(1-\delta\rho p)p^{T_s}}{\lambda+(1-\lambda)(1-\delta\rho p)p^{T_s-1}}\delta\rho\varepsilon$. By construction, $w_t^{s\prime}$ is nondecreasing (by $w_{T_s+2}^s > w_{T_s+1}^s$) and $V_N'=V_N$. Therefore, (22) still holds. Now consider the change of the LHS of (21):

$$(\delta\rho)^{T_s-2}[-\Delta+\delta\rho(\varepsilon+\Delta)-(\delta\rho)^2\varepsilon]\sim\varepsilon(1-\frac{\lambda+(1-\lambda)(1-\delta\rho p)p^{T_s}}{\lambda+(1-\lambda)(1-\delta\rho p)p^{T_s-1}})>0.$$

Thus under $\{w_t^{s\prime}\}$, (21) holds with strict inequality. Therefore, both $\{w_t^{s\prime}\}$ and $\{w_t^s\}$ are not optimal.

Next we show that $1-w_{T_s+1}^s=(1-\delta)V_N$. Inequality (22) and w_t^s being constant after tenure period T_s+1 imply that $1-w_{T_s+1}^s\geq(1-\delta)V_N$. So we only need to rule out $1-w_{T_s+1}^s>(1-\delta)V_N$. Suppose there is a self-enforcing and nondecresing contract such that w_t^s is constant after tenure period T_s+1 and $1-w_{T_s+1}^s>(1-\delta)V_N$. We design another contract $\{w_t^{s\prime}\}$ as follows: $w_t^{s\prime}=w_t^s+\varepsilon$ for any $t\geq T_s+1$ and $w_{T_s}'=w_{T_s}-\Delta$, where $\Delta=\frac{\lambda/(1-\delta\rho)+(1-\lambda)p^{T_s}}{\lambda+(1-\lambda)(1-\delta\rho p)p^{T_s-1}}\delta\rho\varepsilon$. By construction, $w_t^{s\prime}$ is nondecreasing and $V_N'=V_N$. By the fact that $1-w_{T_s+1}^s>(1-\delta)V_N$, for ε small enough (22) still holds under $\{w_t^{s\prime}\}$. Now consider the change of the LHS of (21):

$$(\delta\rho)^{T_s-2}[-\Delta+\delta\rho(\varepsilon+\Delta)]\sim\varepsilon(1-\frac{\lambda+(1-\lambda)(1-\delta\rho)p^{T_s}}{\lambda+(1-\lambda)(1-\delta\rho p)p^{T_s-1}})>0.$$

Thus under $\{w_t^{s\prime}\}$, (21) holds with strict inequality. Therefore, both $\{w_t^{s\prime}\}$ and $\{w_t^s\}$ are not optimal.

■

The class of contracts described in Lemma 7 is characterized by T_s and w_{T_s}. Given T_s and w_{T_s}, $w_{T_{s+1}}$ is determined by (subject to $w_{T_s} \le w_{T_s+1}$):

$$0 = \frac{1 - w_{T_{s+1}}}{1-\delta\rho} - \frac{1}{1-\delta\rho(1-\lambda)}\{\lambda[\frac{1-(\delta\rho)^{T_s-1}}{1-\delta\rho} + (\delta\rho)^{T_s-1}(1-w_{T_s}) + (\delta\rho)^{T_s}\frac{1-w_{T_{s+1}}}{1-\delta\rho}]$$
$$+(1-\lambda)[p-(1-\delta\rho p)[(\delta\rho p)^{T_s-1}w_{T_s} + \frac{(\delta\rho p)^{T_s}w_{T_{s+1}}}{1-\delta\rho p}]\} \tag{23}$$

From (23), we can see that $w_{T_{s+1}}$ is increasing in w_{T_s} and decreasing in T_s. Define the LHS of (21) as

$$G_s(T_s, w_{T_s}) = (\delta\rho)^{T_s-2}w_{T_s} + (\delta\rho)^{T_s-1}(w_{T_{s+1}} - w_{T_s}). \tag{24}$$

It can be verified that $G_s(T_s, w_{T_s})$ is increasing in w_{T_s}. Subject to $w_{T_s} \le w_{T_s+1}$, $G_s(T_s, w_{T_s})$ is maximized when $w_{T_s} = w_{T_s+1}$. Now define $g_s(T_s) \equiv \max_{w_{T_s}} G_s(T_s, w_{T_s})$. More specifically,

$$g_s(T_s) = (\delta\rho)^{T_s-2}w_{T_s};\ \text{where } w_{T_s} \text{ satisfies } 0 = \frac{1-w_{T_s}}{1-\delta\rho} - \frac{1}{1-\delta\rho(1-\lambda)}$$
$$\times\{\lambda[\frac{1-(\delta\rho)^{T_s-1}}{1-\delta\rho} + (\delta\rho)^{T_s-1}\frac{1-w_{T_s}}{1-\delta\rho}] + (1-\lambda)[p-(\delta\rho p)^{T_s-1}w_{T_s}]\}. \tag{25}$$

From (25), we see that w_{T_s} is decreasing in T_s. Therefore, $g_s(T_s)$ is decreasing in T_s. Therefore, self-enforcing separating contracts exist if and only if $g_s(2) \ge \widehat{c}$. This condition can be written more explicitly as

$$\frac{c}{\delta\rho(1-p)^2} \le \frac{(1-\lambda)}{1-(1-\lambda)\delta\rho p}. \tag{26}$$

Note that the RHS of (26) is decreasing in λ. Therefore, a necessary condition for (26) to be satisfied is that it is satisfied for $\lambda = 0$, or equivalently

$$\frac{c}{\delta\rho(1-p)^2} \le \frac{1}{1-\delta\rho p} \tag{27}$$

Note that the RHS of (26) is 0 when $\lambda = 1$. Thus (26) cannot be satisfied if $\lambda = 1$. If condition (27) is satisfied, then there is a $\widehat{\lambda} \in [0, 1)$ such that (26) is satisfied if and only if $\lambda \in [0, \widehat{\lambda}]$. Equation (10) defines λ as a function of β, and $\lambda(\beta)$ is decreasing in β. Therefore, (26) is satisfied if and only if $\beta \in [\widehat{\beta}, 1]$, where $\widehat{\beta}$ is defined as $\lambda(\widehat{\beta}) = \widehat{\lambda}$.

Now suppose (26) is satisfied. To search for optimal separating contracts, which is characterized by $(T_s^*, w_{T_s}^*)$, we first identify T_s^*. Specifically, T_s^* is determined by $g_s(T_s^*) \ge \widehat{c}$ and $g_s(T_s^* + 1) < \widehat{c}$. Note that such a T_s^* is unique since $g_s(T_s)$ is decreasing in T_s. After T_s^* is determined, then $w_{T_s}^*$ is determined by $G_s(T_s^*, w_{T_s}^*) = \widehat{c}$. Thus, we have the following proposition.

Proposition 4 *A quasi high-effort equilibrium with separating contracts exist if and only if (27) holds, $\beta \in [\widehat{\beta}, 1]$ with $\widehat{\beta} \in (0, 1)$, and $V_N \ge p/(1-\delta)$. If the equilibrium exists, the optimal separating contract is unique and has the following form: $w_t = 0$ for $t < T_s^*$, $G_s(T_s^*, w_{T_s}^*) = \widehat{c}$, w_t is constant for $t \ge T_s^* + 1$, $w_{T_s}^*$ and $w_{T_s+1}^*$ satisfy (23), and T_s^* is determined by $g_s(T_s^*) \ge \widehat{c}$ and $g_s(T_s^* + 1) < \widehat{c}$.*

Proposition 4 indicates that self-enforcing separating contracts exist if and only if there are enough L type workers. The wage increases of $\{w_t^s\}$ has to be big enough to motivate H type workers. To prevent firms from reneging, there must be enough punishment for reneging. This punishment comes from the scarcity of H type workers who generate higher profits for firms: after reneging, firms must match with new workers who might be L type workers. The more L type workers, the lower the probability to match with a H type worker in the unmatched pool, hence the bigger the punishment for reneging. Recall Proposition 2. Self-enforcing pooling contracts exist if and only if the proportion of L type workers is not too low or too high. The difference comes from the fact that with pooling contracts, the wage increases cannot exceed the speed of learning. When the proportion of L type workers is too high, the belief updating will be very slow initially, and due to discounting, not enough incentives can be provided to H type workers.[17]

The forces that determine the optimal separating contract are similar to those that govern the optimal pooling contract. To provide incentives to high type workers, the discounted sum of wage increases must be big enough. To reduce informational rent to low type workers, firms try to backload wages as much as possible. However, firms' ability to backload wages is limited by firms' no-reneging conditions. The last two forces lead to constant stage profits in later tenure periods and constant (zero) wage in early tenure periods.

The wage dynamics in the optimal separating contract and the optimal pooling contract share a similar feature: wage is low and remains constant in earlier tenure periods. The difference is that in the optimal separating contract, wage increases at most in two tenure periods, and then wage remains constant afterwards. This difference comes from the fact that learning is completed in the first tenure period under separating contracts. Thus constant stage profits in later tenure periods implies constant wage.

6 Comparison

Now we compare pooling contracts and separating contracts.

Lemma 8 *If a self-enforcing pooling contract exists, then a self-enforcing separating contract exists. On the other hand, for some parameter values, only self-enforcing separating contracts exist.*

Proof. First, we show that the necessary condition for a self-enforcing pooling contract to exist, (20), is more stringent than the necessary condition for a self-enforcing separating contract to exist, (27). Specifically, if (20) holds, then $1 > \frac{\widehat{c}}{(1-p)}$, since $f(\phi_1^*) < 1$. This implies that $\frac{1}{1-\delta\rho p} > \frac{\widehat{c}}{(1-p)}$, hence (27) holds as well.

[17] When β is close to 1, we do not expect a quasi high-effort equilibrium with separating contracts to exist. This is because if there are too many low type workers, it is not worthwhile for firms to motivate high type workers and give informational rents to low type workers. Instead, all firms will offer the zero-wage contract.

Now suppose that (20) holds. Recall that self-enforcing pooling contracts exist when $\beta \in [\underline{\beta}, \overline{\beta}]$, and self-enforcing separating contracts exist when $\beta \in [\widehat{\beta}, 1]$. Thus it is sufficient to show that $\underline{\beta} > \widehat{\beta}$. By (1) and (10), $\phi_1(\beta) > \lambda(\beta)$, thus it is enough to show that $\underline{\phi_1} \geq \widehat{\lambda}$. Recall that $\underline{\phi_1}$ is the smaller root of the equation $f(\phi_1) = \frac{\widehat{c}}{1-p}$, and $\widehat{\lambda}$ is the solution to equation $f_s(\lambda) \equiv \frac{(1-\lambda)}{1-(1-\lambda)\delta\rho p} = \frac{\widehat{c}}{1-p}$ and $f_s(\lambda)$ is decreasing in λ. Therefore, to show $\underline{\phi_1} \geq \widehat{\lambda}$, it is sufficient to show that $f(\phi_1) \leq f_s(\phi_1)$. More explicitly,

$$\begin{aligned} f(\phi_1) &= \sum_{t=1}^{\infty}(\delta\rho)^{t-1}[\frac{\phi_1}{\phi_1 + p^t(1-\phi_1)} - \frac{\phi_1}{\phi_1 + p^{t-1}(1-\phi_1)}] \\ &\leq \sum_{t=1}^{\infty}[\frac{\phi_1}{\phi_1 + p^t(1-\phi_1)} - \frac{\phi_1}{\phi_1 + p^{t-1}(1-\phi_1)}] = 1 - \phi_1. \end{aligned}$$

On the other hand, it can be easily seen that $f_s(\phi_1) \geq 1 - \phi_1$. Therefore, $f(\phi_1) \leq f_s(\phi_1)$ always holds. ■

Lemma 8 states that self-enforcing separating contracts exist under a wider range of parameter values than self-enforcing pooling contracts do. The intuition for this result is as follows. Under separating contracts, since learning is completed in the first tenure period, subject to the no-reneging conditions the maximum amount of wage increase can occur in the second tenure period. On the other hand, since under pooling contracts learning occurs gradually, the same amount of wage increase has to be spread over many tenure periods. Due to discounting, less incentive are provided to H type workers with pooling contracts.

Lemma 9 *Suppose (20) holds and $\beta \in [\underline{\beta}, \overline{\beta}]$, so that both types of self-enforcing contracts exist. Suppose $\phi_1 = \lambda$. Let V_1 under the optimal pooling contract $\{w_t^*\}$ be V_1^*, and V_N under the optimal separating contract $(w_L^*, \{w_t^{s*}\})$ be V_N^*. Then we must have $V_1^* < V_N^*$.*

Proof. Suppose the opposite, $V_1^* \geq V_N^*$. From the ex ante point of view, V_1^* can be written as

$$V_1 = \phi_1[\sum_{t=1}^{\infty}(\delta\rho)^{t-1}(1-w_t^*) + \frac{\delta(1-\rho)}{1-\delta\rho}V_1] + (1-\phi_1)[\sum_{t=1}^{\infty}(\delta\rho p)^{t-1}(p-w_t^*) + \frac{\delta(1-\rho p)}{1-\delta\rho p}V_1]. \quad (28)$$

With $\phi_1 = \lambda$, suppose a firm adopts the following separating contract: $w_t^s = w_t^*$ for all t and $w_L = (1-\delta\rho p)\sum_{t=1}^{\infty}(\delta\rho p)^{t-1}w_t^*$. Note that this separating contract is self-enforcing. To see this, first note that $\{w_t^*\}$ satisfying the no-shirking conditions means that the separating contract also satisfies the no-shirking conditions. By the fact that $\{w_t^*\}$ satisfies the no-reneging conditions, we have $w_\infty^* = (1-\delta)V_1^*$. Given that $V_1^* \geq V_N^*$, we have $w_\infty^* \geq (1-\delta)V_N^* \geq (1-\delta)V_N$. Thus the separating contract satisfies the no-reneging conditions. Under this separating contract, a firm's V_N can be written as

$$V_N = \phi_1[\sum_{t=1}^{\infty}(\delta\rho)^{t-1}(1-w_t^*) + \frac{\delta(1-\rho)}{1-\delta\rho}V_N] + (1-\phi_1)[(1-\delta\rho p)\sum_{t=1}^{\infty}(\delta\rho p)^{t-1}(p-w_t^*) + \delta V_N]. \quad (29)$$

Note that V_N for a positive ϕ_1 is strictly greater than the value if $\phi_1 = 0$, that is, $V_N > \frac{1-\delta\rho p}{1-\delta}\sum_{t=1}^{\infty}(\delta\rho p)^{t-1}(p - w_t^*)$. Now using this inequality, by (29) we have

$$V_N > \phi_1[\sum_{t=1}^{\infty}(\delta\rho)^{t-1}(1 - w_t^*) + \frac{\delta(1-\rho)}{1-\delta\rho}V_N] + (1-\phi_1)[\sum_{t=1}^{\infty}(\delta\rho p)^{t-1}(p - w_t^*) + \frac{\delta(1-\rho p)}{1-\delta\rho p}V_N]. \quad (30)$$

Now compare (28) and (30), we can clearly see that $V_N > V_1^*$. Therefore we have $V_N^* \geq V_N > V_1^*$, a contradiction. ■

Lemma 9 implies that for any initial belief ϕ_1, each individual firm always has incentive to offer a separating contract instead of a pooling contract. In other words, optimal separating contracts dominate optimal pooling contracts.[18] Since by Lemma 8, a self-enforcing separating contract exists whenever a self-enforcing pooling contract exists, so there is no high-effort equilibrium with pooling contracts. This leads to the following proposition.

Proposition 5 *If a high-effort equilibrium exists, the associated equilibrium contract must be the optimal separating contract. Specifically, if condition (27) holds, $\beta \in [\widehat{\beta}, 1]$, and $V_N \geq p/(1-\delta)$, then a high-effort equilibrium exists.*

The intuition for separating contracts to dominate pooling contracts is as follows. Under both types of contracts, firms' ability to backload wages are more or less the same. Under pooling contracts, firms' ability to backload wages is dictated by the gradual increase of the beliefs about workers as tenure period increases. Under separating contracts, though learning is completed in tenure period 1, firms are able to backload wages since in tenure period 1, workers are very likely to be of low type. Comparing separating contracts and pooling contracts, there is an additional effect that favors separating contracts. With separating contracts, a firm is able to learn the type of a new worker immediately. In contrast, with pooling contracts it takes a longer time for a firm to learn a worker's type. Thus, with the same initial beliefs, on average it takes a shorter time for a firm to match with a H type worker with separating contracts. This fast screening effect favors separating contracts.

An implicit reason for separating contracts to dominate pooling contracts is that there is no cost of separating in our setting, as we mentioned earlier. This result crucially depends on the feature that after leaving the current relationship, in the next period a worker can always match with another firm. Actually, this result of zero separating costs no longer holds if there is unemployment in the economy. Suppose the measure of workers is still 1, but the measure of firms is $\alpha \in (0, 1)$. That is, workers are on the long side of the market. Now workers in the unmatched pool may not

[18] Consider a numerical example with $\delta = 0.95$, $\rho = 0.9$, $c = 0.25$, $\beta = 0.354$ and $p = 0.3$. Under pooling contracts, $\phi_1 = 0.2$. The optimal pooling contract is characterized by $T^* = 3$, $w_3^* = 0.3916$, $w_4^* = 0.4513$, and $(1-\delta)V_1^* = 0.4825$. With $\lambda = 0.2$, the optimal separating contract is characterized by $T_s^* = 3$, $w_3^{s*} = 0.455$, $w_4^{s*} = 0.4942$, and $(1-\delta)V_N^* = 0.5058$. Clearly, $V_1^* < V_N^*$.

get a match. Let $\gamma \in (0,1)$ be the probability that a worker in the unmatched pool matches with a firm. Obviously, γ is increasing in α. Suppose firms offer both types a pooling contract $\{w_t\}$. Define U_u^L as a L type's payoff when he is in the unmatched pool, and U_1^L as a L type's payoff when he is just matched with a firm and always chooses the H contract.

$$\begin{aligned} U_u^L &= \gamma U_1^L + (1-\gamma)\delta U_u^L; \\ U_1^L &= \sum_{t=1}^{\infty} (\delta\rho p)^{t-1} w_t + \frac{\delta(1-\rho p)}{1-\delta\rho p} U_u^L. \end{aligned}$$

Now to induce immediate type revelation,

$$w_L = U_1^L - \delta U_u^L = \frac{1-\delta\rho p}{1-(1-\gamma)\delta\rho p} \sum_{t=1}^{\infty} (\delta\rho p)^{t-1} w_t > (1-\delta\rho p) \sum_{t=1}^{\infty} (\delta\rho p)^{t-1} w_t,$$

where $(1-\delta\rho p)\sum_{t=1}^{\infty}(\delta\rho p)^{t-1}w_t$ is the average per-period wage that firms pay to L type workers in the pooling contract. Thus the separating cost now is positive. It is easy to see that as γ (α) decreases, w_L increases and thus the cost of separating increases. Intuitively, when the unemployment rate increases, an L type has a stronger incentive to pool with H types in the current match to avoid separation, since separation now will lead to a lower continuation payoff. As a result, to induce immediate type revelation, the firm has to pay a higher w_L to compensate for separation. Another way to understand this result is that as the unemployment rate increases, the continuation payoff after separation decreases. Thus to induce immediate type revelation, the payment to an L type becomes closer to the discounted sum of informational rents in the current relationship.

This observation suggests that separating contracts will be dominated by pooling contracts if the unemployment rate is high enough, as the cost of separating might outweigh the benefit of separating (the fast screening effect).[19] The existence of unemployment in labor markets is probably the reason that in the real world we see more pooling contracts than separating contracts. Another reason that favors pooling contracts is that selecting an L contract (self-revealing as the low type) might involve some physcological cost, which leads to a positive separating cost.

7 Conclusions

This paper studies nonstationary relational contracts driven by the presence of adverse selection. The internal wage dynamics is pinned down by the tension between incentive provision and contractual enforcement. The paper contributes to the understanding of how contractual enforcement restricts firms' ability in offering long-term contracts in nonstationary environments. Moreover, the

[19] Introducing unemployment would not change the optimal pooling contract qualitatively, though low type workers would get less informational rent.

paper shows that when contractual enforcement is an issue and agents are free to change partners in markets, adverse selection can alleviate moral hazard.

Two features restrict the applications of our model. First, workers' performance should not be publicly observed. This means that this model does not apply to the labor markets for professionals (e.g. lawyers and academia). Second, our model is more relevant for small firms than for large firms, since large firms, which might have external reputation at stake, may be able to commit to long-term contracts to some degree. In reality, small firms and large firms do exhibit different wage-tenure effects: Pearce (1990) found that the wage-tenure profile is steeper for firms with larger average establishment size.

Although our model is framed in a labor market setting, it can be applied to broader settings. In fact, it applies to markets in which both moral hazard and adverse selection exist, and contractual enforcement is an issue. The internal wage dynamics derived in our model can be generalized as *internal contract dynamics*. Two relevant examples are lending markets and buyer-seller relationships. In the context of lending markets,[20] our model implies that as a lending relationship continues, the contractual terms should become more favorable to the borrower, who has the moral hazard problem. This is consistent with the phenomenon of relationship lending: borrowers with longer relationships with a bank pay lower interest rates and are less likely to pledge collateral (Berger and Udell, 1995; Bodenhorn, 2003).

Appendix: Proofs

Proof of Theorem 1.

Proof. The proof is by construction. We show it in several steps.

Step (1) (construction of a new contract). Suppose that a self-enforcing contract $\{w_t\}$ decreases (strictly) from tenure period i to $i+j$, and are nondecreasing elsewhere. (The proof for $\{w_t\}$ decreases strictly more than one place is essentially the same). First, define w_i' as the following:

$$\sum_{t=i}^{i+j} (\delta\rho)^{t-i} \frac{\phi_i}{\phi_t} w_i' = \sum_{t=i}^{i+j} (\delta\rho)^{t-i} \frac{\phi_i}{\phi_t} w_t.$$

Define a new contract $\{w_t'\}$ as $\{w_1, w_2, ...w_{i-1}, w_i', w_i', ...w_i', w_{i+j+1}, ...\}$. That is, $\{w_t'\}$ differs from $\{w_t\}$ only from tenure period i to $i+j$, in which range $\{w_t'\}$ is constant. If $w_{i-1} \leq w_i'$ and $w_{i+j+1} \geq w_i'$, then $\{w_t'\}$ is nondecreasing; and this is the new contract that we are looking for. If either of these two inequalities is not satisfied, we need to redefine contract $\{w_t'\}$.

[20] Specifically, consider a lending market with two types of borrowers (firms): high type and low type. High type firms can choose to implement one project from two available projects: one bad project which is more risky, and one good project with a safe and higher expected return. Low type firms only have access to the bad project.

Case (1): If $w_{i-1} > w'_i$ and $w_{i+j+1} \geq w'_i$, define w'_{i-1} as

$$\sum_{t=i-1}^{i+j} (\delta\rho)^{t-i-1} \frac{\phi_{i-1}}{\phi_t} w'_{i-1} = \sum_{t=i-1}^{i+j} (\delta\rho)^{t-i-1} \frac{\phi_{i-1}}{\phi_t} w_t.$$

And redefine $\{w'_t\}$ as $\{w_1, w_2, ...w_{i-2}, w'_{i-1}, w'_{i-1}, ...w'_{i-1}, w_{i+j+1}, ...\}$.

Case (2): If $w_{i-1} \leq w'_i$ and $w_{i+j+1} < w'_i$, redefine w'_i as

$$\sum_{t=i}^{i+j+1} (\delta\rho)^{t-i} \frac{\phi_i}{\phi_t} w'_i = \sum_{t=i}^{i+j+1} (\delta\rho)^{t-i} \frac{\phi_i}{\phi_t} w_t.$$

And redefine $\{w'_t\}$ as $\{w_1, w_2, ...w_{i-1}, w'_i, w'_i, ...w'_i, w_{i+j+2}, ...\}$.

Case (3): If $w_{i-1} > w'_i$ and $w_{i+j+1} < w'_i$, define w'_{i-1} as

$$\sum_{t=i-1}^{i+j+1} (\delta\rho)^{t-i-1} \frac{\phi_{i-1}}{\phi_t} w'_{i-1} = \sum_{t=i-1}^{i+j+1} (\delta\rho)^{t-i-1} \frac{\phi_{i-1}}{\phi_t} w_t.$$

And redefine$\{w'_t\}$ as $\{w_1, w_2, ...w_{i-2}, w'_{i-1}, w'_{i-1}, ...w'_{i-1}, w_{i+j+2}, ...\}$.

Repeat this procedure until $\{w'_t\}$ is nondecreasing. This is always feasible, because a constant wage contract cannot be self-enforcing. Suppose that the eventual $\{w'_t\}$ differs from $\{w_t\}$ from tenure period k to $k+n$, with $k \leq i$ and $k+n \geq i+j$. Specifically, $\{w'_t\} = \{w_1, ..., w_{k-1}, w'_k, ..., w'_k, w_{k+n+1}, ...\}$. According to the construction procedure, w'_k is defined as

$$\sum_{t=k}^{k+n} (\delta\rho)^{t-k} \frac{\phi_k}{\phi_t} w'_k = \sum_{t=k}^{k+n} (\delta\rho)^{t-k} \frac{\phi_k}{\phi_t} w_t. \tag{31}$$

Step (2) (An important property). From the construction of $\{w'_t\}$, we must have $w_k > w'_k$, otherwise w_k needs not to be redefined in $\{w'_t\}$. Similarly, we must have $w'_k > w_{k+n}$. Moreover, there is an integer z ($i < z < i+j$) such that $w_t \geq w'_k$ for $k \leq t \leq z$ and $w_t \leq w'_k$ for $z \leq t \leq k+n$. To see this, since w_t is monotonically decreasing from i to $i+j$, there is a z such that $w_t \geq w'_k$ for $i \leq t \leq z$ and $w_t \leq w'_k$ for $z \leq t \leq i+j$. For $k \leq t < i$, $w_t \geq w'_k$, otherwise w_t needs not be redefined in $\{w'_t\}$. Similar argument shows that $w_t \leq w'_k$ for $i+j < t < k+n$. Following this property, it can be readily shown that the following inequality holds for any $1 \leq m \leq n$,

$$\sum_{t=k+m}^{k+n} (\delta\rho)^{t-k-m} \frac{1}{\phi_t} (w'_k - w_t) > 0. \tag{32}$$

By the fact that $w_k > w'_k$ and (31), (32) holds for $m = 1$. It follows that (32) holds for $k+m \leq z$, since by removing one negative term the inequality should also hold. For $k+m > z$ (32) obviously holds since all the terms are positive and the last term is strictly positive. Now we are ready to derive an important property. By (31) and (32),

$$\begin{aligned} w_k - w'_k &= \sum_{t=k+1}^{k+n} (\delta\rho)^{t-k} \frac{\phi_k}{\phi_t}(w'_k - w_t) < \sum_{t=k+1}^{k+n} (\delta\rho)^{t-k} \frac{\phi_{k+1}}{\phi_t}(w'_k - w_t) \\ &< \delta\rho(w'_k - w_{k+1}) + \sum_{t=k+2}^{k+n} (\delta\rho)^{t-k} \frac{\phi_{k+2}}{\phi_t}(w'_k - w_t) < ... < \sum_{t=k+1}^{k+n} (\delta\rho)^{t-k}(w'_k - w_t) \\ &\Rightarrow \sum_{t=k}^{k+n}(\delta\rho)^{t-k} w_t < \sum_{t=k}^{k+n}(\delta\rho)^{t-k} w'_k. \end{aligned} \tag{33}$$

Step (3) (The no-reneging conditions). Define value functions (firms' expected profits) under contract $\{w'_t\}$ as V'_t. Note that an equation similar to (6) holds for V'_t. By construction, we have

$$\sum_{l=1}^{\infty}(\delta\rho)^{l-1} \frac{\phi_1}{\phi_l} w_l = \sum_{l=1}^{\infty}(\delta\rho)^{l-1} \frac{\phi_1}{\phi_l} w'_l. \tag{34}$$

Hence, $V_1 = V'_1$. Thus two contracts have the same expected payments. Similarly, one can show that $V_t = V'_t$ for all $t \le k$ and $t > k+n$. Because $V_t \ge V_1$ by assumption, $V'_t \ge V'_1$ for all $t \le k$ and $t > k+n$. Now what remains to be shown is $V'_t \ge V'_1 = V_1$ for $k+1 \le t \le k+n$. Suppose the opposite is true, i.e. $V'_{k+1} < V'_1$. Note that

$$\begin{aligned} V'_k &= (\frac{\phi_k}{\phi_{k+1}} - w'_k) + \delta[\rho \frac{\phi_k}{\phi_{k+1}} V'_{k+1} + (1 - \rho\frac{\phi_k}{\phi_{k+1}})V'_1] \ge V'_1 \\ &\Rightarrow (\frac{\phi_k}{\phi_{k+1}} - w'_k) \ge [1 - \delta(1 - \rho\frac{\phi_k}{\phi_{k+1}})]V'_1 - \delta\rho\frac{\phi_k}{\phi_{k+1}} V'_{k+1} \\ &\Rightarrow (\frac{\phi_k}{\phi_{k+1}} - w'_k) > [1 - \delta(1 - \rho\frac{\phi_k}{\phi_{k+1}})]V'_1 - \delta\rho\frac{\phi_k}{\phi_{k+1}} V'_1 = (1-\delta)V'_1. \end{aligned}$$

Then

$$\begin{aligned} V'_{k+n} &= (\frac{\phi_{k+n}}{\phi_{k+n+1}} - w'_k) + \delta[\rho\frac{\phi_{k+n}}{\phi_{k+n+1}} V'_{k+n+1} + (1 - \rho\frac{\phi_{k+n}}{\phi_{k+n+1}})V'_1] \\ &> (\frac{\phi_{k+n}}{\phi_{k+n+1}} - w'_k) + \delta V'_1 > (1-\delta)V'_1 + \delta V'_1 = V'_1. \end{aligned}$$

In the derivation, we used the fact that $\frac{\phi_{k+n}}{\phi_{k+n+1}} > \frac{\phi_k}{\phi_{k+1}}$ and $V'_{k+n+1} \ge V'_1$. By the same procedure, we can prove that

$$V'_{k+n-1} > V'_1 \Rightarrow V'_{k+n-2} > V'_1 \Rightarrow ... \Rightarrow V'_{k+1} > V'_1.$$

A contradiction. By similar arguments, we can prove that $V'_{k+m} > V'_1$ for any $1 \le m \le n$.

Step (4) (The no-shirking conditions). Define U'_t as type H worker's value function if he follows the equilibrium strategy under contract $\{w'_t\}$. By (38), the difference between U_t and U_1 is

$$U_t - U_1 = \sum_{l=t}^{\infty}(\delta\rho)^{l-t} w_l - \sum_{l=1}^{\infty}(\delta\rho)^{l-t} w_l \ge \widehat{c}.$$

Since $\{w'_t\}$ is nondecreasing,

$$U'_{t+1} - U'_t = \sum_{l=t}^{\infty} (\delta\rho)^{l-t}(w'_{l+1} - w'_l) \geq 0.$$

Then what remains to be shown is $U'_2 - U'_1 > \widehat{c}$. Note that if $k \geq 2$,

$$\begin{aligned} (U'_2 - U'_1) - (U_2 - U_1) &= \\ (\delta\rho)^{k-2}(1-\delta\rho)[\sum_{l=k}^{k+n}(\delta\rho)^{l-k}w'_k - \sum_{l=k}^{k+n}(\delta\rho)^{l-k}w_l] &> 0. \end{aligned}$$

The last inequality comes from (33). If $k = 1$, then

$$\begin{aligned} (U'_2 - U'_1) - (U_2 - U_1) &= \sum_{l=2}^{1+n}(\delta\rho)^{l-2}w'_1 - \sum_{l=2}^{1+n}(\delta\rho)^{l-2}w_l - [\sum_{l=1}^{1+n}(\delta\rho)^{l-1}w'_1 - \sum_{l=1}^{1+n}(\delta\rho)^{l-1}w_l] \\ &= (w_1 - w'_1) + (1-\delta\rho)[\sum_{l=2}^{1+n}(\delta\rho)^{l-2}w'_1 - \sum_{l=2}^{1+n}(\delta\rho)^{l-2}w_l] > 0. \end{aligned}$$

The last inequality comes from (33) and the fact that $w'_1 < w_1$. Therefore, $U'_2 - U'_1 > U_2 - U_1 \geq \widehat{c}$. The strict inequality implies that the no-shirking conditions are not binding for any t. ■

Proof of Theorem 2.

Proof. The proof is by construction. Suppose there is a $k > T$ such that $\pi_k > \pi_{k+1}$. The proof is divided into three cases. For different cases we use different constructions.

Case (1): $\pi_k = \frac{\phi_k}{\phi_{k+1}} - w_k < (1-\delta)V_1$**.** Our goal is to construct another self enforcing contract $\{w'_t\}$ with $\pi'_k = \pi'_{k+1}$. For that purpose, define

$$w'_{k+1} = \frac{\phi_{k+1}}{\phi_{k+2}} - (\frac{\phi_k}{\phi_{k+1}} - w_k); \; w'_{k+2} = w_{k+2} + (\delta\rho\frac{\phi_{k+1}}{\phi_{k+2}})^{-1}(w_{k+1} - w'_{k+1}).$$

Define a new contract $\{w'_t\} = \{w_1, ..., w_k, w'_{k+1}, w'_{k+2}, w_{k+3,}...\}$. Note that $0 < w_k < w'_{k+1} < w_{k+1}$, $w'_{k+2} > w_{k+2}$. By construction,

$$\pi_k = \frac{\phi_k}{\phi_{k+1}} - w_k = \frac{\phi_{k+1}}{\phi_{k+2}} - w'_{k+1} = \pi'_{k+1}.$$

Also notice that $\{w'_t\}$ is nondecreasing from 1 to $k+1$ and from $k+3$ on. The only concern is that w'_{k+2} may be bigger than w_{k+3}. If $w'_{k+2} \leq w_{k+3}$, then $\{w'_t\}$ is nondecreasing. Otherwise, redefine w'_{k+2} as the following

$$w'_{k+1} + \delta\rho\frac{\phi_{k+1}}{\phi_{k+2}}w'_{k+2}(1 + \delta\rho\frac{\phi_{k+2}}{\phi_{k+3}}) = w_{k+1} + \delta\rho\frac{\phi_{k+1}}{\phi_{k+2}}w_{k+2} + (\delta\rho)^2\frac{\phi_{k+1}}{\phi_{k+3}}w_{k+3}.$$

And redefine $\{w'_t\}$ as $\{w_1, w_{2,...}, w_k, w'_{k+1}, w'_{k+2}, w'_{k+2}, w_{k+4}, ...\}$. Under the new contract, if $w'_{k+2} \leq w_{k+4}$, then $\{w'_t\}$ is nondecreasing. Otherwise, redefine w'_{k+2} accordingly. Repeat this procedure until $\{w'_t\}$ is nondecreasing. Suppose that w_{k+n} is the last wage component that needs to be redefined. Note that it is necessary that $w_{k+j} < w'_{k+2}$ for all $2 \leq j \leq n$. By the construction, it immediately follows that

$$w'_{k+1} + \sum_{t=k+2}^{k+n} (\delta\rho)^{t-k} \frac{\phi_{k+1}}{\phi_t} w'_{k+2} = w_{k+1} + \sum_{t=k+2}^{k+n} (\delta\rho)^{t-k} \frac{\phi_{k+1}}{\phi_t} w_t. \tag{35}$$

According to (6), $V'_1 = V_1$, and $V'_t = V_t$ for all $t \leq k+1$ and $t > k+n$. By the fact that $\{w_t\}$ is self-enforcing, $V'_t \geq V'_1$ for all $t \leq k+1$ and $t > k+n$. To prove that $\{w'_t\}$ satisfies firms' no-reneging conditions, what remains to be shown is that $V'_{k+j} \geq V'_1 = V_1$ for $2 \leq j \leq n$. But

$$\begin{aligned}
V'_{k+1} &= (\frac{\phi_{k+1}}{\phi_{k+2}} - w'_{k+1}) + \delta[\rho\frac{\phi_{k+1}}{\phi_{k+2}} V'_{k+2} + (1-\rho\frac{\phi_{k+1}}{\phi_{k+2}})V_1] \geq V_1 \\
&\Rightarrow \delta\rho\frac{\phi_{k+1}}{\phi_{k+2}} V'_{k+2} \geq (1-\delta+\delta\rho\frac{\phi_{k+1}}{\phi_{k+2}})V_1 - (\frac{\phi_k}{\phi_{k+1}} v - w_k) \\
&\geq (1-\delta+\delta\rho\frac{\phi_{k+1}}{\phi_{k+2}})V_1 - (1-\delta)V_1 = \delta\rho\frac{\phi_{k+1}}{\phi_{k+2}} V_1 \\
&\Rightarrow V'_{k+2} \geq V_1 = V'_1.
\end{aligned}$$

The second line uses the fact that $\frac{\phi_k}{\phi_{k+1}} - w_k = \frac{\phi_{k+1}}{\phi_{k+2}} - w'_{k+1}$, and the third line uses the fact that $\frac{\phi_k}{\phi_{k+1}} - w_k \leq (1-\delta)V_1$.

Now suppose that $V'_{k+3} < V_1$. Then

$$\begin{aligned}
V'_{k+2} &= (\frac{\phi_{k+2}}{\phi_{k+3}} - w'_{k+2}) + \delta[\rho\frac{\phi_{k+2}}{\phi_{k+3}} V'_{k+3} + (1-\rho\frac{\phi_{k+2}}{\phi_{k+3}})V_1] \geq V_1 \\
&\Rightarrow (\frac{\phi_{k+2}}{\phi_{k+3}} - w'_{k+2}) \geq [1-\delta(1-\rho\frac{\phi_{k+2}}{\phi_{k+3}})]V_1 - \delta\rho\frac{\phi_{k+2}}{\phi_{k+3}} V'_{k+3} > (1-\delta)V_1.
\end{aligned}$$

This implies that

$$\begin{aligned}
V'_{k+n} &= (\frac{\phi_{k+n}}{\phi_{k+n+1}} - w'_{k+2}) + \delta[\rho\frac{\phi_{k+n}}{\phi_{k+n+1}} V'_{k+n+1} + (1-\rho\frac{\phi_{k+n}}{\phi_{k+n+1}})V_1] \\
&> (\frac{\phi_{k+n}}{\phi_{k+n+1}} - w'_{k+2}) + \delta V_1 > (1-\delta)V_1 + \delta V'_1 = V_1.
\end{aligned}$$

In the derivation, we use the fact that $\frac{\phi_{k+n}}{\phi_{k+n+1}} > \frac{\phi_{k+2}}{\phi_{k+3}}$ and $V'_{k+n+1} \geq V_1$. By the same procedure, we can prove that

$$V'_{k+n} > V_1 \Rightarrow V'_{k+n-1} > V_1 \Rightarrow ... \Rightarrow V'_{k+3} > V_1.$$

A contradiction. Therefore, $V'_{k+3} \geq V'_1$. By similar arguments, we can prove that $V'_{k+j} \geq V'_1$ for any $3 \leq j \leq n$.

Now we show that $\{w'_t\}$ also (strictly) satisfies H workers' no-shirking conditions. By the fact that $\{w'_t\}$ is nondecreasing, we only needs to show that $(U'_2-U'_1)-(U_2-U_1)>0$, which is essentially equivalent to

$$w'_{k+1}+\sum_{t=k+2}^{k+n}(\delta\rho)^{t-k-1}w'_{k+2}>w_{k+1}+\sum_{t=k+2}^{k+n}(\delta\rho)^{t-k-1}w_t. \tag{36}$$

By (35),

$$(w_{k+1}-w'_{k+1})=\sum_{t=k+2}^{k+n}(\delta\rho)^{t-k-1}\frac{\phi_{k+1}}{\phi_t}(w'_{k+2}-w_t).$$

By the fact that $w_{k+j}<w'_{k+2}$ for all $2\leq j\leq n$, we get

$$(w_{k+1}-w'_{k+1})<\sum_{t=k+2}^{k+n}(\delta\rho)^{t-k-1}(w'_{k+2}-w_t).$$

Thus (36) is satisfied. Therefore, $\{w'_t\}$ is self-enforcing. Moreover, the strict inequality in (36) implies that the no-shirking conditions do not bind at any $t\geq 2$.

Case (2): $\pi_{k+1}=\frac{\phi_{k+1}}{\phi_{k+2}}-w_{k+1}\geq(1-\delta)V_1$**.**

The construction of $\{w'_t\}$ is a mirror image of case (1). Define

$$w'_k=\frac{\phi_k}{\phi_{k+1}}-(\frac{\phi_{k+1}}{\phi_{k+2}}-w_{k+1});\ w'_{k-1}=w_{k-1}-(\delta\rho\frac{\phi_{k-1}}{\phi_k})(w'_k-w_k).$$

Define a new contract $\{w'_t\}=\{w_1,...,w_{k-2},w'_{k-1},w'_k,w_{k+1},..\}$. Note that $w'_k>w_k$ and $w'_{k-1}<w_{k-1}$. Following the construction,

$$\pi'_k=\frac{\phi_k}{\phi_{k+1}}-w'_k=\frac{\phi_{k+1}}{\phi_{k+2}}-w_{k+1}=\pi_{k+1}.$$

Also note that $\{w'_t\}$ is nondecreasing from 1 to $k-2$ and from $k-1$ on. The only problem is that w'_{k-1} may be less than w_{k-2}. If $w'_{k-1}\geq w_{k-2}$, then $\{w'_t\}$ is nondecreasing and $w'_{k-1}\geq 0$. Otherwise, redefine w'_{k-1} as the following

$$w'_{k-1}(1+\delta\rho\frac{\phi_{k-2}}{\phi_{k-1}})+(\delta\rho)^2\frac{\phi_{k-2}}{\phi_k}w'_k=w_{k-2}+\delta\rho\frac{\phi_{k-2}}{\phi_{k-1}}w_{k-1}+(\delta\rho)^2\frac{\phi_{k-2}}{\phi_k}w_k.$$

And redefine $\{w'_t\}$ as $\{w_1,...,w_{k-3},w'_{k-1},w'_{k-1},w'_k,w_{k+1},..\}$. Repeat this procedure until $\{w'_t\}$ is nondecreasing. Suppose that w_{k-n} is the last wage component that needs to be redefined. Note that it is necessary that $w_{k-j}>w'_{k-1}$ for all $1\leq j\leq n$.

Due to the fact that $w_t=0$ for all $t<T$, if $k-n<T$ then the redefined $w'_{k-1}<0$, which violates the constraint that $w_t\geq 0$ for any t. Therefore, we need to consider two subcases.

Subcase (a): $k-n\geq T$.

In this subcase, the redefined $w'_{k-1} \geq 0$. Hence the constructed $\{w'_t\}$ is nondecreasing and satisfies the non-negativity constraints. By the construction, it immediately follows that

$$\sum_{t=k-n}^{k-1} (\delta\rho)^{t-k+n} \frac{\phi_{k-n}}{\phi_t} w'_{k-1} + (\delta\rho)^n \frac{\phi_{k-n}}{\phi_k} w'_k = \sum_{t=k-n}^{k} (\delta\rho)^{t-k+n} \frac{\phi_{k-n}}{\phi_t} w_t. \tag{37}$$

According to (6), $V'_1 = V_1$, and $V'_t = V_t$ for all $t \leq k-n$ and $t > k$. Moreover, $V'_t \geq V'_1$ for all $t \leq k-n$ and $t > k$. To prove that $\{w'_t\}$ satisfies firms' no-reneging conditions, what remains to be shown is that $V'_{k-j} \geq V'_1 = V_1$ for $0 \leq j \leq n-1$. But

$$\begin{aligned} V'_k &= \left(\frac{\phi_k}{\phi_{k+1}} - w'_k\right) + \delta[\rho \frac{\phi_k}{\phi_{k+1}} V'_{k+1} + (1 - \rho \frac{\phi_k}{\phi_{k+1}}) V_1] \\ &\geq (1-\delta)V_1 + \delta[\rho \frac{\phi_k}{\phi_{k+1}} V_1 + (1 - \rho \frac{\phi_k}{\phi_{k+1}}) V_1] = V_1. \end{aligned}$$

In deriving this, we use the facts that $V'_{k+1} \geq V_1$ and $(\frac{\phi_k}{\phi_{k+1}} - w'_k) = \frac{\phi_{k+1}}{\phi_{k+2}} - w_{k+1} \geq (1-\delta)V_1$. Using arguments similar to those in case (1), we can prove that $V'_{k-j} > V_1$ for any $1 \leq j \leq n-1$.

To prove that $\{w'_t\}$ strictly satisfies H type worker's no-shirking conditions, it is enough to show that $(U'_2 - U'_1) - (U_2 - U_1) > 0$, which is equivalent to

$$\sum_{t=k-n}^{k-1} (\delta\rho)^{t-k+n} w'_{k-1} + (\delta\rho)^n w'_k > \sum_{t=k-n}^{k} (\delta\rho)^{t-k+n} w_t.$$

By (37),

$$\begin{aligned} & \sum_{t=k-n}^{k-1} (\delta\rho)^{t-k+n} \frac{\phi_{k-n}}{\phi_t} (w_t - w'_{k-1}) = (\delta\rho)^n \frac{\phi_{k-n}}{\phi_k} (w'_k - w_k) \\ \Rightarrow \quad & \sum_{t=k-n}^{k-1} (\delta\rho)^{t-k+n} \frac{\phi_k}{\phi_t} (w_t - w'_{k-1}) = (\delta\rho)^n (w'_k - w_k) \\ \Longrightarrow \quad & \sum_{t=k-n}^{k-1} (\delta\rho)^{t-k+n} (w_t - w'_{k-1}) < (\delta\rho)^n (w'_k - w_k). \end{aligned}$$

The last inequality uses the fact that $w_{k-j} > w'_{k-1}$ for all $1 \leq j \leq n$. Therefore, $\{w'_t\}$ also strictly satisfies H type worker's no-shirking conditions.

Subcase (b): $k - n < T$.

In this case, the redefined $w'_{k-1} < 0$. We need to use another construction $\{w'_t\}$. Let $w'_t = 0$ for all $T \leq t \leq k-1$ and

$$w'_k = w_k + \frac{\sum_{t=T}^{k-1} (\delta\rho)^{t-T} \frac{\phi_T}{\phi_t} w_t}{(\delta\rho)^{k-T} \frac{\phi_T}{\phi_k}};$$

and all the other wages remain the same. By the construction, we have $\pi'_k > \pi'_{k+1}$, since otherwise we would have subcase (a). This also implies that $w'_k < w'_{k+1}$, thus the constructed $\{w'_t\}$ is nondecreasing and satisfies the non-negativity constraints.

By the construction and (6), we have $V'_1 = V_1$, and $V'_t = V_t$ for all $t \leq T$ and $t > k$. By the fact that $\{w_t\}$ is self-enforcing, using similar argument as in subcase (a), we can show that $V'_t \geq V'_1 = V_1$ for $T < t \leq k$. Thus $\{w'_t\}$ satisfies firms' no-reneging conditions. Applying a similar argument to that in subcase (a), we can show that $\{w'_t\}$ also strictly satisfies H type worker's no-shirking conditions.

Case (3): $\pi_{k+1} = \frac{\phi_{k+1}}{\phi_{k+2}} - w_{k+1} < (1-\delta)V_1 < \frac{\phi_k}{\phi_{k+1}} - w_k = \pi_k$**.**

The construction in this case is a combination of those in case (1) and case (2). The proof is also a combination of Case (1) and Case (2). Therefore, it is omitted.

By repeating the procedure specified above, for any nondecreasing and self-enforcing contract $\{w_t\}$, we can construct a quasi-monotonic contract $\{w'_t\}$ such that it is self-enforcing, $V'_1 = V_1$, and H-type workers' no-shirking conditions are strictly satisfied.

Finally, suppose there is a nondecreasing and self-enforcing contract $\{w_t\}$ satisfying $\pi_{t+1} \geq \pi_t$ for any $t > T$, $\pi_T > \pi_{T+1}$, and $\pi_T < (1-\delta)V_1$. Note that this contract is not quasi-monotonic. By applying the construction in Case (1), we can find a quasi-monotonic and self-enforcing contract $\{w'_t\}$, with $\pi_{t+1} \geq \pi_t$ for any $t \geq T$. ■

Proof of Lemma 3.

Proof. (i) The no-shirking conditions can be reduced to (15). First we show that under $\{w_t\}$, $U_t \geq U_2$ for any $t \geq 3$. By (3),

$$U_t = \frac{\delta(1-\rho)}{1-\delta\rho}U_1 - \frac{1}{1-\delta\rho}c + \sum_{l=t}^{\infty}(\delta\rho)^{l-t}w_l. \tag{38}$$

By (38), for $t \geq 3$

$$U_t - U_2 = \sum_{j=t}^{\infty}(\delta\rho)^{j-t}w_j - \sum_{j=2}^{\infty}(\delta\rho)^{j-2}w_j = \sum_{j=0}^{\infty}(\delta\rho)^j(w_{t+j} - w_{2+j}) \geq 0,$$

since w_t is nondecreasing in t. Therefore, the no-shirking conditions boil down to $U_2 - U_1 \geq \widehat{c}$, which can be explicitly written as (15).

(ii) If $\pi_{t+1} \geq \pi_t$ for any $t \geq T$, then $V_T \geq V_1$ implies firms' no-reneging conditions (9). To prove this claim, we first show that $V_t \geq V_1$ for any $t > T$. Suppose $V_{T+1} < V_1$. Then, combining with $V_T = \pi_T + \delta[\rho\frac{\phi_T}{\phi_{T+1}}V_{T+1} + (1-\rho\frac{\phi_T}{\phi_{T+1}})V_1] \geq V_1$, we have $\pi_T > (1-\delta)V_1$. But

$$\begin{aligned} V_{T+1} &= \sum_{j=T+1}^{\infty}(\delta\rho)^{j-(T+1)}\frac{\phi_{T+1}}{\phi_j}\pi_j + \delta[1-(1-\delta)\sum_{j=T+1}^{\infty}(\delta\rho)^{j-T}\frac{\phi_{T+1}}{\phi_{j+1}}]V_1 \\ &\geq \sum_{j=T+1}^{\infty}(\delta\rho)^{j-(T+1)}\frac{\phi_{T+1}}{\phi_j}(1-\delta)V_1 + \delta[1-(1-\delta)\sum_{j=T+1}^{\infty}(\delta\rho)^{j-T}\frac{\phi_{T+1}}{\phi_{j+1}}]V_1 = V_1, \end{aligned}$$

where the inequality follows $\pi_{t+1} \geq \pi_t$ for any $t \geq T$ and $\pi_T > (1-\delta)V_1$. A contradiction. Therefore, $V_{T+1} \geq V_1$. By similar arguments, we can recursively show that $V_t \geq V_1$ for any $t > T$.

Next we show that $V_t \geq V_1$ for any $t < T$ (this step is necessary only if $T > 2$). Suppose $V_{T-1} < V_1$. Then $V_{T-1} = \pi_{T-1} + \delta[\rho\frac{\phi_{T-1}}{\phi_T}V_T + (1-\rho\frac{\phi_{T-1}}{\phi_T})V_1] < V_1$ and $V_T \geq V_1$ implies that $\pi_{T-1} < (1-\delta)V_1$. Since π_t is increasing in the domain from 1 to T, We have $V_{T-2} = \pi_{T-2} + \delta[\rho\frac{\phi_{T-2}}{\phi_{T-1}}V_{T-1} + (1-\rho\frac{\phi_{T-2}}{\phi_{T-1}})V_1] < V_1$. Apply this argument recursively, eventually we have, $V_1 = \pi_1 + \delta[\rho\frac{\phi_1}{\phi_2}V_2 + (1-\rho\frac{\phi_1}{\phi 2})V_1] < V_1$, a contradiction. Therefore, $V_{T-1} \geq V_1$. By similar arguments, we can show that $V_t \geq V_1$ for any $t < T$.

(3) If $\pi_{t+1} \geq \pi_t$ for any $t > T$ and $\pi_T > \pi_{T+1}$, then $V_{T+1} \geq V_1$ implies that $V_t \geq V_1$ for any t. Recall that this type of quasi-monotonic contract must have $\pi_T \geq (1-\delta)V_1$. Combining with $V_{T+1} \geq V_1$, we have $V_T \geq V_1$. The rest of the proof is the same as those in (2). ■

Proof of Lemma 6.

Proof. It is easy to verify that $f(0) = f(1) = 0$, since $\phi_t = 0$ (for all t) if $\phi_1 = 0$ and $\phi_t = 1$ (for all t) if $\phi_1 = 1$. Expand $f(\phi_1)$ and take derivative with respect to ϕ_1,

$$\begin{aligned}\frac{df}{d\phi_1} &= -1 + (1-\delta\rho)\sum_{t=1}^{\infty}(\delta\rho)^{t-1}\frac{1}{\phi_1 + p^t(1-\phi_1)} \\ &\quad +\phi_1(1-\delta\rho)\sum_{t=1}^{\infty}(\delta\rho)^{t-1}\frac{p^t-1}{[\phi_1+p^t(1-\phi_1)]^2}.\end{aligned}$$

Note that

$$\begin{aligned}\frac{df}{d\phi_1}(0) &= (1-\delta\rho)\sum_{t=1}^{\infty}(\delta\rho)^{t-1}(\frac{1}{p^t}-1) > 0; \\ \frac{df}{d\phi_1}(1) &= (1-\delta\rho)\sum_{t=1}^{\infty}(\delta\rho)^{t-1}(p^t-1) < 0.\end{aligned}$$

Take second derivative,

$$\frac{d^2f}{d\phi_1^2} = 2(1-\delta\rho)\sum_{t=1}^{\infty}(\delta\rho)^{t-1}\{\frac{p^t-1}{[\phi_1+p^t(1-\phi_1)]^2} + \phi_1\frac{(p^t-1)^2}{[\phi_1+p^t(1-\phi_1)]^3}\}.$$

Note that the term in the bracket

$$\frac{(p^t-1)}{[\phi_1+p^t(1-\phi_1)]^2} + \phi_1\frac{(p^t-1)^2}{[\phi_1+p^t(1-\phi_1)]^3} = \frac{-(1-p^t)p^t}{[\phi_1+p^t(1-\phi_1)]^3} < 0$$

Therefore, $\frac{d^2f}{d\phi_1^2} < 0$. ■

References

[1] Abowd, John, Francis Kramarz and David Margolis. "High Wage and High Wage Firms." *Econometrica*, 1999, Mar., 251-333.

[2] Abraham, Katharine and Henry Farber. "Job Duration, Seniority, and Earnings." *American Economic Review*, 1987, Jun., 278-97.

[3] Altonji, Joseph and Robert Shakotko. "Do Wages Rise with Job Seniority." *Review of Economic Studies*, 1987, Jul., 437-59.

[4] Baker, George, Robert Gibbons and Kevin Murphy. "Subjective Performance Measures in Optimal Incentive Contracts," *Quarterly Journal of Economics*, 1994, Nov., 1125-56.

[5] Becker, Gary. "Investment in Human Capital: A Theoretical Analysis." *Journal of Political Economy*, 1962, Oct., 9-49.

[6] Becker, Gary. *Human Capital*. Second edition, 1975, Chicago, University of Chicago Press.

[7] Berger, Allen and Gregory Udell. "Relationship Lending and Lines of Credit in Small Firm Finance." *Journal of Business*, 1995, Jun., 351-81.

[8] Bernhardt, Dan. "Strategic Promotion and Compensation." *Review of Economic Studies*, 1995, Apr., 315-39.

[9] Bodenhorn, Howard. "Short-Term Loans and Long-Term Relationships: Relationship Lending in Early American." *Journal of Money, Credit and Banking*, 2003, Aug., 485-505.

[10] Brown, James. "Why Do Wages Increase with Tenure? On-the-Job Training and Life-Cycle Wage Growth Observed Within Firms." *American Economic Review*, 1989, Dec., 971-91.

[11] Bull, Clive. "The Existence of Self-Enforcing Implicit Contracts." *Quarterly Journal of Economics*, 1987, Feb., 147-59.

[12] Burdett, Ken and Melvyn Coles. "Equilibrium Wage-Tenure Contracts." *Econometrica*, 2003, Sept., 1377-1404.

[13] Dutta, "Building Trust", 1993, Mimeo. London School of Economics.

[14] Farber, Henry and Robert Gibbons. "Learning and Wage Dynamics." *Quarterly Journal of Economics*, 1996, Nov., 1007-47.

[15] Felli, Leonardo and Chris Harris. "Learning, Wage Dynamics, and Firm-Specific Human Capital." *Journal of Political Economy*, 1996, Nov., 838-68.

[16] Fuchs, William. "Contracting with Repeated Moral Hazard and Private Evaluations," *American Economic Review*, 97(4), 2007, 1432-1448.

[17] Ghosh, P. and Ray, D. "Cooperation in Community Interaction without Information Flows", *Review of Economic Studies*, 1996, 63(3), 491-519.

[18] Greenwald, Bruce. "Adverse Selection in the Labor Market." *Review of Economic Studies*, 1986, Jun., 325-47.

[19] Harris, Milton and Bengt Holmstrom. "A Theory of Wage Dynamics." *Review of Economic Studies*, 1982, Jul., 315-33.

[20] Hashimoto, Masanori. "Firm-Specific Human Capital as a Shared Investment." *American Economic Review*, 1981, Jun., 475-82.

[21] Holmstrom, Bengt. "Managerial Incentive Problems: A Dynamic Perspective." *Review of economic Studies*, 1999, Jan., 169-82.

[22] Jovanovic, Boyan. "Job Matching and the Theory of Turnover." *Journal of Political Economy*, 1979, Oct., 972-90.

[23] Laffont, J.J. and J. Tirole. "Adverse Selection and Renegotiation in Procurement," Review of Economic Studies, 1990, 597-626.

[24] Lazear, Edward. "Why Is There Mandatory Retirement." *Journal of Political Economy*, 1979, Dec., 1261-84.

[25] Levin, Jonathan. "Relational Incentive Contracts." *American Economic Review*, 2003, Jun.

[26] MacLeod, Bentley and James Malcomson. "Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment." *Econometrica*, 1989, Mar., 447-80.

[27] MacLeod, Bentley and James Malcomson. "Motivation and Markets." *American Economic Review*, 1998, Jun., 388-411.

[28] Medoff, James and Karharine Abraham. "Experience, Performance, and Earnings." *Quarterly Journal of Economics*, 1980, Dec., 703-36.

[29] Mailath, George and Larry Samuelson. "Who Wants a Good Reputation." *Review of Economic Studies*, Apr. 2001, 415-441.

[30] Mincer, Jacob. *Schooling, Experience, and Earnings*. 1974, New York, NBER.

[31] Pearce, James. "Tenure, Unions, and the Relationship between Employer Size and Wages." *Journal of Labor Economics*, 1990, summer, 251-269.

[32] Rob, Rafael and Huanxing Yang. "Long-Term Relationships as Safeguards," 2005, Working paper, University of Pennsylvania.

[33] Waldman, Michael. "Job Assignments, Signaling and Efficiency." *Rand Journal of Economics*, 1984, Summer, 255-67.

[34] Yang, Huanxing. "Efficiency Wages and Subjective Performance Pay." *Economic Inquiry*, 46(2), 2008, 179-196.