Online Appendix to "Nonstationary Relational Contracts with Adverse Selection"

Huanxing Yang, Ohio State University

January 2012

1 Missing Proofs

Proof of Proposition 1:

We prove Proposition 1 by proving two lemmas.

Lemma A1: If a contract $\{w_t\}$ is self-enforcing, then there is another self-enforcing contract $\{w'_t\}$ such that: (i) w'_t is nondecreasing in t, (ii) the no-shirking conditions, $U_t - U_1 \geq \widehat{c}$ for all $t \geq 2$, do not bind for any $t \geq 2$, (iii) firms' expected (discounted) profits are the same under two contracts, $V_1 = V'_1$.

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 $\{w'_t\}$.

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

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

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 $1 \leq m \leq n$,

$$\sum_{t=k+m}^{k+n} (\delta\rho)^{t-k-m} \frac{1}{\phi_t} (w'_k - w_t) > 0. \tag{2}$$

By the fact that $w_k > w'_k$ and (1), (2) holds for $m = 1$. It follows that (2) holds for $k+m \leq z$, since by removing one negative term the inequality should also hold. For $k+m > z$ (2) 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 (1) and (2),

$$\begin{aligned} w_k - w'_k &= \sum_{t=k+1}^{k+n} (\delta\rho)^{t-k} \frac{\phi_k}{\phi_t} (w'_k - w_t) < \sum_{t=k+1}^{k+n} (\delta\rho)^{t-k} \frac{\phi_{k+1}}{\phi_t} (w'_k - w_t) \\ &< \delta\rho(w'_k - w_{k+1}) + \sum_{t=k+2}^{k+n} (\delta\rho)^{t-k} \frac{\phi_{k+2}}{\phi_t} (w'_k - w_t) < ... < \sum_{t=k+1}^{k+n} (\delta\rho)^{t-k} (w'_k - w_t) \\ &\Rightarrow \sum_{t=k}^{k+n} (\delta\rho)^{t-k} w_t < \sum_{t=k}^{k+n} (\delta\rho)^{t-k} w'_k. \end{aligned} \tag{3}$$

Step (3) (The no-reneging conditions). Define value functions (firms' expected discounted profits) under contract $\{w'_t\}$ as V'_t. Recall the expression for V_t:

$$V_t = (\frac{\phi_t}{\phi_{t+1}} - w_t) + \delta[\rho \frac{\phi_t}{\phi_{t+1}} V_{t+1} + (1 - \rho \frac{\phi_t}{\phi_{t+1}}) V_1]. \tag{4}$$

Note that an equation similar to (4) holds for V'_t. By construction, we have

$$\sum_{l=1}^{\infty} (\delta\rho)^{l-1} \frac{\phi_1}{\phi_l} w_l = \sum_{l=1}^{\infty} (\delta\rho)^{l-1} \frac{\phi_1}{\phi_l} w'_l. \tag{5}$$

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

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

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 $V'_{k+n+1} \geq V'_1$. 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 $1 \leq m \leq n$.

Step (4) (The no-shirking conditions). Define U'_t as type H worker's value function if he follows the equilibrium strategy under contract $\{w'_t\}$. We can rewrite U_t as:

$$U_t = \frac{\delta(1-\rho)}{1-\delta\rho}U_1 - \frac{1}{1-\delta\rho}c + \sum_{l=t}^{\infty}(\delta\rho)^{l-t}w_l. \tag{6}$$

By (6), the difference between U_t and U_1 can be written as

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

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 (3). If $k = 1$, then

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

The last inequality comes from (3) 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. ■

By Lemma A1, without loss of generality we can focus on nondecreasing contracts.

Lemma A2: 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, $V_1 = V_1'$.

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

$$w_{k+1}' + \sum_{t=k+2}^{k+n}(\delta\rho)^{t-k}\frac{\phi_{k+1}}{\phi_t}w_{k+2}' = w_{k+1} + \sum_{t=k+2}^{k+n}(\delta\rho)^{t-k}\frac{\phi_{k+1}}{\phi_t}w_t. \tag{7}$$

According to (4), $V_1' = V_1$, 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

$$w'_{k+1} + \sum_{t=k+2}^{k+n} (\delta\rho)^{t-k-1} w'_{k+2} > w_{k+1} + \sum_{t=k+2}^{k+n} (\delta\rho)^{t-k-1} w_t. \tag{8}$$

By (7),

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

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 (8) is satisfied. Therefore, $\{w'_t\}$ is self-enforcing. Moreover, the strict inequality in (8) 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

$$\sum_{t=k-n}^{k-1}(\delta\rho)^{t-k+n}\frac{\phi_{k-n}}{\phi_t}w_{k-1}' + (\delta\rho)^n\frac{\phi_{k-n}}{\phi_k}w_k' = \sum_{t=k-n}^{k}(\delta\rho)^{t-k+n}\frac{\phi_{k-n}}{\phi_t}w_t. \tag{9}$$

According to (4), $V_1' = V_1$, 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

$$\begin{aligned} V_k' &= (\frac{\phi_k}{\phi_{k+1}} - w_k') + \delta[\rho\frac{\phi_k}{\phi_{k+1}}V_{k+1}' + (1-\rho\frac{\phi_k}{\phi_{k+1}})V_1] \\ &\geq (1-\delta)V_1 + \delta[\rho\frac{\phi_k}{\phi_{k+1}}V_1 + (1-\rho\frac{\phi_k}{\phi_{k+1}})V_1] = V_1. \end{aligned}$$

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 (9),

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

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 (4), 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 $t\geq T$. ■

2 Separating Contracts

2.1 Analysis

Given that there are two types of workers, theoretically speaking firms could offer separating contracts: firms offers two contracts and let workers 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 contract designed for H type workers, the payment plan evolves according to $\{w^s_t\}$. In short, we denote a separating contract as $(w_L,\{w^s_t\})$. In the remaining of this section we just provide the main results regarding separating contracts. The detailed analysis can be found in an online appendix.

With separating contracts $(w_L,\{w^s_t\})$, 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 1 *A (trigger strategy) high-effort equilibrium with separating contract* $(w_L, \{w_t^s\})$ *satisfies: (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* $\bar{e}$ *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 offering the zero-wage contract.*

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 an 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 the corresponding lemma with pooling contracts.

Lemma 1 *A high-effort equilibrium with separating contracts exists if and only if: (i) there is a self-enforcing separating contract $(w_L, \{w_t^s\})$ (it satisfies (11)-(14)) (ii) under the contract $(w_L, \{w_t^s\})$, $V_N \geq p/(1-\delta)$.*

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:

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

$$\sum_{t=j}^{\infty}(\delta\rho)^{t-j}(1-w_t^s) - \frac{1}{1-\delta\rho(1-\lambda)}\{\lambda\sum_{t=1}^{\infty}(\delta\rho)^{t-1}(1-w_t^s) \tag{16}$$
$$+(1-\lambda)[p-(1-\delta\rho p)\sum_{t=1}^{\infty}(\delta\rho p)^{t-1}w_t^s]\} \geq 0, \text{ for } j > T_s.$$

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 2 *If a separating contract* $(w_L, \{w_t^s\})$ *is self-enforcing (satisfies (15) and (16)), 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 (15) 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 nondecreasing 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, (16) still holds. Now consider the change of the LHS of (15):

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

Thus under $\{w_t^{s\prime}\}$, (15) 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 (16) 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 nondecreasing 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 (16) still holds under $\{w_t^{s\prime}\}$. Now consider the change of the LHS of (15):

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

Thus under $\{w_t^{s\prime}\}$, (15) holds with strict inequality. Therefore, both $\{w_t^{s\prime}\}$ and $\{w_t^s\}$ are not optimal. ■

2.2 Optimal Separating Contracts

The class of contracts described in Lemma 2 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 $w_{T_s} \leq w_{T_s+1}$):

$$0 = \frac{1-w_{T_{s+1}}}{1-\delta\rho} - \frac{1}{1-\delta\rho(1-\lambda)}\{\lambda[\frac{1-(\delta\rho)^{T_s-1}}{1-\delta\rho} + (\delta\rho)^{T_s-1}(1-w_{T_s}) + (\delta\rho)^{T_s}\frac{1-w_{T_{s+1}}}{1-\delta\rho}]$$
$$+(1-\lambda)[p - (1-\delta\rho p)[(\delta\rho p)^{T_s-1}w_{T_s} + \frac{(\delta\rho p)^{T_s}w_{T_{s+1}}}{1-\delta\rho p}]\} \quad (17)$$

From (17), we can see that $w_{T_{s+1}}$ is increasing in w_{T_s} and decreasing in T_s. Define the LHS of (15) as

$$G_s(T_s, w_{T_s}) = (\delta\rho)^{T_s-2}w_{T_s} + (\delta\rho)^{T_s-1}(w_{T_{s+1}} - w_{T_s}). \quad (18)$$

It can be verified that $G_s(T_s, w_{T_s})$ is increasing in w_{T_s}. Subject to $w_{T_s} \leq w_{T_s+1}$, $G_s(T_s, w_{T_s})$ 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,

$$g_s(T_s) = (\delta\rho)^{T_s-2}w_{T_s};\ \text{where } w_{T_s} \text{ satisfies } 0 = \frac{1-w_{T_s}}{1-\delta\rho} - \frac{1}{1-\delta\rho(1-\lambda)}$$
$$\times\{\lambda[\frac{1-(\delta\rho)^{T_s-1}}{1-\delta\rho} + (\delta\rho)^{T_s-1}\frac{1-w_{T_s}}{1-\delta\rho}] + (1-\lambda)[p - (\delta\rho p)^{T_s-1}w_{T_s}]\}. \quad (19)$$

From (19), we see that w_{T_s} is decreasing in T_s. Thus $g_s(T_s)$ is decreasing in T_s. Therefore, self-enforcing separating contracts exist if and only if $g_s(2) \geq \widehat{c}$. This condition can be written more explicitly as

$$\frac{c}{\delta\rho(1-p)^2} \leq \frac{(1-\lambda)}{1-(1-\lambda)\delta\rho p}. \tag{20}$$

Note that the RHS of (20) is decreasing in λ. Therefore, a necessary condition for (20) to be satisfied is that it is satisfied for $\lambda = 0$, or equivalently

$$\frac{c}{\delta\rho(1-p)^2} \leq \frac{1}{1-\delta\rho p} \tag{21}$$

Note that the RHS of (20) is 0 when $\lambda = 1$. Thus (20) cannot be satisfied if $\lambda = 1$. If condition (21) is satisfied, then there is a $\widehat{\lambda} \in [0, 1)$ such that (20) 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, (20) is satisfied if and only if $\beta \in [\widehat{\beta}, 1]$, where $\widehat{\beta}$ is defined as $\lambda(\widehat{\beta}) = \widehat{\lambda}$.

Now suppose (20) 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 $g_s(T_s^*) \geq \widehat{c}$ 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 1 *A high-effort equilibrium with separating contracts exist if and only if (21) holds, $\beta \in [\widehat{\beta}, 1]$ with $\widehat{\beta} \in (0, 1)$, and $V_N \geq p/(1-\delta)$. If 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 $t \geq T_s^* + 1$, $w_{T_s}^*$ and $w_{T_s+1}^*$ satisfy (17), and T_s^* is determined by $g_s(T_s^*) \geq \widehat{c}$ and $g_s(T_s^* + 1) < \widehat{c}$.*

Proposition 1 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 **??**. 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.

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.

2.3 Comparison

Now we compare pooling contracts and separating contracts.

Lemma 3 *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,

$$f(\phi_1^*) \geq \frac{\widehat{c}}{(1-p)} = \frac{c}{\delta\rho(1-p)^2}, \tag{22}$$

is more stringent than the necessary condition for a self-enforcing separating contract to exist, (21). Specifically, if (22) 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 (21) holds as well.

Now suppose that (22) 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 the fact that $\phi_1(\beta) = \frac{(1-\rho)(1-\beta)}{(1-\rho)(1-\beta)+(1-\rho p)\beta}$ 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,

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

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 3 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 4 *Suppose (22) 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

$$V_1 = \phi_1\left[\sum_{t=1}^{\infty}(\delta\rho)^{t-1}(1-w_t^*) + \frac{\delta(1-\rho)}{1-\delta\rho}V_1\right] + (1-\phi_1)\left[\sum_{t=1}^{\infty}(\delta\rho p)^{t-1}(p-w_t^*) + \frac{\delta(1-\rho p)}{1-\delta\rho p}V_1\right]. \tag{23}$$

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

$$V_N = \phi_1[\sum_{t=1}^{\infty}(\delta\rho)^{t-1}(1-w_t^*) + \frac{\delta(1-\rho)}{1-\delta\rho}V_N] + (1-\phi_1)[(1-\delta\rho p)\sum_{t=1}^{\infty}(\delta\rho p)^{t-1}(p-w_t^*) + \delta V_N]. \tag{24}$$

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 (24) we have

$$V_N > \phi_1[\sum_{t=1}^{\infty}(\delta\rho)^{t-1}(1-w_t^*) + \frac{\delta(1-\rho)}{1-\delta\rho}V_N] + (1-\phi_1)[\sum_{t=1}^{\infty}(\delta\rho p)^{t-1}(p-w_t^*) + \frac{\delta(1-\rho p)}{1-\delta\rho p}V_N]. \tag{25}$$

Now compare (23) and (25), we can clearly see that $V_N > V_1^*$. Therefore we have $V_N^* \geq V_N > V_1^*$, a contradiction. ■

Lemma 4 implies that for any initial belief ϕ_1 such that high-effort equilibria exist under pooling contracts, high effort equilibria exist under separating contracts as well. Moreover, firms' expected discounted profits are higher under the optimal separating contract than under the optimal pooling contract.[1] The following proposition summarizes the comparison of pooling contracts and separating contracts.

Proposition 2 *Compared to pooling contracts, high-effort equilibria exist for a wider range of parameter values under separating contracts. Moreover, if higher-effort equilibria exist under both contracts, the optimal separating contract yields a higher profit for firms.*

The intuition for Proposition 2 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.

[1] 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, $V_1^* < V_N^*$.