Internet Appendix for "Adverse Selection and Assortative Matching in Labor Markets"

January 2017

1. Mixed-strategy Equilibria

We relax Assumption A2 to allow for the possibility of mixed-strategy equilibria. In a mixed-strategy equilibrium, a type-τ_i worker who is indifferent between accepting or rejecting a poaching offer (i.e., an offer such that $w^p(w_i) = w_i$) rejects the poaching offer with probability $p_i(w_i)$. We then obtain the following result:

Result 1 *In any equilibrium, $p_i(w_i)$ is non-decreasing in w_i.*

Proof. Suppose that there is an equilibrium in which $w_i' = w^p(w_i') > w_i = w^p(w_i)$. In such an equilibrium,

$$E[\tau_i \mid w_i'] \equiv \int_0^{\overline{\tau}} \tau dF(\tau \mid w_i') > \int_0^{\overline{\tau}} \tau dF(\tau \mid w_i) \equiv E[\tau_i \mid w_i], \quad (1)$$

(because of $w^p(w, i, W) = \theta\gamma\left(\int_0^{\overline{\tau}} \tau dF^W(\tau \mid w, i) - \mu\right)$ and Bayesian rationality on the equilibrium path). Suppose now that $p_i(w_i') < p_i(w_i)$. Then an incumbent firm facing a worker with type $\tau_i' \geq E[\tau_i \mid w_i']$ could deviate from the equilibrium and offer this worker w_i. The worker has now a strictly lower probability of being poached and receives a strictly lower wage if retained. The incumbent firm is strictly better off after this deviation. Thus, $p_i(w_i)$ must be non-decreasing in equilibrium. ■

Result 1 implies that higher types are more likely to be retained in any equilibrium. Because the first-best allocation implies that only the best types should be poached (if anyone should be poached at all), Result 1 implies that mixed-strategy equilibria are also talent-allocation inefficient. Furthermore, mixed-strategy equilibria also typically involve the inefficient poaching of mediocre workers. Thus, allowing for mixed-strategy equilibria does not restore efficiency, and our qualitative results are not affected by Assumption A2.

Now we fully characterize equilibria involving strictly mixed strategies in the case in which $1 \geq \theta\gamma$. For brevity, we only characterize the equilibrium poaching of workers from l firms, and thus to simplify notation we now drop the subscript $i = l$. From

$$\tau^{\#} = \begin{cases} \overline{\tau} & \text{if } \theta\gamma \leq 1 \\ \min\left\{(\theta - 1)\gamma\mu / (\theta\gamma - 1), \overline{\tau}\right\} & \text{if } \theta\gamma > 1 \end{cases} . \tag{2}$$

we have that $\tau^{\#} = \overline{\tau}$, thus poaching is always inefficient. Because equilibria in which workers play strictly-mixed strategies must involve some poaching, it follows trivially that such equilibria will also be inefficient. Furthermore, the source of inefficiency is the same as in the pure-strategy equilibria: there is too much poaching. Thus, the policy implications are also unchanged.

Although the equilibrium still involves excessive poaching, mixed strategies may improve allocational efficiency by allowing for the retention of some types in $[\gamma\mu, \tilde{\tau}]$ with some positive probability (but not with probability 1).

An equilibrium is characterized in the same way as in the pure-strategy case, except that we now need to describe the equilibrium behavior of a worker who faces two equivalent offers. Whenever an equilibrium with strictly-mixed strategies exists, there exists a function $p(w)$ that maps incumbent wage offers into probabilities of acceptance. Here we describe the equilibrium properties of this function.

Define $w(\tau)$ as the equilibrium wage offer that an incumbent makes to a worker of type τ and let $p(\tau) \equiv p(w(\tau))$. Result (1) shows that $p(w)$ is nondecreasing in w, which trivially implies that $p(\tau)$ is also non-decreasing in τ. Another equilibrium property of $p(\tau)$ is as follows:

Result 2 *Function $p(\tau)$ is continuous for all τ such that $p(\tau) > 0$.*

Proof. Consider τ' and let $\lim_{\varepsilon \to 0} p(\tau') - p(\tau' - \varepsilon) \equiv \delta$. For a deviation not to be profitable, we need

$$p(\tau')(\tau' - \varepsilon - \gamma\mu - w(\tau')) \leq p(\tau' - \varepsilon)(\tau' - \varepsilon - \gamma\mu - w(\tau' - \varepsilon)) \tag{3}$$

and

$$p(\tau')(\tau' - \gamma\mu - w(\tau')) \geq p(\tau' - \varepsilon)(\tau' - \gamma\mu - w(\tau' - \varepsilon)) \tag{4}$$

We take the limit as $\varepsilon \to 0$ and let $\tilde{w}(\tau') \equiv \lim_{\varepsilon \to 0} w(\tau' - \varepsilon)$. Then

$$p(\tau')(\tau' - \gamma\mu - w(\tau')) \leq (p(\tau') - \delta)(\tau' - \gamma\mu - \tilde{w}(\tau')) \tag{5}$$

and

$$p(\tau')(\tau' - \gamma\mu - w(\tau')) \geq (p(\tau') - \delta)(\tau' - \gamma\mu - \tilde{w}(\tau')), \tag{6}$$

which implies that $\delta = 0$, i.e., $p(\tau)$ must be continuous. ■

The next result follows directly from Results 1 and 2:

Corollary 1 *For $\tau \in [\tau', \overline{\tau}]$ such that $p(\tau') > 0$, we can find sets $A_1, A_2, ...$ such that $\bigcup_i A_i = [\tau', \overline{\tau}]$ and that, for each A_i, either $p(\tau)$ is constant for $\tau \in A_i$ or $p(\tau)$ is strictly increasing for $\tau \in A_i$.*

In other words, $p(\tau)$ is defined over regions of *pooling* (i.e., $p(\tau)$ is constant over an interval) and *fully-revealing separation* (i.e., $p(\tau)$ is strictly increasing over an interval, so that types in this interval are fully revealed in equilibrium).

Suppose that the interval $[a, b]$ is an equilibrium pooling region with $p(\tau) \in (0, 1)$ for $\tau \in [a, b]$, and assume that this interval is not contained in any other pooling interval. The equilibrium wage must be

$$w(\tau) = w^p = \theta\gamma \left(\int_a^b \frac{\tau f(\tau)}{F(b) - F(a)} d\tau - \mu \right) \text{ for } \tau \in [a, b]. \tag{7}$$

To find $p(\tau)$ for $\tau \in [a, b]$ notice there must exist at least one separating interval to the right or to the left of $[a, b]$. From continuity,

$$\lim_{\tau \to a} p(\tau) = \lim_{\tau \to b} p(\tau), \tag{8}$$

which implies that we can characterize $p(\tau)$ for $\tau \in [a, b]$ by the limit of $p(\tau)$ over any fully-revealing separation region in the neighborhood of $[a, b]$. This implies that it suffices to characterize $p(\tau)$ over separation regions.

Let $[c, d]$ denote a fully-revealing separation interval, so that type $\tau \in [c, d]$ is fully revealed in equilibrium. Due to competition among poachers, $w^p(w(\tau)) = \theta\gamma(\tau - \mu)$. In order to obtain separation, the probability schedule must be such that it prevents an incumbent employer with a worker of type τ from pretending that the worker is of type $\hat{\tau} \in [c, d]$ and $\hat{\tau} \neq \tau$. Thus, the following incentive compatibility constraint must hold for any such $\hat{\tau}$:

$$p(\tau)[\tau - \gamma\mu - \theta\gamma(\tau - \mu)] \geq p(\hat{\tau})([\tau - \gamma\mu - \theta\gamma(\hat{\tau} - \mu)]. \tag{9}$$

Define

$$U(\tau) = \max_{x \in [c,d]} p(x)[\tau - \gamma\mu - \theta\gamma(x - \mu)] + \gamma\mu. \tag{10}$$

By the envelope theorem we obtain:

$$\frac{\partial U(\tau)}{\partial \tau} = p(x^*) = p(\tau), \tag{11}$$

where the second equality follows from the IC condition in (9): If τ is fully revealed in equilibrium, then $x^* = \tau$.

Integrating (11) yields

$$U(\tau) = U(d) - \int_{\tau}^{d} p(x)dx. \tag{12}$$

For simplicity we assume that the function $p(\tau)$ is twice differentiable over the interval $[c, d]$. Then the next result allows us to solve for $p(\tau)$.

Result 3 *All incentive constraints are satisfied if and only if the following two sets of constraints hold:*

(i) Local incentive compatibility:

$$p'(\tau)\left[\tau - \gamma\mu - \theta\gamma(\tau - \mu)\right] - \theta\gamma p(\tau) = 0 \tag{13}$$

(ii) Monotonicity:

$$p'(\tau) \geq 0. \tag{14}$$

Proof. Assume first that all incentive compatibility constraints are satisfied, then it must be that the following first and second order conditions are satisfied at $x^* = \tau$

$$FOC: \; p'(x^*)\left[\tau - \gamma\mu - \theta\gamma(x^* - \mu)\right] - \theta\gamma p(x^*) = 0 \tag{15}$$

$$SOC: \; p''(x^*)\left[\tau - \gamma\mu - \theta\gamma(x^* - \mu)\right] - 2\theta\gamma p'(x^*) \leq 0 \tag{16}$$

Replacing x^* with τ and totally differentiating the local incentive compatibility constraint with respect to τ, we obtain:

$$p''(\tau)\left[\tau - \gamma\mu - \theta\gamma(\tau - \mu)\right] - 2\theta\gamma p'(\tau) + p'(\tau) = 0. \tag{17}$$

From the second order condition, this equation implies that $p'(\tau) \geq 0$.

Now, suppose that both the monotonicity and local incentive compatibility conditions hold. This must imply that all incentive compatibility constraints are satisfied:

$$p(\tau)\left[\tau - \gamma\mu - \theta\gamma(\tau - \mu)\right] \geq p(\hat{\tau})\left[\tau - \gamma\mu - \theta\gamma(\hat{\tau} - \mu)\right] \text{ for any } \tau \neq \hat{\tau}. \tag{18}$$

This equation can be rewritten as:

$$p(\tau)\left[\tau-\gamma\mu-\theta\gamma(\tau-\mu)\right]\geq p(\hat{\tau})\left[\hat{\tau}-\gamma\mu-\theta\gamma(\hat{\tau}-\mu)\right]-(\hat{\tau}-\tau)p(\hat{\tau})$$

or

$$(\hat{\tau}-\tau)p(\hat{\tau})\geq p(\tau)\left[\tau-\gamma\mu-\theta\gamma(\tau-\mu)\right]-p(\hat{\tau})\left[\hat{\tau}-\gamma\mu-\theta\gamma(\hat{\tau}-\mu)\right], \tag{19}$$

which implies

$$\begin{aligned}\int_{\tau}^{\hat{\tau}} p(\hat{\tau})dx \quad &\geq \quad \int_{\tau}^{d}\left\{p(x)+p'(x)\left[x-\gamma\mu-\theta\gamma(x-\mu)\right]-\theta\gamma p(x)\right\}dx \\ &\quad -\int_{\hat{\tau}}^{d}\left\{p(x)+p'(x)\left[x-\gamma\mu-\theta\gamma(x-\mu)\right]-\theta\gamma p(x)\right\}dx. \end{aligned}\tag{20}$$

If the local incentive compatibility constraint holds and $\hat{\tau}\geq\tau$, this condition becomes:

$$\int_{\tau}^{\hat{\tau}} p(\hat{\tau})dx\geq\int_{\tau}^{\hat{\tau}} p(x)dx, \tag{21}$$

which always holds for $p'(\tau)\geq 0$. If $\hat{\tau}<\tau$, the condition becomes:

$$\int_{\hat{\tau}}^{\tau} p(x)dx\geq\int_{\hat{\tau}}^{\tau} p(\hat{\tau})dx, \tag{22}$$

which always holds for $p'(\tau)\geq 0$. ■

This result allows us to characterize $p(\tau)$ by solving the differential equation in (13):

Corollary 2 *In any mixed-strategy equilibrium, the probability that type τ is retained is*

$$p(\tau)=K\left[(1-\theta\gamma)\tau+\gamma\mu(\theta-1)\right]^{\frac{\theta\gamma}{1-\theta\gamma}}, \tag{23}$$

where $K\geq 0$ is a constant.

The constant K is pinned down by the boundaries of $[c,d]$. The indeterminacy of K reflects the potential multiplicity of equilibria. Once a boundary condition is chosen, K is uniquely determined. For example, if $d=\overline{\tau}$ and type $\overline{\tau}$ is retained with probability 1, then

$$K=\left[(1-\theta\gamma)\overline{\tau}+\theta\mu(\theta-1)\right]^{\frac{\theta\gamma}{\theta\gamma-1}}. \tag{24}$$

2. Changing the Timing of the Offers

In the paper, the timing of the game is such that the uninformed party (the poacher) moves last. We now introduce the case in which the informed party (the incumbent) moves last.

We modify the original timing slightly by adding a date between Dates 2 and 3:

Date $2\frac{1}{2}$. Each firm i independently makes a counter offer w_i^c.

At Date 3, a worker from a firm i who holds an initial offer w_i, a poaching offer $w^p(w_i)$, and a counter offer w_i^c, accepts the poaching offer if and only if $w^p(w_i) > \max\{w_i, w_i^c\}$.

We now characterize the equilibrium under this modified timing. For the sake of brevity, we focus only on the equilibrium that displays the maximum amount of retention by the incumbent firm.* First, define the set $Y_i \equiv \{y \in Y_i : H_i(y) = 0\}$ where

$$H_i(y) \equiv y - \frac{\theta\gamma}{i}\left(\frac{\int_{\gamma\mu}^{y} \tau dF(\tau)}{F(y) - F(\gamma\mu)} - \mu\right) - \gamma\mu. \tag{25}$$

We then have the following result:

Result 4 *The (maximum-retention) equilibrium has the following properties:*

1. *There is a unique* $\tilde{\tau}_i' \in [\gamma\mu, \overline{\tau}]$ *such that all types* $\tau_i \geq \tilde{\tau}_i'$ *are retained. Threshold* $\tilde{\tau}_i'$ *is given by*

$$\tilde{\tau}_i' = \begin{cases} \text{the largest element in } \{\gamma\mu\} \cup Y_i & \text{if } H_i(\overline{\tau}) \geq 0 \\ \overline{\tau} & \text{if } H_i(\overline{\tau}) \leq 0 \end{cases}. \tag{26}$$

All retained workers are offered wage

$$w_i^{*\prime} = \max\left\{\theta\gamma\left(\frac{\int_{\gamma\mu}^{\tilde{\tau}_i'} \tau dF(\tau)}{F(\tilde{\tau}_i') - F(\gamma\mu)} - \mu\right), 0\right\}. \tag{27}$$

2. *All types* $\tau_i \in [0, \gamma\mu]$ *are fired in equilibrium.*

3. *All types* $\tau_i \in [\gamma\mu, \tilde{\tau}_i']$ *are poached in equilibrium.*

Proof. As before, we assume that E1 and E2 hold.

To find the equilibrium, we work backwards. At Date $2\frac{1}{2}$, the incumbent observes a poaching wage w_i^p. The incumbent pays the poaching wage and retains type τ if and only if $\tau - \frac{w_i^p}{i} \geq \gamma\mu$.

*In the original game, the most-efficient equilibrium is also the equilibrium that maximizes retention. By contrast, in the modified game, these two properties ("most-efficient" and "maximum-retention") may not lead to the same equilibrium. For comparing the two games, we choose the maximum retention criterion as the most natural. However, our conclusions are not sensitive to using alternative equilibrium-selection criteria.

At Date 2, a worker with a wage offer w_i receives a poaching offer equal to

$$\theta\gamma\left(\int_0^{\overline{\tau}} \tau dF\left(\tau \mid w_i, i\right) - \mu\right). \tag{28}$$

The beliefs represented by $F\left(\tau \mid w_i, i\right)$ must be Bayesian on the equilibrium path and consistent with E2.

At Date 1, the incumbent chooses w_i. We argue that an incumbent offers a unique wage $w_i = 0$ to any retained employee, i.e., an employee with talent $\tau_i \geq \gamma\mu$. The argument is similar to the one used to prove Lemma 1. Suppose that there are two types $\tau' > \tau''$ and that an incumbent i wants to retain both of them. Suppose the incumbent offers two different wages $w_i' > w_i''$ and suppose the poacher's offers are $w^p(w_i') > w^p(w_i'')$. Then, there is a profitable deviation for the incumbent, which is to offer w_i'' to both types. Now, suppose that $w_i > 0$. Then, the incumbent could deviate and offer $w_i' = 0$; Assumption E2 implies that $w^p(0) < w^p(w_i)$. Thus, $w_i = 0$. E1 implies that all $\tau < \gamma\mu$ receive negative offers. Maximum retention implies that the incumbent offers $w_i = 0$ to all $\tau_i \geq \gamma\mu$. This proves Part 2 of the result and that there is a unique $\tilde{\tau}_i' \in [\gamma\mu, \overline{\tau}]$ such that all types $\tau_i > \tilde{\tau}_i'$ are retained. Then, it follows that the equilibrium poaching wage is given by

$$w_i^p = \theta\gamma\left(\frac{\int_{\gamma\mu}^{\tilde{\tau}_i'} \tau dF\left(\tau\right)}{F\left(\tilde{\tau}_i'\right) - F\left(\gamma\mu\right)} - \mu\right), \tag{29}$$

and thus all retained workers are offered wage

$$w_i^{*\prime} = \max\left\{\theta\gamma\left(\frac{\int_{\gamma\mu}^{\tilde{\tau}_i'} \tau dF\left(\tau\right)}{F\left(\tilde{\tau}_i'\right) - F\left(\gamma\mu\right)} - \mu\right), 0\right\}, \tag{30}$$

because the incumbent only needs to offer $w_i^c = \max\{w_i^p, 0\}$. If w_i^p is strictly positive, then clearly all types $\tau_i \in (\gamma\mu, \tilde{\tau}_i')$ are poached in equilibrium. If $w_i^p \leq 0$, then no one is poached and thus $\tilde{\tau}_i' = \gamma\mu$. This proves Part 3.

To prove Part 1, suppose first that $H_i\left(\overline{\tau}\right) < 0$. Then, the incumbent does not wish to retain any type, implying that $\tilde{\tau}_i' = \overline{\tau}$.

Suppose now that $H_i\left(\overline{\tau}\right) \geq 0$. If $H_i\left(\tau_i\right) \geq 0$ for all τ_i, then the incumbent can retain any type for a given equilibrium w_i^p and still make a net profit. Thus, all types higher than $\gamma\mu$ are retained. Finally, if $H_i\left(\tau_i\right) < 0$ for some τ_i, then the set Y_i is non-empty and the equilibrium threshold must be in Y_i (which has at least two elements because $H_i\left(0\right) > 0$). Consider a candidate equilibrium threshold $\tau_i^* \in Y_i$, with respective equilibrium poaching wage w_i^{p*}, and assume that τ_i^* is not the largest element of Y_i. Then, a single poacher may deviate and offer an alternative poaching wage equal to

$$w_i^{p\prime} = \tilde{\tau}_i' - \alpha - \gamma\mu, \tag{31}$$

where $\tilde{\tau}_i'$ is the largest element in Y_i and $\alpha > 0$ is sufficiently small so that $w_i^{p*} < w_i^{p\prime}$. This poacher would be successful at poaching all types $[\gamma\mu, \tilde{\tau}_i' - \alpha)$ at a wage that is strictly lower than the one implied by the zero net profit condition. Thus, this deviation is profitable. Thus, the equilibrium threshold must be $\tilde{\tau}_i'$, i.e., the largest element of Y_i. ■

The equilibrium outcome is qualitatively similar to the outcome in Proposition 4: All types above a threshold are retained, and only mediocre types are poached. Thus, our main result that asymmetric information creates inefficiencies in talent allocation does not depend on whether the informed party moves last or not. In particular, we note that not only inefficient retention is possible, but also that inefficient poaching will often occur because at least a subset of types in $[\gamma\mu, \tilde{\tau}_i']$ should be retained in the first-best allocation.

An important property of this equilibrium is as follows:

Result 5 *In the modified game in which the incumbent moves last, fewer types are poached in equilibrium:*

$$\tilde{\tau}_i' \leq \tilde{\tau}_i. \tag{32}$$

Proof. Threshold $\tilde{\tau}_i$ is defined by the lowest value that solves

$$\tilde{\tau}_i - \gamma\mu = \frac{\theta\gamma}{i}\left(\frac{\int_{\tilde{\tau}_i}^{\overline{\tau}} \tau dF(\tau)}{1 - F(\tilde{\tau}_i)} - \mu\right). \tag{33}$$

(We assume an interior solution for simplicity; if the solution is not interior, then there is no retention, and the result is trivially proven).

Threshold $\tilde{\tau}_i'$ is defined by the largest value that solves

$$\tilde{\tau}_i' - \gamma\mu = \frac{\theta\gamma}{i}\left(\frac{\int_{\gamma\mu}^{\tilde{\tau}_i'} \tau dF(\tau)}{F(\tilde{\tau}_i') - F(\gamma\mu)} - \mu\right). \tag{34}$$

Suppose that $\tilde{\tau}_i' > \tilde{\tau}_i$. Then, it must be that

$$\frac{\int_{\gamma\mu}^{\tilde{\tau}_i'} \tau dF(\tau)}{F(\tilde{\tau}_i') - F(\gamma\mu)} > \frac{\int_{\tilde{\tau}_i}^{\overline{\tau}} \tau dF(\tau)}{1 - F(\tilde{\tau}_i)}, \tag{35}$$

which cannot be true. ■

This result demonstrates that when the incumbent has the option to make the last offer, it is able to retain the worker more often. This result is unsurprising because this modified timing gives more market power to the incumbent. One interpretation for this timing of offers is that if the worker accepts the incumbent's offer at date $2\frac{1}{2}$, this offer becomes binding and the worker can no longer accept a poaching offer.

Because of (32), the modified game is more likely to display inefficient retention than the original game. The modified game is less likely to display inefficient poaching than the original game for the same reason. Thus, by giving the incumbent the option to make a final binding offer, poaching inefficiencies can be reduced and sometimes eliminated.

3. Deferred compensation

The solution to the planner's problem reveals that the negative assortative matching result is a consequence of information asymmetries alone and not of any artificial restriction on the space of contracts. It is nevertheless instructive to consider the case in which the incumbent may use deferred compensation as a means to reduce mobility.

Proposition 2 immediately implies that, from a social welfare perspective, such bonding contracts may either improve or worsen efficiency. However, even when it is optimal to ban these contracts, incumbent firms may still choose to write such contracts. Here we show that such contracts, even when feasible, may not be voluntarily adopted by firms.

Consider the following contract: Before the incumbent firm learns its worker's type (at $t = 0$), the firm commits to a fixed wage $\overline{w}_i$ to be paid at the end of the game, but only if the worker remains with the firm or if the worker is fired. To retain types $\tau_i \geq \gamma\mu$, the lowest wage that must be offered is $\overline{w}_i = w^p(\overline{w}_i) = \theta\gamma\left(\int_{\gamma\mu}^{\overline{\tau}} \frac{\tau f(\tau)}{1-F(\gamma\mu)} d\tau - \mu\right)$. Under commitment to $\overline{w}_i$, expected profit (at $t = 1$) to the incumbent is thus

$$E[\pi_{ic}] = F(\gamma\mu)\gamma\mu + [1 - F(\gamma\mu)] \int_{\gamma\mu}^{\overline{\tau}} \frac{\tau f(\tau)}{1 - F(\gamma\mu)} d\tau - \overline{w}_i. \tag{36}$$

Without commitment, we know that the equilibrium implies that the incumbent chooses some $\tilde{\tau}_i \geq \gamma\mu$, and thus its expected profit at $t = 1$ is

$$E[\pi_{inc}] = F(\tilde{\tau}_i)\gamma\mu + [1 - F(\tilde{\tau}_i)]\left[\int_{\tilde{\tau}_i}^{\overline{\tau}} \frac{\tau f(\tau)}{1 - F(\tilde{\tau}_i)} d\tau - \theta\gamma\left(\int_{\tilde{\tau}_i}^{\overline{\tau}} \frac{\tau f(\tau)}{1 - F(\tilde{\tau}_i)} d\tau - \mu\right)\right]. \tag{37}$$

It can be shown, through simple examples, that $E[\pi_{inc}] \lesseqgtr E[\pi_{ic}]$ depending on the parameters. The intuition for this result is that deferred compensation schemes (such as restricted shares or vesting of stock options) are costly to the firm because some workers who are fired are still paid $\overline{w}_i$, which may leave rents to dismissed workers (for example, if wages at $t = 0$ cannot be negative). Thus, the expected excess cost of such a scheme is $F(\gamma\mu)\overline{w}_i$. Without such a scheme, the overall surplus may be higher or lower, but the profit could still be larger *even when the surplus is lower.* Hence, deferred compensation contracts may not be chosen by firms even when they are feasible.

4. An Infinite-Horizon Model

We now extend the analysis to more than two periods, which allows for a more natural interpretation of the model.

The economy is populated with many infinitely-lived firms. Again, firms can be of one of two types, L or H, and these represent both the type and the mass of firms of each type. To simplify the notation and the exposition, we denote a representative firm of each type by $i \in \{l, h\}$, which also denotes the profitability parameter, i.e., $h = \theta$ and $l = 1$, where $\theta > 1$.

Workers live for two periods: young age and old age. Firms and workers are risk-neutral and share a common discount factor $\delta \in [0, 1)$. At each period t ($t = 0, 1, 2, ...$), a mass M of young workers enter the labor market. Young workers are in excess supply: $M > H + L$. The outside option of an unemployed agent (young or old) is normalized to zero.

As before, we assume that bonding arrangements (fines, transfer fees, non-compete clauses, etc.) are not feasible. To focus on the role of asymmetric learning, we assume that workers are not protected by limited liability, because limited liability would generate inefficiencies in the infinite-horizon model even when learning is symmetric.

At the beginning of a period, a firm can be in one of the following states:

(i) The firm has a vacant position, because its worker retired at the end of the previous period (that is, the worker was old).

(ii) The firm does not have a vacant position because its worker was young in the previous period.

Both types of firms may have incumbent workers and may also become poachers.

In each period t, the timing of actions is as follows:

Timing.

Date 1. Each type-i firm with an incumbent (old) worker who is known to be of type τ independently chooses wage $w_i(\tau) \in \mathbb{R}$.

Date 2. After observing all wage offers w_i, all firms that have a vacant position simultaneously make offers according to the function $w^p(w_i)$.

Date 3. A worker who holds an offer w accepts all poaching offers such that $w^p(w) > w$ and rejects all poaching offers such that $w^p(w) \leq w$ (as described in Assumption A2).

Date 4. All firms that do not have a worker at this date randomly select one young agent from the outside pool and offer the wage $w_i^y \in \mathbb{R}$ (i.e., w_i^y could be positive or negative), for $i \in \{l, h\}$.[†]

Date 5. Payoffs are realized and a young worker's type is revealed to the worker's employer.

In sum, a firm that hires a young worker (whose average type is μ) learns about the type of its worker only at the end of the period (at Date 5) and thus begins the subsequent period with an old worker whose type is known. For simplicity, we also assume that firm performance is not observed by outsiders, or (equivalently) that observed firm performance is noisy, and thus not informative about a young worker's type.

A type-h firm can attempt to poach a worker from a type-l firm or from another type-h firm. In general, we also allow type-l firms to make poaching offers. However, for simplicity, we (implicitly) restrict our analysis to a set of parameters for which, in equilibrium, workers would strictly prefer poaching offers from type-h firms. Thus, without loss of generality, we assume that type-l firms cannot poach workers.

4.1. Benchmark: Symmetric Learning

Under symmetric learning, all firms have the same information about an old worker's type, i.e., they learn the employed young workers' types at Date 5 of each period. As the equilibrium will be time-invariant, for simplicity we ignore time subscripts. At Date 1 of each period, a type-i firm with an incumbent worker who is of a known type τ offers the wage:

$$w_i^S = \begin{cases} \text{any } w < 0 & \tau \leq \underline{\tau}_i \\ 0 & \tau \in [\underline{\tau}_i, \hat{\tau}_i] \\ w^{pS}(\tau) & \tau \in \left[\hat{\tau}_i, \tau_i^{\#}\right] \\ \text{any } w < w^{pS}(\tau) & \tau \in \left[\tau_i^{\#}, \overline{\tau}\right) \end{cases}, \tag{38}$$

where $\underline{\tau}_i$, $\hat{\tau}_i$, $\tau_i^{\#}$ and function $w^{pS}(\tau)$ are to be determined in equilibrium.

Because poachers compete à la Bertrand, their equilibrium value function, $V_h^{pS}(\tau)$, when poaching a worker of type τ should be equal to the value they derive from hiring a young worker, V_h^{yS}:

$$V_h^{pS}(\tau) - V_h^{yS} = 0, \tag{39}$$

[†]It is always in a firm's interest to hire a young – as opposed to an old – employee from the pool of the unemployed.

where

$$V_h^{pS}(\tau) = \theta\gamma\tau - w^{pS}(\tau) + \delta \max\left\{V_h^{yS}, V_h^{pS}(\tau)\right\}, \tag{40}$$

$$V_h^{yS} = \theta\gamma\mu - w^{yS} + \delta V_h^{oS}, \tag{41}$$

and

$$V_h^{oS} = F(\underline{\tau}_h)\max\left\{V_h^{yS}, V_h^{pS}(\tau)\right\} + \theta\int_{\underline{\tau}_h}^{\overline{\tau}} \tau f(\tau)d\tau - \int_{\widehat{\tau}_h}^{\overline{\tau}} w^{pS}(\tau)f(\tau)d\tau + \delta(1-F(\underline{\tau}_h))V_h^{yS}. \tag{42}$$

By replacing (40) and (41) into (39), we obtain the following expression for the poaching wage (recall that this is only defined for non-negative wages):

$$w^{pS}(\tau) = \theta\gamma(\tau-\mu) + w^{yS} - \delta(V_h^{oS} - V_h^{yS}). \tag{43}$$

The threshold $\hat{\tau}_i$ corresponds to the level of talent above which a poacher offers a positive wage to a worker of type $\tau > \hat{\tau}_i$. Because information is symmetric, the poaching wage depends only on a worker's talent, therefore we set $\hat{\tau}_l = \hat{\tau}_h = \hat{\tau}$, and thus threshold $\hat{\tau}$ is given by $w^{pS}(\hat{\tau}) = 0$.

Using (41) in (42), we obtain:

$$\begin{aligned} V_h^{oS} &= F(\underline{\tau}_h)\left[\theta\gamma\mu - w^{yS} + \delta V_h^{oS}\right] + \theta\int_{\underline{\tau}_h}^{\overline{\tau}} \tau f(\tau)d\tau \\ &\quad - \int_{\widehat{\tau}_h}^{\overline{\tau}} w^{pS}(\tau)f(\tau)d\tau + \delta(1-F(\underline{\tau}_h))V_h^{yS}. \end{aligned} \tag{44}$$

Subtracting V_h^{yS} from both sides yields

$$\begin{aligned} V_h^{oS} - V_h^{yS} &= -\left[1-F(\underline{\tau}_h)\right]\left[\theta\gamma\mu - w^{yS} + \delta V_h^{oS}\right] + \theta\int_{\underline{\tau}_h}^{\overline{\tau}} \tau f(\tau)d\tau \\ &\quad - \int_{\widehat{\tau}_h}^{\overline{\tau}} w^{pS}(\tau)f(\tau)d\tau + \delta(1-F(\underline{\tau}_h))V_h^{yS}, \end{aligned} \tag{45}$$

or

$$V_h^{oS} - V_h^{yS} = \frac{\int_{\underline{\tau}_h}^{\overline{\tau}}(\theta\tau - \theta\gamma\mu)f(\tau)d\tau - \int_{\hat{\tau}}^{\overline{\tau}} w^{pS}(\tau)f(\tau)d\tau + (1-F(\underline{\tau}_h))w^{yS}}{1+\delta(1-F(\underline{\tau}_h))}. \tag{46}$$

The first-period wage of a young worker is given by the first-period participation constraint:[‡]

$$w^{yS} = -\delta\int_{\hat{\tau}}^{\overline{\tau}} w^{pS}(\tau)f(\tau)d\tau. \tag{47}$$

[‡]Note that the wage of a young worker is independent of the type of the firm.

Therefore, we can replace $\int_{\hat{\tau}}^{\overline{\tau}} w^{pS}(\tau)f(\tau)d\tau$ by $-w^{yS}/\delta$ in (46) to obtain:

$$V_h^{oS} - V_h^{yS} = \frac{\int_{\underline{\tau}_h}^{\overline{\tau}} (\theta\tau - \theta\gamma\mu)f(\tau)d\tau}{1+\delta(1-F(\underline{\tau}_h))} + \frac{w^{yS}}{\delta}. \tag{48}$$

Now, plug (48) into (43) to find the poaching wage offered to a worker with talent τ (this function is defined only for values of τ such that $w^{pS}(\tau) \geq 0$):

$$w^{pS}(\tau) = \theta\gamma(\tau - \mu) - \frac{\delta}{1+\delta(1-F(\underline{\tau}_h))}\int_{\underline{\tau}_h}^{\overline{\tau}} (\theta\tau - \theta\gamma\mu)f(\tau)d\tau. \tag{49}$$

In the infinite-horizon model, for a given τ, the offer made by a poacher is lower than that in the two-period model. In the infinite-horizon setting, hiring a young worker has an option value: At Date 5 of the first period of employment, the firm learns the worker's type and thus has the option to retain this worker for the subsequent period. The value of this option is given by the second term on the right-hand side of (49). Thus, poaching an old worker comes at an opportunity cost, which is the value of this option. The existence of this option is the main qualitative difference between the static model and the infinite-horizon model.

The threshold $\hat{\tau}_i$ corresponds to the level of talent above which a poacher offers a positive wage to a worker of type $\tau > \hat{\tau}_i$. Because information is symmetric, the poaching wage depends only on a worker's talent; therefore, we set $\hat{\tau}_l = \hat{\tau}_h = \hat{\tau}$, and thus the threshold $\hat{\tau}$ is given by $w^{pS}(\hat{\tau}) = 0$.

The first-period wage w^{yS} of a young worker is given by

$$w^{yS} = -\delta \int_{\hat{\tau}}^{\overline{\tau}} w^{pS}(\tau)f(\tau)d\tau. \tag{50}$$

Note that this wage is always negative and equal to the discounted expected wage received by this worker in the second period. In other words, young workers have zero expected surplus. This result is a consequence of our assumptions that the worker's outside option is zero and that there is no limited liability. We know from Terviö (2009) that, in a dynamic model with symmetric learning, limited liability creates inefficiencies: There is excessive retention of mediocre types. Because we want to isolate the effect of asymmetric learning on welfare, we choose not to impose limited liability, which also implies that, unlike Terviö (2009), the first-best allocation is obtained in our benchmark model with symmetric learning.

Threshold $\underline{\tau}_i$ from (38) is determined by

$$V_i^{oS}(\underline{\tau}_i) - V_i^{yS} = 0, \tag{51}$$

where $V_i^{oS}(\tau)$ is the value function a type-i firm from retaining an incumbent (old) worker with talent τ, and V_i^{yS} is the value from hiring a young worker. For a type-h firm, this is given by

$$V_h^{oS}(\underline{\tau}_h) - V_h^{yS} = 0, \tag{52}$$

where

$$V_h^{oS}(\underline{\tau}_h) = \theta\underline{\tau}_h + \delta V_h^{yS}, \tag{53}$$

and V_h^{yS} is given by equation (41) (Recall that in equilibrium $V_h^{yS} = V^{pS}(\tau)$ for any $\tau \geq \hat{\tau}$). We can rewrite (52) as

$$\begin{aligned} &\theta\underline{\tau}_h - \theta\gamma\mu + w^{yS} - \delta(V_h^{oS} - V_h^{yS}) = 0 \\ \iff &\theta\underline{\tau}_h - \theta\gamma\mu - \frac{\delta}{1+\delta(1-F(\underline{\tau}_h))}\int_{\underline{\tau}_h}^{\overline{\tau}}(\theta\tau - \theta\gamma\mu)f(\tau)d\tau = 0 \\ \iff &\underline{\tau}_h - \gamma\mu - \delta\int_{\underline{\tau}_h}^{\overline{\tau}}(\tau - \underline{\tau}_h)f(\tau)\,d\tau = 0. \end{aligned} \tag{54}$$

The equilibrium threshold $\underline{\tau}_h$ is given by the unique solution to (54) (note that the left-hand side of (54) is increasing in $\underline{\tau}_h$ and is negative for $\underline{\tau}_h = 0$ and positive for $\underline{\tau}_h = \overline{\tau}$). Then, we have a closed form solution for the poaching wage in (49). By setting $w^{pS}(\widehat{\tau}) = 0$ in (49), we then obtain a unique equilibrium value for $\hat{\tau}$.

So, the threshold is given by:

$$\underline{\tau}_h = \gamma\mu + \delta\int_{\underline{\tau}_h}^{\overline{\tau}}(\tau - \underline{\tau}_h)f(\tau)d\tau. \tag{55}$$

The decision to retain a worker is given by the following trade-off. The left-hand side of (55) is the immediate gain from retaining an old worker of type $\underline{\tau}_h$; the right-hand side is the benefit from hiring a young worker from the outside pool. This benefit has two components. First, a young worker from the outside pool produces (in expectation) $\gamma\mu$ during the first year of employment. Second, hiring a young worker again gives the firm the option to retain this worker in the subsequent period. The value of this option is given by the second term on the right-hand side of (55).

We now need to find threshold $\underline{\tau}_l$. An l-firm is willing to retain a worker of type $\underline{\tau}_l$ for a wage of zero if the following condition holds:

$$V_l^{oS}(\underline{\tau}_l) - V_l^{yS} = 0,$$

where

$$V_l^{oS}(\underline{\tau}_l) = \underline{\tau}_l + \delta V_l^{yS}, \tag{56}$$

$$V_l^{yS} = \gamma\mu - w^{yS} + \delta V_l^{oS}, \tag{57}$$

and

$$\begin{aligned} V_l^{oS} &= (F(\underline{\tau}_l) + 1 - F(\tau_l^{\#}))V_l^{yS} + \int_{\underline{\tau}_l}^{\tau_l^{\#}} \tau f(\tau) d\tau \\ &- \int_{\widehat{\tau}_l}^{\tau_l^{\#}} w^{pS}(\tau) f(\tau) d\tau + \delta(F(\tau_l^{\#}) - F(\underline{\tau}_l))V_l^{yS}. \end{aligned} \tag{58}$$

We use (56), (57), and (58) to obtain:

$$V_l^{oS}(\underline{\tau}_l) - V_l^{yS} = 0 \quad \Leftrightarrow \underline{\tau}_l - \gamma\mu - \frac{\delta\left(\int_{\underline{\tau}_l}^{\tau_l^{\#}} (\tau - \gamma\mu) f(\tau) d\tau + \int_{\tau_l^{\#}}^{\overline{\tau}} w^p(\tau) f(\tau) d\tau\right)}{1 + \delta(F(\tau_l^{\#}) - F(\underline{\tau}_l))} = 0 \tag{59}$$

$$\Leftrightarrow \underline{\tau}_l - \gamma\mu - \delta \int_{\underline{\tau}_l}^{\tau_l^{\#}} (\tau - \underline{\tau}_l) f(\tau) d\tau - \delta \int_{\tau_l^{\#}}^{\overline{\tau}} w^p(\tau) f(\tau) d\tau = 0, \tag{60}$$

which again determines a unique $\underline{\tau}_l$ for a given $\tau_l^{\#}$.

For a type-l firm the retention threshold $\underline{\tau}_l$ is thus:

$$\underline{\tau}_l = \gamma\mu + \delta \int_{\underline{\tau}_l}^{\tau_l^{\#}} (\tau - \underline{\tau}_l) f(\tau) d\tau + \delta \int_{\tau_l^{\#}}^{\overline{\tau}} w^{pS}(\tau) f(\tau) d\tau. \tag{61}$$

The first two terms on the right-hand side of (61) are analogous to those in (55). The key difference between these two conditions is the last term on the right-hand side of (61), which represents the present value of the wages paid to those workers who are poached in equilibrium in the second year of employment. A firm of type l is able to capture such surplus by offering a negative wage to young workers. Thus, these firms are compensated for being talent discoverers; even if their best workers leave to work for other firms, type-l firms capture all the surplus generated by an efficient allocation of talent.

Now, we only need to find $\tau_l^{\#}$. Poaching exists only if the incremental surplus to the poacher is larger than the net loss to the incumbent firm:

$$V_l^{oS}(\tau_l) - V_l^{yS} \leq V_h^{pS}(\tau_l) - V_h^{yS}. \tag{62}$$

To see that this must hold in any equilibrium with poaching, note that if it did not hold, the incumbent could offer a slightly larger wage and profitably prevent poaching. Thus, if an interior $\tau_l^{\#}$ exists, it is determined by one of the solutions to (62) with equality, which yields:

$$\tau_l^{\#} - \gamma\mu - \frac{\delta \int_{\underline{\tau}_l}^{\tau_l^{\#}} (\tau - \gamma\mu) f(\tau) d\tau + \delta \int_{\tau_l^{\#}}^{\overline{\tau}} w^p(\tau) f(\tau) d\tau}{1 + \delta(F(\tau_l^{\#}) - F(\underline{\tau}_l))} = \theta\gamma(\tau_l^{\#} - \mu) - \frac{\delta \int_{\underline{\tau}_h}^{\overline{\tau}} (\theta\tau - \theta\gamma\mu) f(\tau) d\tau}{1 + \delta(1 - F(\underline{\tau}_h))}. \tag{63}$$

(If there is no interior solution, the equilibrium is such that no one is poached). If there is more than one solution, only one of such solutions is an equilibrium. To see this, note that if τ_l is poached in any equilibrium, then $\tau_l' > \tau_l$ will also be poached because $\tau_l - \gamma\mu - \theta\gamma(\tau_l - \mu)$ is strictly decreasing in τ_l (note that the value of future options do not change with τ_l). Thus, there is a unique set of values $\left(\underline{\tau}_l, \underline{\tau}_h, \tau_l^{\#}, \widehat{\tau}, w^{yS}\right)$ and function $w^{pS}(\tau)$ that characterize the equilibrium.

We now discuss two important properties of the equilibrium. First, we have the following result:

Result 6 $\underline{\tau}_l \geq \underline{\tau}_h$.

Proof. Begin by rewriting (60) as

$$\underline{\tau}_l - \gamma\mu = \delta \int_{\underline{\tau}_l}^{\overline{\tau}} (\tau - \underline{\tau}_l) f(\tau) d\tau + \delta \int_{\tau_l^{\#}}^{\overline{\tau}} (w^p(\tau) - \tau + \underline{\tau}_l) f(\tau) d\tau. \tag{64}$$

The left-hand side of equation (64) increases with $\underline{\tau}_l$ and the right-hand side (RHS) decreases with $\underline{\tau}_l$. If $\tau_l^{\#} = \overline{\tau}$, then the conditions defined by equations (64) and (54) are the same and $\underline{\tau}_l = \underline{\tau}_h$. If $\tau_l^{\#} < \overline{\tau}$, then $\delta \int_{\tau_l^{\#}}^{\overline{\tau}} (w^p(\tau) - \tau + \underline{\tau}_l) f(\tau) d\tau > 0$, which increases the RHS and thus increases the value for $\underline{\tau}_l$. ■

This result indicates that l-firms are more likely to fire workers with low talent than are h-firms. The intuition is as follows: It is more efficient for l-firms to act as talent discoverers than as producers because l-firms are as efficient as h-firms in discovering talent, but less efficient at producing output. Thus, l-firms have a comparative (but not absolute) advantage at discovering talent and should thus do more of it in an efficient allocation.

We also have the following result:

Result 7 *The unique equilibrium under symmetric learning is efficient (in the Kaldor-Hicks sense).*

To prove this result formally, we proceed as follows. We first state the necessary and sufficient conditions for an allocation, here fully characterized by thresholds $\left(\tau_l^{\#*}, \underline{\tau}_l^*, \underline{\tau}_h^*\right)$, to be (Kaldor-Hicks) efficient (i.e., to maximize a social welfare function with equal weights to all players). We then show that we can construct a set of prices (wages) that sustains such an allocation as a decentralized equilibrium of our game. Thus, an efficient allocation is also a decentralized equilibrium. Because the decentralized equilibrium is unique, it is thus always efficient.

Proof. For simplicity, without loss of generality we consider only symmetric allocations in which all firms and workers of the same type and in identical situations are assigned the same surplus by a hypothetical social planner. Under this assumption, to derive the efficiency conditions we can work with an alternative interpretation of the model in which there is only one firm of each type.

Consider an allocation associated with the thresholds $\left(\tau_l^{\#*}, \underline{\tau}_l^*, \underline{\tau}_h^*\right)$. Let $S^*(\tau_l, \tau_h)$ denote the total surplus generated by this allocation, conditional on knowing the incumbent workers' types (τ_l, τ_h) (if one or both firms do not have incumbent workers, define the surplus accordingly as being conditional only on the type of the existing incumbent worker, if any). This allocation is efficient if and only if, for any other allocation with conditional surplus $S'(\tau_l, \tau_h)$,

$$S^*(\tau_l, \tau_h) \geq S'(\tau_l, \tau_h) \text{ for all } (\tau_l, \tau_h). \tag{65}$$

We can focus on conditional surplus because, under the current interpretation, there are only two firms and at most two incumbent workers.

To maximize (conditional) surplus, we list three necessary conditions:[§]

(1) For any given τ_l, firm l retains this type instead of hiring a young worker if and only if:

$$U_l^o(\tau_l) + u(\tau_l) + U_h^{ol} + u(\tau_h, \tau_l) \geq U_l^y + u_f(\tau_l) + u^{yl} + U_h^{yl} + u_f(\tau_h, \tau_l), \tag{66}$$

where $U_i^o(\tau_i)$ is the expected payoff to i of retaining τ_i under the allocation, $u(\tau_i)$ is the expected payoff to worker τ_i of being retained by i, U_h^{ol} is the expected payoff to h of l retaining τ_l, $u(\tau_h, \tau_j)$ is the expected payoff to worker τ_h, who currently works for firm h, if worker τ_l is retained by l (if h has no incumbent worker, we set this value to zero), U_i^y is the expected payoff to i of hiring a young worker, $u_f(\tau_l)$ is the expected payoff to a worker of type τ_l of being fired by l, u^{yi} is the expected payoff to a young worker of being hired by i, $U_h^{yl}$is the expected payoff to h of l hiring a young worker, and $u_f(\tau_h, \tau_l)$ is the expected payoff to worker τ_h if worker τ_l is fired by firm l (if firm h has no incumbent worker, we set this value to zero).

(2) If firm h has a vacancy, h poaches a worker of type τ_l instead of hiring a young worker if and only if:

$$U_h^p(\tau_l) + u^h(\tau_l) + U_l^y + u^{yl} \geq U_h^y + u^{yh} + \max\left\{U_l^o(\tau_l) + u(\tau_l), U_l^y + u^{yl}\right\}. \tag{67}$$

[§]In what follows, for simplicity we assume that all workers who remain unemployed are assigned zero net surplus by the social planner. This is without loss of generality. In addition, in line with the previous assumption that only h firms can be poachers, we focus on the cases where there are no job transitions from h to l.

where $U_h^p(\tau_l)$ is the expected payoff to h of poaching τ_l and $u^h(\tau_l)$ is the expected payoff to worker τ_l of being hired by h.

(3) For any given τ_h and τ_l, firm h retains this type if and only if:

$$\begin{aligned} U_h^o(\tau_h) + u(\tau_h) + \max\left\{U_l^o(\tau_l) + u(\tau_l), U_l^y + u^{yl}\right\} \geq \\ \max\left\{U_h^y + u^{yh} + \max\left\{U_l^o(\tau_l) + u(\tau_l), U_l^y + u^{yl}\right\}, U_h^p(\tau_l) + u^h(\tau_l) + U_l^y + u^{yl}\right\}. \end{aligned} \tag{68}$$

Now, consider the efficient allocation, which is determined by the thresholds $\left(\tau_l^{\#*}, \underline{\tau}_l^*, \underline{\tau}_h^*\right)$. Note first that these thresholds fully determine the following wages:

$$w^{p*}(\tau) = \theta\gamma\tau - \theta\gamma\mu - \frac{\delta}{1+\delta(1-F(\underline{\tau}_h^*))}\int_{\underline{\tau}_h^*}^{\overline{\tau}}(\theta\tau - \theta\gamma\mu)f(\tau)d\tau, \tag{69}$$

$$w^{y^*} = -\delta\int_{\hat{\tau}^*}^{\overline{\tau}} w^{pS}(\tau)f(\tau)d\tau, \tag{70}$$

where $\hat{\tau}^*$ is the threshold for which $w^{p*}(\hat{\tau}^*) = 0$. Given these wages, then we can easily verify that we can uniquely define $V_h^{p*}(\tau), V_h^{o*}(\tau), V_h^{y*}, V_h^{o*}, V_l^{o*}(\tau), V_l^{y*}$, and V_l^{o*} as the value functions as before, but taking the thresholds $\left(\tau_l^{\#*}, \underline{\tau}_l^*, \underline{\tau}_h^*\right)$ as given.

We now need to show that such wages can sustain a decentralized equilibrium such that Conditions (1)-(3) hold. Start with (66). First, if $u^{yl} \neq 0$, then use (positive or negative) lump-sum transfers from the worker to firm l to create a new allocation on the right-hand side of (66), without changing its total surplus, so that U_l^y under this new allocation is equal to the old U_l^y plus the old u^{yl}, and thus the new u^{yl} becomes zero:

$$U_l^o(\tau_l) + u(\tau_l) + U_h^{ol} + u(\tau_h, \tau_l) \geq U_l^y + u_f(\tau_l) + U_h^{yl} + u_f(\tau_h, \tau_l). \tag{71}$$

Second, consider U_h^{ol}. Suppose that h has a vacancy. If $U_h^{ol} \neq V_h^{y*}$, make transfers to or from all the other players until $U_h^{ol} = V_h^{y*}$ and the surplus on left-hand side is unchanged.¶ Make similar transfers in the analogous case in which h has a worker of type τ_h until $U_h^{ol} = \max\{V_h^{y*}, V_h^{o*}(\tau_h)\}$. Make similar transfers on the right-hand side until $U_h^{yl} = V_h^{y*}$ or $U_h^{yl} = \max\{V_h^{y*}, V_h^{o*}(\tau_h)\}$, depending on which case is relevant. Then, we can rewrite the condition above as

$$U_l^o(\tau_l) + u(\tau_l) + u(\tau_h, \tau_l) \geq U_l^y + u_f(\tau_l) + u_f(\tau_h, \tau_l). \tag{72}$$

Third, consider $u(\tau_h, \tau_l)$. This term is zero if h has a vacancy. If instead h has an incumbent who is retained (i.e., if if $\tau_h \geq \widetilde{\tau}_h^*$), then $u(\tau_h, \tau_l) = u(\tau_h)$. Suppose in this case that $\tau_h \leq \hat{\tau}^*$.

¶Notice that such transfers can always be made because the initial allocation is assumed to be efficient and thus has the maximum possible conditional surplus. If, counterfactually, V_h^{y*} was higher than the maximum surplus, an allocation that delivered V_h^{y*} (which is possible by construction) would be superior to the efficient allocation, which is a contradiction.

If $u(\tau_h) \neq 0$, make transfers so that $u(\tau_h) = 0$. If instead $\tau_h > \hat{\tau}^*$, if $u(\tau_h) \neq V_h^{p*}(\tau_h) - V_h^{y*}$, make transfers until $u(\tau_h) = V_h^{p*}(\tau_h) - V_h^{y*}$. Similarly, we have that $u_f(\tau_h, \tau_l) = u(\tau_h)$ if $\tau_h \geq \widetilde{\tau}_h^*$ and $u_f(\tau_h, \tau_l) = 0$ otherwise. Make transfers on the right-hand side so that $u(\tau_h) = 0$ or $u(\tau_h) = V_h^{p*}(\tau_h) - V_h^{y*}$, depending on which case is relevant. Then, we can rewrite the condition above as

$$U_l^o(\tau_l) + u(\tau_l) \geq U_l^y + u_f(\tau_l). \tag{73}$$

Finally, suppose first that $\tau_l \leq \hat{\tau}^*$. If $u(\tau_l) \neq 0$, make transfers to or from l so that $u^l(\tau_l) = 0$. Suppose now that $\tau_l > \hat{\tau}^*$. If $u(\tau_l) \neq V_h^{p*}(\tau_l) - V_h^{y*}$, make transfers to or from l so that $u(\tau_l) = V_h^{p*}(\tau_l) - V_h^{y*}$. Similarly, make transfers on the right-hand side so that $u_f(\tau_l) = 0$ or $u_f(\tau_l) = V_h^{p*}(\tau_l) - V_h^{y*}$, depending on which case is relevant. Then, we can rewrite the condition above as

$$U_l^o(\tau_l) \geq U_l^y, \tag{74}$$

which by construction is equivalent to $V_l^{o*}(\tau_l) \geq V_l^{y*}$. But this is also a necessary condition for the retention of type τ_l in a competitive equilibrium given thresholds $\left(\tau_l^{\#*}, \underline{\tau}_l^*, \underline{\tau}_h^*\right)$. Thus, condition (66) is compatible with a decentralized equilibrium with thresholds $\left(\tau_l^{\#*}, \underline{\tau}_l^*, \underline{\tau}_h^*\right)$.

It is possible to replicate this argument for the other two conditions (i.e., (67) and (68)), and similarly show that none of these conditions impose restrictions on the equilibrium. The steps are tedious but simple; we omit them here for brevity.

We then conclude that, for any given efficient allocation $\left(\tau_l^{\#*}, \underline{\tau}_l^*, \underline{\tau}_h^*\right)$, it is possible to construct prices (i.e., wages) that support this allocation as a decentralized equilibrium. Because we showed earlier that the decentralized equilibrium is unique, then this equilibrium must be efficient. ■

The intuition for this proposition is straightforward. As there are no labor market frictions, perfect competition for talent implies that the allocation of workers to firms is efficient. Thus, the only potential source of inefficiency is the choice between the retention of an old worker and the hiring of a young worker. Hiring a young worker is a potential source of externalities, as everyone learns about the talent of a young worker, which increases the number of options available to all players. However, because a firm that hires a young worker can extract all of the worker's surplus by charging a negative wage, and because Bertrand competition implies that poachers obtain zero net surplus from their poaching activity, a firm extracts all of the expected surplus from its decision to hire a young worker. Thus, the firm internalizes all of the potential costs and benefits of such a decision, and thus the firm's optimal private decision is also socially optimal.

Although it is not surprising that under symmetric learning the first-best outcome is achieved, we note that, unlike the static case, a hypothetical social planner has to consider two different trade-offs. First, we require an efficient allocation of workers to firms. As discussed above, the social planner would then choose the poaching threshold $\tau^{\#}$ by trading off the loss in firm-specific skills and the gain from assigning a worker to a more productive firm. Second, the social planer must find the optimal rate of talent discovery. The social planner chooses the retention threshold $\underline{\tau}_i$ by trading off the loss in firm-specific skills and the gain from sampling a young worker and learning about the worker's type in the subsequent period.

4.2. Asymmetric Learning

We begin by noting that Lemmas 1 and 2 continue to hold; these results are invariant to the number of periods in the game. Therefore, in any equilibrium, only the best workers are retained, and all retained workers are offered the same wage.

To characterize the equilibrium, in each period, we need to find three types of thresholds. As discussed above, $\tilde{\tau}_i$, $i \in \{l, h\}$, denotes the threshold such that all types $\tau \geq \tilde{\tau}_i$ are retained. Here, the only difference from the static case is that both types of firms can retain workers. We define $\underline{\tau}_i$ (as in the symmetric learning case) as the threshold for which all types $\tau \leq \underline{\tau}_i$ are fired.[‖] Finally, we define $\hat{\tau}_i$ as the lowest type that is poached in equilibrium.

An equilibrium is fully determined by a sequence of thresholds $\{\tilde{\tau}_l, \tilde{\tau}_h, \hat{\tau}_l, \hat{\tau}_h, \underline{\tau}_l, \underline{\tau}_h\}_t$, $t = 0, 1, .., \infty$. For simplicity, we focus only on equilibria in which these thresholds are time-invariant. Thus, we can drop the time subscript from the analysis that follows.

Result 8 *Any equilibrium* $\{\tilde{\tau}_l, \tilde{\tau}_h, \hat{\tau}_l, \hat{\tau}_h, \underline{\tau}_l, \underline{\tau}_h\}$ *has the following properties:*

Proposition 1 *1.*

2. *For a given* $\{\hat{\tau}_l, \hat{\tau}_h, \underline{\tau}_l, \underline{\tau}_h\}$, *threshold* $\tilde{\tau}_i$ *is either* $\overline{\tau}$ *or the least element of the set of fixed points of*

$$G_i(\tau) \equiv \gamma\mu + \frac{w_i^* - w_i^y - \delta i \int_{\tau}^{\overline{\tau}} (x - \gamma\mu) dF(x)}{i\left[1 + \delta(1 - F(\tau))\right]}. \tag{75}$$

(Recall that $i \in \{1, \theta\}$*). All retained workers are offered wage*

$$w_i^* = \max\left\{\theta\gamma\left(\int_{\tilde{\tau}_i}^{\overline{\tau}} \frac{\tau f(\tau) d\tau}{1 - F(\tilde{\tau}_i)} - \mu\right) + w_h^y - \frac{\delta\int_{\tilde{\tau}_h}^{\overline{\tau}}(\theta\tau - w_h^* - \theta\gamma\mu + w_h^y) f(\tau) d\tau}{\left[1 + \delta(1 - F(\tilde{\tau}_h))\right]}, 0\right\}, \tag{76}$$

[‖] As above, there could be a subset P of types that are poached in equilibrium. For simplicity, we focus only on cases in which P is an interval.

where

$$w_h^* = \max\left\{\theta\gamma\left(\int_{\tilde{\tau}_h}^{\overline{\tau}} \frac{\tau f(\tau)d\tau}{1-F(\tilde{\tau}_h)} - \mu\right) + w_h^y - \delta\theta\int_{\tilde{\tau}_h}^{\overline{\tau}} (1-\gamma)\tau f(\tau)d\tau, 0\right\}, \tag{77}$$

all workers who are poached (if any) are paid

$$w_i^{**} = \theta\gamma\left(\int_{\hat{\tau}_i}^{\tilde{\tau}_i} \frac{\tau f(\tau)d\tau}{F(\tilde{\tau}_i) - F(\hat{\tau}_i)} - \mu\right) + w_h^y - \frac{\delta\int_{\tilde{\tau}_h}^{\overline{\tau}} (\theta\tau - w_h^* - \theta\gamma\mu + w_h^y) f(\tau)d\tau}{1+\delta(1-F(\tilde{\tau}_h))}, \tag{78}$$

and all young workers who agree to work for a type-i firm are offered wage

$$w_i^y = -\delta(1-F(\tilde{\tau}_i))w_i^* - \delta(F(\tilde{\tau}_i) - F(\hat{\tau}_i))\max\{w_i^{**}, 0\}. \tag{79}$$

3. *All types* $\tau_i \in [0, \underline{\tau}_i]$ *are fired in equilibrium.*

4. *If* $\underline{\tau}_i < \tilde{\tau}_i$*, then types* $[\hat{\tau}_i, \tilde{\tau}_i]$*, with* $\hat{\tau}_i \geq \underline{\tau}_i$*, are poached in equilibrium and are paid* w_i^{**}*.*

Proof. In what follows, we take $\{\hat{\tau}_l, \hat{\tau}_h, \underline{\tau}_l, \underline{\tau}_h\}$ as given. Our goal is to find the unique $\{\tilde{\tau}_l, \tilde{\tau}_h\}$ conditional on the other thresholds. Because many of the steps are similar to those in the proof of Proposition 4, we refer the reader to that proof in some instances.

Lemma 2 implies that an equilibrium with retention must have a threshold $\tilde{\tau}_i$. Lemma 1 implies that all types in $[\tilde{\tau}_i, \overline{\tau}]$ are paid the same wage. To prevent poaching, this wage must be such that $w_i^* \geq w^p(w_i^*)$ ($w^p(.)$ will be derived below). Because poachers know that all types in $[\tilde{\tau}_i, \overline{\tau}]$ are offered w_i^*, their beliefs must be given by $F(\tau \mid \tau \geq \tilde{\tau}_i)$ upon observing w_i^*. The poaching wage offered by a type-H company with a vacant position is implicitly determined by the following condition:

$$V_h^p(\tau \geq \tilde{\tau}_i) - V_h^y = 0, \tag{80}$$

where

$$V_h^p(\tau \geq \tilde{\tau}_i) = \theta\gamma\int_{\tilde{\tau}_i}^{\overline{\tau}} \frac{\tau f(\tau)d\tau}{1-F(\tilde{\tau}_i)} - w^p(w_i) + \delta V_h^y, \tag{81}$$

$$V_i^y = i\gamma\mu - w_i^y + \delta V_i^o, \tag{82}$$

and

$$V_i^o = F(\tilde{\tau}_i)V_i^y + (1-F(\tilde{\tau}_i))\left(\int_{\tilde{\tau}_i}^{\overline{\tau}} \frac{i\tau f(\tau)}{(1-F(\tilde{\tau}_i))}d\tau - w^p(w_i) + \delta V_i^y\right). \tag{83}$$

From equations (82) and (83), we obtain:

$$V_i^o - V_i^y = \frac{1}{1+\delta(1-F(\tilde{\tau}_i))}\left(\int_{\tilde{\tau}_i}^{\overline{\tau}} (i\tau - w^p(w_i) - i\gamma\mu)f(\tau)d\tau + (1-F(\tilde{\tau}_i))w_i^y\right). \tag{84}$$

The poaching wage offered by a type-H firm upon observing w_i^* is

$$w^p(w_i^*) = \theta\gamma\left(\int_{\tilde{\tau}_i}^{\overline{\tau}} \frac{\tau f(\tau)d\tau}{1-F(\tilde{\tau}_i)} - \mu\right) + w_h^y - \frac{\delta\int_{\tilde{\tau}_h}^{\overline{\tau}}(\theta\tau - w_h^* - \theta\gamma\mu + w_h^y)f(\tau)d\tau}{[1+\delta(1-F(\tilde{\tau}_h))]}. \tag{85}$$

Using this poaching wage, we can now proceed exactly as in the proof of Proposition 4 to show that $w_i^* = \max\{w^p(w_i^*), 0\}$ if the equilibrium threshold is $\tilde{\tau}_i$ for $i \in \{l, h\}$.** Solving it for w_h^*, we obtain

$$w_h^* = \max\left\{\theta\gamma\left(\int_{\tilde{\tau}_h}^{\overline{\tau}} \frac{\tau f(\tau)d\tau}{1-F(\tilde{\tau}_h)} - \mu\right) + w_h^y - \delta\theta\int_{\tilde{\tau}_h}^{\overline{\tau}}(1-\gamma)\tau f(\tau)d\tau, 0\right\}, \tag{86}$$

which can be plugged into (85) to find w_l^*.

Because $w_i^* = \max\{w^p(w_i^*), 0\}$, a necessary condition for an incumbent with type $\tau \in [\tilde{\tau}_i, \overline{\tau}]$ not to deviate and fire the worker is:

$$V_i^o(\tilde{\tau}_i) \geq V_i^y, \tag{87}$$

where

$$V_i^o(\tilde{\tau}_i) = i\tilde{\tau}_i - w_i^* + \delta V_i^y. \tag{88}$$

Hence, after some rearranging, condition (87) becomes:

$$i\left(\tilde{\tau}_i - \gamma\mu\right) - \frac{w_i^* - w_i^y + \delta i\int_{\tilde{\tau}_i}^{\overline{\tau}}\left(\tau - \gamma\mu\right)f\left(\tau\right)d\tau}{1+\delta(1-F(\tilde{\tau}_i))} \geq 0. \tag{89}$$

The wage offered by type-H firms to workers with talent $\tau \in [\hat{\tau}_i, \tilde{\tau}_i]$ is determined by the following condition (from Bertrand competition):

$$V_h^p(\tau \in [\hat{\tau}_i, \tilde{\tau}_i]) = V_h^y, \tag{90}$$

where

$$V_h^p(\tau \in [\hat{\tau}_i, \tilde{\tau}_i]) = \theta\gamma\int_{\hat{\tau}_i}^{\tilde{\tau}_i} \frac{\tau f(\tau)d\tau}{F(\tilde{\tau}_i) - F(\hat{\tau}_i)} - w_i^{**} + \delta V_h^y. \tag{91}$$

We use equations (82), (83), and (91) to derive the wage for those workers who are poached

**Formally, we need to modify Assumption E2 slightly to fit the dynamic setup: After observing an off-the-equilibrium-path wage w_i', poachers believe that the probability that type $\tau' \geq w_i' + i\gamma\mu - w_i^y + \delta\left(V_i^o - V_i^y\right)$ deviates is no less than the probability that type $\tau'' > \tau'$ deviates. The application of this equilibrium refinement thus depends on some other equilibrium values (w_i^y, V_i^o, and V_i^y); this creates no difficulties as the condition can always be checked for each candidate equilibrium.

in equilibrium:

$$w_i^{**} = \theta\gamma\left(\int_{\hat{\tau}_i}^{\tilde{\tau}_i} \frac{\tau f(\tau)d\tau}{F(\tilde{\tau}_i) - F(\hat{\tau}_i)} - \mu\right) + w_h^y - \frac{\delta\int_{\tilde{\tau}_h}^{\overline{\tau}}(\theta\tau - w_h^* - \theta\gamma\mu + w_h^y)f(\tau)d\tau}{1 + \delta(1 - F(\tilde{\tau}_h))}. \tag{92}$$

The participation constraint of a young worker is given by

$$w_i^y = -\delta(1 - F(\tilde{\tau}_i))w_i^* - \delta(F(\tilde{\tau}_i) - F(\hat{\tau}_i))\max\{w_i^{**}, 0\}. \tag{93}$$

We now discuss the existence and uniqueness of the threshold $\tilde{\tau}_i$. Define the function:

$$G_i(\tau) = \gamma\mu + \frac{w_i^* - w_i^y + \delta i\int_{\tau}^{\overline{\tau}}(x - \gamma\mu)f(x)\,dx}{i\left[1 + \delta(1 - F(\tau))\right]}. \tag{94}$$

The existence of an equilibrium with retention requires this function to be non-negative for some $\tilde{\tau}_i$. Because, $G_i(\tau)$ is continuous and $G_i(0) > 0$, at least one fixed point exists if and only if $\max_{\tau\in[0,\overline{\tau})}\tau - G_i(\tau) \geq 0$. As before, if this latter condition does not hold, then no type is retained by firm i in equilibrium, i.e., $\tilde{\tau}_i = \overline{\tau}$. If $\max_{\tau\in[0,\overline{\tau})}\tau - G_i(\tau) \geq 0$, this proves the existence of at least one threshold τ' such that $\tau' = G_i(\tau')$.

Among all such τ', we define $\tilde{\tau}_i$ as the lowest one. To show that this threshold is part of an equilibrium, notice that because $G_i(0) > 0$, $\tau - G_i(\tau)$ crosses zero from below at $\tilde{\tau}_i$, which is also a necessary condition for an equilibrium. To show that no other $\tau' > \tilde{\tau}_i$ can be an equilibrium, we use the same argument as in the the proof of Proposition 4. Thus, $\tilde{\tau}_i$ is uniquely determined given $\{\hat{\tau}_l, \hat{\tau}_h, \underline{\tau}_l, \underline{\tau}_h\}$. ■

From this proposition we conclude that the equilibrium displays the same type of talent misallocation as in the two-period model: The best types $[\tilde{\tau}_i, \overline{\tau})$ are retained and the mediocre types $[\hat{\tau}_i, \tilde{\tau}_i]$ are poached. Thus, our main conclusions continue to hold in the infinite-horizon model.

To understand the differences between the two versions, consider the poaching wage given by (76). This wage differs from the poaching wage in the two-period model only because of

$$w_h^y - \frac{\delta\int_{\tilde{\tau}_h}^{\overline{\tau}}(\theta\tau - w_h^* - \theta\gamma\mu + w_h^y)f(\tau)d\tau}{[1 + \delta(1 - F(\tilde{\tau}_h))]}. \tag{95}$$

This expression has two terms: the present value of payments to a young worker and the option value of hiring a young worker. The first term is a consequence of our assumption of unlimited liability and the second term reflects the value of learning. Because w_h^y is negative, both of these terms reduce the incentives to poach and thus make it less likely that inefficient poaching occurs in equilibrium. However, these terms are often not large enough to eliminate poaching entirely. Furthermore, if we impose limited liability, we obtain $w_h^y = 0$, and expressions (75), (76) and (78) are simplified and remain valid.

In the infinite-horizon model, a new result is that inefficient poaching is less likely to be sustained in equilibrium. An equilibrium without poaching is more likely if firms are more patient (i.e., if δ is high). Conversely, if firms are very impatient (i.e., if δ is low), inefficient poaching is more likely.[††] Firms might become more impatient because they experience a higher death rate, which could be a consequence of tougher competition in product markets. Thus, increasing "short-termism" makes inefficient poaching more likely.

[††] As $\delta \to 0$, the equilibrium converges to an equilibrium of the static case.