Employer Learning and General Human Capital*

Jan Eeckhout †

March 2006

Abstract

We develop a model where competing employers gradually learn about a worker's productivity, like in the standard Jovanovic (1979) learning model. Rather than assuming that productivity is match-specific, we allow for general human capital productivity. We consider competitive wage determination through matching wage offers and counter-offers. We can also characterize the wage dynamics. Some of the results we obtain are similar to those under match-specific human capital. Others differ considerably.

Keywords. Labor Markets. Matching. Wage Setting. General Human Capital. Turnover.

*I am grateful to numerous colleagues and seminar audiences for discussion and comments.

†Department of Economics, University of Pennsylvania, 3718 Locust Walk, Philadelphia, PA 19104; `eeckhout@ssc.upenn.edu`; `http://www.ssc.upenn.edu/~eeckhout/`

1 Introduction

In this paper we propose a new framework for studying employer learning in the labor market. The basic structure resembles that of the standard models of learning about worker productivity in the sense that employers only gradually learn a worker's productivity and that firms compete for the same worker. We keep the assumption of employer competition because it seems a convincing way to pin down wage determination, which in turn is key to understanding the efficient allocation of workers to firms and the mobility between firms. The approach in this paper is nonetheless fundamentally different from existing models of learning in the labor market. Rather than assuming that employers gradually learn about the match-specific productivity, we assume that employers learn about general human capital productivity. We will establish that it is feasible to develop a theory of employer learning in which wage determination by competing firms is preserved but the assumption of match-specific productivity is dropped.

While the learning model with match-specific human capital is a particularly tractable and well tested model of the labor market, we study an environment in which firms learn about general human capital productivity to see whether the results from the previous learning model with match-specific human capital extend. It turns out that some results extend, while others do not. The main point is that broadly speaking, wage dynamics in an environment with learning do not hinge exclusively on the assumption that human capital is match-specific. At the same time we can identify a number of characteristics of the general human capital model that produces different predictions.

It has of course long been recognized that there is general human capital component to learning. Firms who learn about an employee's productivity that is valued in other firms as well. Consider for example a programmer who is good at writing computer code for Microsoft. It is likely that she will be good at it for Sun as well. Therefore, when an employer learns about the productivity of its employee, it realizes that the worker will also be more productive in other firms. Moreover, empirical evidence supports the claim that wage growth is attributed to general human capital.[1] But also from a theoretical viewpoint, gradual learning about general human capital is a necessary requirement for some standard theories. Labor market signalling for example implicitly assumes the gradual accumulation of general human capital. If firms exclusively learn about match-specific human capital, there is no return to investing in education to signal high productivity.

The basic point of departure is the learning model of Jovanovic (1979) (see also Felli and Harris (1996) and Moscarini (2005) amongst others). Employers gradually obtain information about the productivity of a worker and compete in wage offers for that worker. In contrast to Jovanovic (1979), each firm receives private information about the worker's general human capital productivity and the underlying information generating process is common to all competing employers. In this private information environment, employers who hire a worker steadily accumulate information about her productivity. We allow for a general formulation of the model in which employers accumulate different amounts of information. It is quite plausible that an incumbent firm that has employed a worker for a

[1] See amongst other Barlevy (2004) and Altonji and Shakotko (1986). Manovskii and Kambourov (2004) provide evidence that wage inequality and occupational mobility are closely related. This fits our model as long as the competing firms are vying for the worker in the same occupation.

certain time will have more precise information about that worker's productivity. And while lot of information can be gathered by an outsider firm about that worker's productivity (through interviews, references, resume, visits,...), it seems natural to allow for the fact that the current employer has superior information. In this environment with potentially differential private information, we investigate whether the worker can exploit the wage competition between employers: are workers able to extract all rents? Full rents are extracted in the match-specific learning model (Jovanovic (1979)),[2] where employees are paid their marginal product, conditional on the information available. We will establish whether this property is maintained in the current model.

We model the wage setting as a second price auction. We believe the second price auction corresponds to a realistic wage setting scenario between competing firms.[3] Consider an employee at the incumbent firm who receives an outside wage offer from a challenging firm. She takes the outside wage offer to the incumbent who either matches the offer, or lets the worker leave to take the job at the challenger firm. An offer is matched as long as, based on the information available, the expected productivity is larger than the wage offer. It is important to note that the wage offer by the other firm conveys information: if a firm is willing to make a certain wage offer, that implies that based on the information that firm has, the expected productivity exceeds that wage offer. Wage offers are matched until wages exceed one of the two competing firms' expected productivity. The winning firm pays the loser's valuation and receives positive expected surplus. This wage setting scenario corresponds to an ascending bid or English auction, which in this environment is equivalent to the second price auction. Observe that with private information about the common value of worker productivity, firms take into account the *winner's curse*. In fact, we show that under almost complete information[4] and sufficient differential information, the winner's curse will lead to extreme conservative bidding despite having received good signals.

In this environment, turnover arises even in the absence of firm specific human capital. The dynamic framework involves a worker employed by a given firm, the incumbent, who gradually accumulates information about productivity. In every period, the worker in addition matches with a new firm, the challenger, who receives a signal about the worker's productivity. The incumbent and the challenger compete for employing the worker using the second price auction. The worker exploits the employer competition and chooses the job with the highest wage, which leads to turnover. This is repeated in each period, where the new incumbent may be either last period's incumbent or last period's challenger. Whichever firm wins the auction will be better informed than any future challenger because it can invert the current competing firm's bid to obtain the signal it received. As the age of a worker increases, more information is revealed but at the same time, the differential between the incumbent and the challenger's information increases. In the case of match-specific learning, more information leads to a decrease in turnover. With learning about general human capital however this remains an open question. The more information is revealed about a worker, the fiercer incumbents and challengers are willing to bid (if the information is good).

[2] See also Felli and Harris (1996) for wage determination games that involve strategic interaction rather than "price-taking" behavior. They show that upon switching firms, the worker gets paid his expected marginal product.

[3] Our model contributes to the growing literature on labor market search with matching wage offers (see amongst many others, Postel-Vinay and Robin (2002)). The main contribution here is to add incomplete information in the wage matching process.

[4] That is, the probability that a worker is of high (low) productivity when a good (bad) signal is observed converges to one.

We will be able to establish whether this tenure-turnover result[5] also holds in the general human capital model.[6]

Finally, the dynamic environment allows us to characterize the equilibrium wage distribution.[7] We show that employer competition will determine the moments of the wage distribution when there is general human capital learning. The equilibrium wage density for the discrete types model we derive is not everywhere upward sloping and it is worth pointing out that the shape of the distribution derived here is not driven by the heterogeneity of outside options or observable productivity differences, but due to the learning experience of different competing firms.

The paper is organized as follows. In section 2, we lay out the model and in section 3 we describe the information filtering problem. In section 4, we solve the subgame: we establish whether the worker can extract all the surplus, we characterize the second price auction equilibrium and we solve the limit case of almost-complete information. In section 5 we study learning and wage determination in the dynamic setting. Section 6 ends with a discussion of the results in comparison with those of match-specific learning models.

2 The Model

The market is populated with a continuum of agents: workers and firms. Firms have exactly one job opening, and workers can perform one job at most at a time. The market operates over time and time is divided in discrete periods, starting at $-\infty$ and ending at ∞. There is a bounded but large measure of infinitely lived firms. Workers live for 3 periods, indexed by $t \in \{0, 1, 2\}$. In each period, a measure 1 of workers is born.

Matching. At the beginning of each period, unmatched firms and *all* workers, including the workers who are already matched, enter in the pool and are uniformly and randomly matched. Because the measure of firms is assumed to be large relative to the measure of workers, each worker is sure to be matched in every period. Newly born workers are matched with one firm, and older workers are matched with two firms $i \in \{I, C\}$, the newly matched firm (the challenger C) and last period's employer, the incumbent firm I.

Payoffs. The value to the firm i of employing the worker is $\pi^i V$, $V \in \mathcal{V} = \left\{\overline{V}, \underline{V}\right\}, \overline{V} > \underline{V}$: the worker is either of high productivity $V = \overline{V}$ or of low productivity $V = \underline{V}$. In each period, π^i is an i.i.d. draw from a non-negative cumulative density function $H(\pi)$. The realization of π^i is common knowledge to all parties in the match at the

[5] In the literature, an additional motive for increasing wage-tenure relationships has been identified, based on intertemporal incentives in the absence of commitment by the worker (see Burdett and Coles (2003), Krueger and Uhlig (2003), and Harris and Holmström (1982)). An increasing wage-tenure profile effectively locks in the risk-averse worker thereby providing insurance. Harris and Holmström (1982) in addition have learning as in Jovanovic (1979), which separately generates increasing wage profiles. One difference with our model is that in theirs, information is symmetric: all firms and the worker observe the information revealed. Wages are therefore determined competitively.

[6] Altnerative studies of turnover in the presence of ex ante heterogeneity include Moscarini (2001).

[7] There is an extensive theoretical labor literature studying wage distributions. In particular, our results relate to the work by Burdett and Mortensen (1998), Burdett and Coles (2003), Postel-Vinay and Robin (2002) and Moscarini (2005) who derive equilibrium wage distribution in the presence of on-the-job search.

time of matching. The firm that gets to employ the worker pays a wage $w^i, i \in \{I, C\}$. The worker's payoff is w^i and the firm's payoff is $\pi^i V - w^i$, if the firm employs the worker and zero otherwise. All agents discount expected future consumption using a common constant discount factor δ.

Information. Firms do not observe V. Both firms have common prior beliefs about V. With probability μ, a newborn is highly productive ($V = \overline{V}$) and with probability $1-\mu$ she is of low productivity, which we normalize to zero ($V = \underline{V}$). Firms observe noisy information about her productivity in the form of a private, real valued signal. A signal to firm i, $\sigma^i \in \{high, low\}$. Signals are distributed according to the conditional density $F\left(\sigma^i \mid V\right), V \in \mathcal{V}$, where $F\left(\sigma^i \mid V\right) = \Pr\left\{\sigma^i \mid V\right\}$. An employee who is of type $V = \overline{V}$ produces a signal *high* with probability $F(\sigma^i = high \mid \overline{V})$ and a signal l with probability $F(\sigma^i = low \mid \overline{V}) = 1 - F(\sigma^i = high \mid \overline{V})$. A type $V = \underline{V}$ produces a signal *high* with probability $F(\sigma^i = high \mid \underline{V})$ and a signal *low* with probability $F(\sigma^i = low \mid \underline{V}) = 1 - F(\sigma^i = high \mid \underline{V})$. Signals are privately observed by a firm, and in keeping with the nature of the problem at hand, signals are informative, i.e. $F(\sigma^i = high \mid \overline{V}) > \frac{1}{2} > F(\sigma^i = low \mid \overline{V})$ and $F(\sigma^i = high \mid \underline{V}) < \frac{1}{2} < F(\sigma^i = low \mid \underline{V})$. We will assume that the distribution of *individual* signals is identical to both firms. Each individual signal is equally informative to all firms i :

Assumption **A1**. $F\left(\sigma^I \mid V\right) = F\left(\sigma^C \mid V\right)$

In what follows and as a result of Assumption **A1**, we will use the shorthand notation $\overline{F} = F\left(\cdot \mid \overline{V}\right) \in \left(\frac{1}{2}, 1\right]$ and $\underline{F} = F\left(\cdot \mid \underline{V}\right) \in \left[0, \frac{1}{2}\right)$. Note that the assumption does not imply that information structures are symmetric. Asymmetry will be derived from the fact that some firm I observes *more* signals than firm C. Denote the number of signals observed by firm I by l and that of firm C by m. Then let s^I and s^C denote the cardinality of signals $\sigma^I = h$ and $\sigma^C = h$ where $s^I \in S^I = \{0, 1, ..., l\}$ and $s^C \in S^C = \{0, 1, ..., m\}$. We can now summarize the information as a double $\left(s^I, s^C\right) = s \in S = S^I \times S^C$.

We assume that workers are passive and always take the highest wage offer. This implies they cannot signal productivity, and that they cannot generate more meetings with other firms/

Beliefs. For a given $\left(s^I, s^C\right) \in S$, let $p\left(s^I, s^C\right)$ denote the joint probability of $\left(s^I, s^C\right)$ occurring. Because the only relevant information is the cardinality of the signals, $p\left(s^I, s^C\right)$ can be derived from the binomial distribution:

$$p\left(s^I, s^C\right) = \binom{l}{s^I}\binom{m}{s^C}\left[\mu \overline{F}^{s^I + s^C}\left(1 - \overline{F}\right)^{l+m-s^I-s^C} + (1-\mu)\,\underline{F}^{s^I+s^C}\left(1 - \underline{F}\right)^{l+m-s^I-s^C}\right]$$

Let $p\left(s^{-i} \mid s^i\right), i \in \{I, C\}$ denote the probability that player i attaches to player $-i$ receiving signal s^{-i} conditional on i receiving signal s^i. Then

$$p^C\left(s^I \mid s^C\right) = \frac{p\left(s^I, s^C\right)}{p\left(s^C\right)} = \frac{p\left(s^I, s^C\right)}{\sum_V p^C\left(s^C \mid V\right) p(V)}$$

Prior beliefs are given by $\mu_0 = \mu$, the probability that a newborn is of productivity $V = \overline{V}$. Then any posterior belief (i.e. the probability attached to the event that a worker is of type $\overline{V}$), given signals (j, k), will be denoted by $\mu(j,k)$. Then $\mu(j,k) = \Pr\left\{V = \overline{V} \mid s^I = j, s^C = k\right\}$ and is given by

$$\mu(j,k) = \frac{\mu\overline{F}^{j+k}\left(1-\overline{F}\right)^{l+m-j-k}}{\mu\overline{F}^{j+k}\left(1-\overline{F}\right)^{l+m-j-k} + (1-\mu)\underline{F}^{j+k}\left(1-\underline{F}\right)^{l+m-j-k}}$$

Given payoffs, the information structure and beliefs, we introduce some further notation to indicate the expectation of V given $\left(s^I, s^C\right)$. Let $v(j,k) = E\left[V \mid s^I = j, s^C = k\right]$, i.e. v is the expectation about V given signals (j,k). Then $v(j,k)$ is given by

$$v(j,k) = \mu(j,k)\overline{V} + (1-\mu(j,k))\underline{V}$$

3 Filtering the Information

In order to derive equilibrium in the dynamic game, it will be useful to consider in a first instance the case of an isolated subgame. Therefore, in this section we will assume that the continuation payoff is not affected by current actions, and without loss of generality, continuation values are normalized to zero. Given the nature of the information generation process, the following Lemma gives an insight into the way information is filtered.

Lemma 1 *For any j, k, j', k' :*

$$v(j,k) = v(j',k') \quad \textit{iff} \quad j+k = j'+k'$$

and

$$v(j,k) < v(j',k') \quad \textit{iff} \quad j+k < j'+k'$$

Proof. In Appendix ■

Given the result in Lemma 1, we introduce the following definitions in order to shorten notation. Let $K = j+k$ and $L = l+m$, then we will use the notation $p_L(K)$ to indicate:

$$p_L(K) = \mu\overline{F}^K\left(1-\overline{F}\right)^{L-K} + (1-\mu)\underline{F}^K\left(1-\underline{F}\right)^{L-K}$$

Note that $p_L(K)$ is not exactly equal to $p(j,k)$ since $p(j,k) = \binom{l}{j}\binom{m}{k}p_L(K)$. We will also define $\mu_L(K) = \mu(j,k)$, which is given by

$$\mu_L(K) = \frac{\mu\overline{F}^K\left(1-\overline{F}\right)^{L-K}}{p_L(K)} = \frac{\mu\overline{F}^K\left(1-\overline{F}\right)^{L-K}}{\mu\overline{F}^K\left(1-\overline{F}\right)^{L-K} + (1-\mu)\underline{F}^K\left(1-\underline{F}\right)^{L-K}}$$

and $v_L(K) = v(j,k) = \mu_L(K)\overline{V} + (1-\mu_L)\underline{V}$. It immediately follows from Lemma 1 that $v_L(K') > v_L(K)$ iff $K' > K$.

Basically, all aggregate information can be summarized in the total number of good signals K that have been observed. Then the following result states that the expected belief and the expected payoff conditional on the current information is constant as more signals are added.

Lemma 2 *The belief $\mu_L(K)$ and the expected payoff $v_L(K)$ are a martingale.*

Proof. In Appendix ■

The information the incumbent has can be ranked relative to that of the Challenger. This can be established by showing that the information structure satisfies Blackwell's sufficiency. We introduce the following notation. Let $\mathbf{i}_{Vj} = \Pr\left\{s^I = j \mid V\right\} \geq 0$ be the chance of j good signals in state $V \in \left\{\overline{V}, \underline{V}\right\}$ and let $\mathbf{I}$ be the Markov matrix of probability densities $\mathbf{I}_{[2,l+1]} \equiv [\mathbf{i}_{Vj}]$, i.e. where $\sum_{j=0}^{l} \mathbf{i}_{Vj} = 1$. Likewise for $\mathbf{c}_{Vk} = \Pr\left\{s^C = k \mid V\right\}$ and $\mathbf{C}_{2\times(m+1)} \equiv [\mathbf{c}_{Vk}]$.

Definition 3 *(Blackwell Sufficiency). The experiment induced by the Markov matrix* $\mathbf{P}$ *is more informative than* $\mathbf{Q}$, *(notation* $\mathbf{P} \supset \mathbf{Q}$*), if any payoff vector attainable with* $\mathbf{Q}$ *is also attainable with* $\mathbf{P}$.

Lemma 4 *If* $l \geq m \geq 0$, *then* $\mathbf{I}$ *is more informative than* $\mathbf{C}$.

Proof. In Appendix ■

The posterior probabilities induced by $\mathbf{C}$ lie in the convex hull of the posteriors induced by $\mathbf{I}$. Because the event in which less signals are observed is less informative in the sense of Blackwell, we can interpret the signal of the least informed as that of the more informed firm with "added noise".

4 Equilibrium of the Subgame

4.1 Matching Wage Offers

Consider the example in the introduction of the programmer who receives job offers from Sun, the challenger, and Microsoft, the incumbent. Both firms have private information and the challenger, based on it's signal makes a wage offer to the programmer who takes it to his current employer. If, conditional on its signal, the incumbent's expected productivity exceeds the outside offer, the offer will be matched. The worker now takes the matching offer to the challenger who will consider making a counter offer. Each firm will make matching counter offers as long as the expected productivity conditional on the signal exceeds the wage offer. This pattern of matching counter offers constitutes an ascending bid or English auction. In this environment with two bidders (the firms) and one auctioneer (the worker) who auctions off a common value commodity, the ascending bid auction is strategically equivalent to the second price auction.[8]

The second price auction is therefore a realistic scenario in which two firms with private information about the general human capital productivity of a worker compete over wages. In this section, we derive and characterize equilibrium wage offers.[9] Both firms simultaneously submit a wage offer, the highest of which wins (in case of a

[8] This equivalence does no longer hold with three or more bidders as the decision for one bidder to drop out reveals information about that bidder's signal to the remaining bidders.

[9] Nothing guarantees that the second price auction will be the constrained best mechanism from the point of view of the auctioneer (i.e. the mechanism that generates the highest expected revenue to the employee). We do know though that in the (different) framework of Milgrom and Weber (1982) the second price auction gives a higher expected profit than the first price auction. Moreover, the second price auction satisfies the Wilson doctrine, i.e. this auction institution is "detail-free".

tie, the worker chooses a firm with uniform probability).

Proposition 5 *(Matching Wage Offers). Let* $\frac{v_L(k+1)}{v_L(k)} > \frac{\pi^i}{\pi^{-i}} > \frac{v_L(k)}{v_L(k+1)}$ *for all* k*. Then in the second price auction, an ex-post Nash equilibrium strategy for any player* i *with signal* j *is to bid* $w^i(j) = \pi^i v_L(\min\{2j, L\})$*.*

Proof. In Appendix ■

The equilibrium payoffs, given signals (j,k) then are $\max\left\{\pi^I v_L(j+k) - \pi^C v_L(2k), 0\right\}$ to firm I and $\max\{\pi^C v_L(j+k) - \pi^I v$ to firm C. It will not be surprising that there is no full rent extraction. The worker gets paid at most her productivity: she gets exactly her productivity if $\pi^I = \pi^C$ and if the cardinality of good signals is the same for I and C (i.e. $j = k$), and strictly less in the case the cardinality of the signals differs. As a result, the expected equilibrium wage is below the marginal product.

It is possible that in addition to this equilibrium, other ex-post Nash equilibria exist. In particular, one can imagine strategies where one firm with j signals bids slightly above the valuation (for the sake of argument, let $\pi^I = \pi^C = 1$) $v_L(2j)$ and the other firm responding accordingly. This is akin to the multiplicity of Nash equilibria in a private values Vickrey auction.

It should also be noted that in the case of $\pi^I \neq \pi^C$, there is misallocation in equilibrium. With positive probability, the firm with the lower π^i will receive more positive signals than the firm with the higher productivity. Since the component $v_L(K)$ is pure common value, efficiency requires that the worker be employed in the high productivity job, whatever the ex post distribution of signals.

While this result is derived for a static wage setting environment, below we show how under certain circumstances (in the absence of observable wage and employment histories), the wage setting procedure extends to the 3 period dynamic model.

4.2 Matching Wage Offers under Almost-Complete Information

Though from Proposition 5 no full rent extraction is guaranteed, we expect that the employee can extract all rents provided the firms are perfectly informed. Suppose that both firms know exactly the productivity of the employee, then Bertrand-like competition (i.e. a second price auction with complete information) will drive down the profits of both firms. The only rents left to the firms will be due to the difference in productivity π^i. Of course, when $\pi^I = \pi^C$ no rents will remain to the firms. For the remainder of the paper, let $\pi^I = \pi^C$. While there is full rent extraction in the case of complete information, the next Proposition shows that this is not the case when information is almost-complete. Quite the contrary: all rents accrue to the firms when there is sufficient difference between the number of signals of the incumbent and the challenger. Almost-complete information implies that a good type draws a signal with probability $\overline{F}$ almost 1 and a bad signal with probability $\underline{F}$ almost 0. We consider the case in which $\overline{F}$ and $1-\underline{F}$ converge to 1 and let $r = \log \overline{F} / \log(1-\underline{F}) = 1$.

Proposition 6 *(Almost-Complete Information). Consider a sequence* $\left\{(\overline{F}, 1-\underline{F})\right\}_n$ *converging to* $(1,1)$ *and a corresponding sequence* $\{v_L(K)\}_n$*. Then,* $\forall l > 3m$*, the equilibrium bid* $v_L(2k)$ *in the second price auction of* C *is* $v_L(2k) = \underline{V}$*, irrespective of the cardinality of good signals* k *received by* C*.*

Proof. In Appendix ■

This result contrasts sharply with the fact that under complete information, all rents accrue to the worker. Given an outside option to either firm of zero, they will bid the worker's marginal product in the case of full information. Still, in the limit as signal precision becomes perfect, all rents accrue to the better informed firm.

In general, for any $x \leq m$, if the limit $\mu_L(2x)$ is zero, then this also holds for all signals $\{0, 1, ..., x\}$. Then x satisfies $l + m - 4x > 0$. To see this, note that

$$\mu_L(2x) = \frac{\mu}{\mu + (1-\mu)\left(\frac{\overline{F}}{1-\overline{F}}\right)^{l+m-4x}}$$

where the $\lim_{(\overline{F},1-\underline{F})\to(1,1)} \mu_L(2x) = 0$ as long as $l + m - 4x > 0$.

Observe that the limit result does not hold when $\pi^I \neq \pi^C$ because the qualifying condition in Proposition 5, that $\frac{v_L(k+1)}{v_L(k)} > \frac{\pi^i}{\pi^{-i}} > \frac{v_L(k)}{v_L(k+1)}$ for all k, is violated. As the sequence $\left\{(\overline{F}, 1-\underline{F})\right\}_n$ converges to $(1,1)$, that condition converges to $1 > \frac{\pi^i}{\pi^{-i}} > 1$ and is violated for any π^i, π^{-i} along the limit sequence n except when $\pi^I = \pi^C$. Even with $\pi^I = \pi^C$, the result in Proposition 6, does not necessarily imply that there is failure of upper hemi continuity of the equilibrium set since other equilibria may exist, the equilibrium of which converges to that of the limit game.

As a benchmark, we consider the opposite case of an entirely uninformative signal, i.e. $\overline{F} = \underline{F} = \frac{1}{2}$. The posterior is equal to the prior $\mu_L(K) = \mu$ for any L and K. To see this, note that

$$\mu_L(K) = \frac{\mu \frac{1}{2}^L}{\mu \frac{1}{2}^L + (1-\mu)\frac{1}{2}^L} = \mu$$

and as a result, $v_L(K) = \mu\overline{V} + (1-\mu)\underline{V}$ for all L and K.

4.3 Turnover and Differential Information

In this environment with differential information between the incumbent and the challenger, define turnover as the likelihood that the worker will get a better offer from the challenger than from the incumbent. It is well known from the standard learning model (Jovanovic (1979)) that turnover decreases as the information accumulated increases. Below, we show that this result also obtains in the current framework. More importantly however, we identify an additional source that affects turnover, namely differential information.

Proposition 7 *(Turnover and Differential Information) Consider any pair l, m such that $l + m = L$ and $l \geq m$. Then turnover decreases the more information is differential: turnover given l, m is greater than turnover given l', m' $(l' + m' = L)$ for all $l' > l$.*

Proof. In Appendix ■

It follows immediately from this result that if information is more differential with tenure (see below), this will lead turnover to decrease with tenure.

In addition to turnover, expected wages are also lower as information is more differential.

Proposition 8 *(Wages and Differential Information) Consider any pair l, m such that $l + m = L$ and $l \geq m$. Then expected wages decrease the more information is differential: expected wages given l, m are higher than expected wages given l', m' ($l' + m' = L$) for all $l' > l$.*

Proof. In Appendix ■

The following Corollary now derives immediately from both Propositions:

Proposition 9 *Turnover and expected wages are positively correlated for a given number of signals.*

The intuition for these results is that having more signals implies having more precise information, which in principle should push up the bid by the challenger. However, crucial in all these results is that the total number of signals L is kept constant. The exercise consists in evaluating the implication of redistributing those signals from the challenger to the incumbent. Such redistribution implies that the informational disadvantage of the challenger is exacerbated. This increases the winners' curse to the challenger who will therefore bid less and leading to lower probabilities of placing the winning bid, thus reducing turnover.

4.4 Full Surplus Extraction

The second price auction typically does not guarantee that the worker gets paid the expected marginal product. Moreover, when $\pi^I \neq \pi^C$ that implies that there is misallocation with positive probability and hence inefficiency. And while the second price auction is a particularly attractive wage determination from a positive point of view, an open question remains whether a more intricate mechanism exists that allows employees to extract full surplus from the bidding firms, thus obtaining an efficient allocation. The following Proposition shows that this is not the case and that employees are *always* paid below their marginal product.

Proposition 10 *(No Full Surplus Extraction). Under the information structure S, there exists no mechanism that can guarantee full extraction of surplus by the employee and hence efficient allocation.*

Proof. In Appendix ■

There exist no auction mechanism that can extract full rent in this environment. In the light of the Crémer-McLean (1988) results, this is surprising. The Crémer-McLean mechanisms involve lotteries in addition to the payment of the bid. The lottery is conditioned on the bids made by the other participants, and in the case of truth telling, the lottery will depend on the true characteristics of all the other bidders. Crémer-McLean show that the auctioneer can design the lotteries in such a way that the expected payoff of participation is zero when the true type is announced (i.e. the participant makes a bid consistent with her type), and negative when a bidder lies. This is possible provided there is enough variation in the probability distribution as an agent's type changes. For the case of rent extraction using dominant strategy mechanisms, this imposes a full rank condition on the matrix of probabilities in function of the signals of each of the bidders. Unfortunately, this full rank condition is not satisfied in our model. This is easily seen in the case of asymmetry in the number of signals between each of the firms, i.e. j signals for the incumbent and k for the challenger firm. Then the full rank condition requires the matrix to be of rank j while the rank of this asymmetric matrix is at most $k < j$. This is of course due to the fact that we restrict attention to mechanisms with one worker and two firms only: there is not enough variation in the probability distribution to condition the bids of the incumbent on as there are only $k < j$ possible lotteries to pin down j different signals.[10]

Under the information structure presented here, even Bayesian Nash mechanisms cannot guarantee full rent extraction. The rank of the matrix of probabilities is at most 2, and as a result, when either of the agents receives more than one signal, there will not exist lotteries that are incentive compatible.

Of course, as is the case throughout the paper, this result hinges on the fact that the worker does not have any information whatsoever about the employers' signals. See further below in Section 6 on the discussion employee actions.

5 Wage Dynamics

As was laid out in section 2, all workers live for three periods $t \in \{0, 1, 2\}$. Any newborn has high ability with probability μ and low ability with probability $1 - \mu$. At the beginning of each period, firms observe one signal. The firm that hires the worker in the first period faces no competition. At the beginning of each consecutive period, the incumbent firm observes one signal in addition to the information it has accumulated in the past. In addition, the worker is matched with a challenger firm, who receives one signal. In this section, we will analyze how wages evolve over the life cycle of an employee while infinitely lived firms compete for them. Firms always observe the age of a worker, and initially we assume that they do not observe employment histories and wages. We relax that assumption further on. In this section we also will assume that firms are identical, i.e. $\pi^i = 1$, for all i.

5.1 Wage Dynamics with Non-observable Employment Histories

Given finite lives of the employees, we will solve backwards. We know exactly what happens at the beginning of the last period of life when two firms are competing for a worker, since we derived the equilibrium strategy profile in the case of continuation payoffs that are unaffected by the current strategies. At the beginning of period 1, each of the firms will take the continuation payoffs as given, and we observe the following.

Observation. When wages are determined using Second Price Auctions, the information partition after realization of the first auction is the same irrespective of whether I wins or whether C wins.

To see why this is true, note that by observing the offer of the losing bid, the winning employers can invert the loser's offer and deduce the information that she had received. Then, conditional upon winning, both the

[10] Suppose there were 3 firms, one of whom observes 3 signals and the two others each observe 2 signals. This is a case of asymmetry, but note that for a general information structure the full rank condition can be satisfied. In particular, the matrix of signals for the first bidder is a 3×4 matrix, i.e. 3 signals against the total number of signals of *all* other players. As a result, lotteries can be constructed that ensure full rent extraction in dominant strategies.

incumbent and the challenger have the same information. Given each employer observes a signal in each period, the total number of signals in period 1 is 3 (2 signals by I from period 0 and period 1, and 1 signal by C in period 1). In period 2, the number of signals conditional upon winning is 5. The winner in period 1 is left with 3 signals and becomes Incumbent in period 2, in which she receives a new signal. While she has 4 signals, the challenger (different from the challenger in period 1) has 1 new signal.

In period 2, payoffs are determined by the second price auction, basically with zero continuation value. The valuation of the object is $v_5(K)$ where K is the total number of good signals observed. For example, when the incumbent has received 2 good signals and the challenger has received one good signal, I will bid $v_5(3)$ and C will bid $v_5(2)$. I wins the auction and gets profits $v_5(3) - v_5(2)$. In period 1, the bids will have to take into account the continuation payoffs. Let $V_1(K)$ denote the continuation payoff in period 1 given K good signals were observed. Then it is given by

$$V_1(K) = v_3(K) + \frac{\delta}{p_3(K)} \left\{ \sum_{J=0}^{1} p_5(K+J) \max\{W_0, v_5(K+J) - v_5(0)\} + \sum_{J=1}^{2} p_5(K+J) \max\{W_0, v_5(K+J) - v_5(2)\} \right\},$$

where W_0 denotes the continuation value of a firm that is unmatched, and W_1, W_2 the continuation value to a firm conditional on matching with a worker of age $1, 2$ respectively (and before observation of the signal).

Observation. 1. $v_L(K) > v_L(J), K, J \in \{0, 1, .., L\}, \forall K > J, \forall L$;

2. $V_1(K) > V_1(J), K, J \in \{0, 1, 2, 3\}, \forall K > J$;

This implies that firms can bid in period 1 in the second price auction where the value of the auction is given by the continuation value. Then in period 1, after matching to an employee, the continuation value to the firm is

$$\begin{aligned} V_0 &= \sum_{K=0}^{1} \left\{ p_1(K) v_1(K) + \delta \left\{ \sum_{J=0}^{1} p_3(K+J)[V_1(K+J) - V_1(0)] + \sum_{J=1}^{2} p_3(K+J) \max\{W_0, V_1(K+J) - V_1(2)\} \right\} \right\} \\ &= p_1(0) v_1(0) + p_1(1) v_1(1) + \delta 2 p_3(1)[V_1(1) - V_1(0)] + \delta p_3(2)[V_1(2) - V_1(0)] + \delta p_3(3)[V_1(3) - V_1(2)] \end{aligned}$$

Remind that there is a measure of $\phi > 5$ firms and a measure 1 of new-born workers in each period. Then expected profits to firms before matching:

$$W_0 = -a + \frac{\phi - 5}{\phi - 2} \delta W_0 + \frac{1}{\phi - 2} \delta [V_0 + \varphi_1 W_1 + \varphi_2 W_2 + (2 - \varphi_1 - \varphi_2) W_0]$$

where φ_t is the probability that a t worker changes employer from the incumbent I to the challenger C and

$$\begin{aligned} \varphi_1 &= \frac{1}{2} p_3(0) + p_3(1) + \frac{1}{2} p_3(2) \\ \varphi_2 &= \frac{1}{2} p_5(0) + p_5(1) + \frac{3}{2} p_5(2) \end{aligned}$$

and where

$$\begin{aligned} W_1 &= p_3(1)[V_1(1) - V_1(0)] + \left(\frac{1}{2} p_3(0) + \frac{1}{2} p_3(2) \right) W_0 \\ W_2 &= p_5(1)[v_5(1) - v_5(0)] + \left(\frac{1}{2} p_5(0) + \frac{3}{2} p_5(2) \right) W_0 \end{aligned}$$

Free entry of firms will drive profits to zero. As long as W_0 is positive, firms will enter the market. Then ϕ solves

$$\phi \in \{\phi > 5 : W_0 = 0\}$$

This is always satisfied provided a is small enough. We will assume that both requirements are fulfilled. This implies

$$\phi = \frac{\delta}{a}\left[V_0 + \varphi_1 W_1 + \varphi_2 W_2\right] + 2$$

The equilibrium wage distribution $G(w)$ with probability density function $g(w)$ is given by:

$$g(w) = \begin{cases} \frac{1}{3} & \text{if } w = 0 \\ \frac{1}{3}\left[p_5(0) + 5p_5(1) + 6p_5(2) + 4p_5(3) + p_5(4)\right] & \text{if } w = v_5(0) \\ \frac{1}{3}\left[p_3(0) + 3p_3(1) + p_3(2)\right] & \text{if } w = V_1(0) \\ \frac{1}{3}\left[4p_5(2) + 6p_5(3) + 4p_5(4) + p_5(5)\right] & \text{if } w = v_5(2) \\ \frac{1}{3}\left[2p_3(2) + p_3(3)\right] & \text{if } w = V_1(2) \end{cases}$$

Note that the average product is μ whereas the average wage is strictly lower than μ. This is due to the fact that there is no full rent extraction. However, some wages are above marginal product (even ex-post marginal product, so not only above expected marginal product). In particular, those wages paid to the employees who received good signals in the first period. The reason is the option value of a good worker combined with the fact that firms don't have to pay the marginal product in each period (i.e. in the future). Because in the future, the informational advantage to the incumbent is even greater, the winner of the auction has a much higher expected continuation value and is willing to bid above marginal product.

5.2 Observable Wages and Employment Histories

One obvious way to extend the model is to allow for more information transmission as tenure increases. It seems reasonable to assume that challengers observe the employment history of employees they consider hiring. In particular, the current wages will convey information about the signals received by past employers. Introducing observable wages is open to potential complications. An immediate reaction may be that wages will make the whole idea of informational advantage to the incumbent evaporate. All information is embodied in the wage and as a result, there is no longer any advantage to the incumbent. While it is true that observable employment histories and wages decreases on average the degree of the informational advantage, it does not disappear. The reason is that wages in a second price auction are equal to the bid of the loser. As such, they reflect the information that was available to the loser. The exact information available to the winner is not revealed through the past wages.

If both bidders can deduce a public component of the information from the employment history, then the second price auction will be similar, except for the public component.

Proposition 11 *If J good signals have been publicly observed, then in the second price auction, an ex-post Nash equilibrium strategy for any player i with private signal j is to bid $w^i(j) = v_L(J + 2j)$.*

Proof. This is an immediate extension of the proof of Proposition 5. ■

Deriving the dynamic wage path in the presence of observable wages and employment histories however is open to a mayor complication. The employee may have an incentive to choose the lower of the two wages offered in order to select that employer (typically the challenger) who has observed less information. The lower current wage is compensated by the fact that the informational advantage of that employer in the future is smaller, implying that the worker can extract more rents. As is discussed below, a worker always has incentives to engage in activities that reduce the informational advantage of the employer. This of course will affect the bidding strategy of both firms in the first place.

6 Discussion

In order to put the results in this paper in the perspective of the match-specific learning model, we discuss several issues on which both models make predictions. Some of our results are similar to those in the match-specific learning model, others differ.

Wages and Marginal Product. One of the main findings in the match-specific learning model is that equilibrium wages exist that are equal to the marginal product conditional on the available information. Jovanovic (1979) shows that such a market equilibrium exists. Felli and Harris (1996) show that when introducing a wage setting mechanism with strategic interaction between the employers, expected marginal product is paid upon switching employers. In the case of general human capital learning, the result no longer holds. No wage setting mechanism exists in which workers are paid their expected marginal product conditional on the available information. For sufficiently differential information between the two competing employers, rents extracted converge to zero and the wage converges the low type's productivity level as information becomes almost-complete. There is a discontinuity in the sense that under complete information, the high type gets paid his marginal product. Because in the case of match-specific learning the worker always gets paid his marginal product, this is also true under almost-complete information.

Wage Distribution. In the dynamic context, because of future rent extraction, employers are not just bidding for the current marginal productivity, but also for all future rents to be extracted. This may lead workers who have generated sufficiently good signals to be paid above their marginal product. This does of course not occur in the match-specific learning model where workers are paid their marginal product in each period.

Turnover – Job Tenure Relation. Exactly as in the case of the match-specific learning model, turnover and job tenure are negatively related. The underlying reasons for this relation are only partly common. In the match-specific learning model, information gradually becomes more precise, which makes it more likely a worker will settle on a particular job. In the general human capital learning model, information becomes more precise over time, but that makes bidding for the worker also more attractive to the challenger firms. If a worker has been in employed at the same firm for a long time, this reveals positive information about the productivity. Whether turnover will increase or decrease is therefore not immediate. However, because over time the information between the incumbent and the challenger becomes more and more differential, the incumbent has an advantage in retaining the worker. Therefore, turnover (and the separation hazard) decreases with tenure.

Turnover – Job Separation Relation. Conditional on a higher wage in the first period (and therefore conditional on better signals realized), the worker will have a lower separation hazard. Good workers are less likely to turn over. This is true also in Jovanovic (1979). Again, the reason why this holds is different here. Rather than having more precise information, the incumbent has a greater informational advantage over the challenger.

Employee Actions. As in the case of the match-specific learning model, the employee remains conspicuously silent in the current model.[11] In addition, we assume that workers have no information about the signals received by the employers. The motivation for this were modelling choices that would keep the analysis tractable. However, realism forces us to admit that workers will not remain passive and that they have access to common information. First, they can – and generally do – affect the meeting probabilities. The gains from offers are potentially huge. Since the worker is paid below her expected marginal product, she has a strong incentive to generate information flows by seeking new offers. Second, the mechanism by which wages are set is predetermined. Whether it is a first-price auction, as second-price auction, wage matching or any other extensive form Bayesian game, the worker will often be able to manipulate the particular wage determination procedure. For example, since most often offers are made to the employee, she can decide whether or not to communicate the offer. If the offer is bad, she may prefer to pretend that no offer has been made, whereas if the offer is good, she will have an interest to signal that the offer reflects good information by the challenger.

[11] For a model of information acquisition of workers about jobs, see Acemoglu and Shimer (2000).

7 Appendix

Proof of Lemma 1

Lemma 1. *For any* j, k, j', k' :

$$v(j,k) = v(j',k') \quad \text{iff} \quad j+k = j'+k'$$

and

$$v(j,k) < v(j',k') \quad \text{iff} \quad j+k < j'+k'$$

Proof. The posterior is given by

$$\mu(j,k) = \frac{\mu \overline{F}^{j+k} \left(1-\overline{F}\right)^{l+m-j-k}}{\mu \overline{F}^{j+k} \left(1-\overline{F}\right)^{l+m-j-k} + (1-\mu) \underline{F}^{j+k} \left(1-\underline{F}\right)^{l+m-j-k}}$$

Let $K = j+k$ then observation of

$$\mu(j,k) = \frac{\mu \overline{F}^{K} \left(1-\overline{F}\right)^{l+m-K}}{\mu \overline{F}^{K} \left(1-\overline{F}\right)^{l+m-K} + (1-\mu) \underline{F}^{K} \left(1-\underline{F}\right)^{l+m-K}}$$

reveals immediately that $\mu(j,k) = \mu(j',k')$ for any $K = j'+k'$. Likewise, $\mu(j,k) < \mu(j',k')$ for any $K < j'+k'$. Since $v(j,k) = \mu(j,k)\overline{V} + (1-\mu(j,k))\underline{V}$, both statements in the lemma follow immediately. ■

Proof of Lemma 2

Lemma 2. *The belief* $\mu_L(K)$ *and the expected payoff* $v_L(K)$ *are a martingale.*

Proof. This amounts to showing that

$$E\left[\mu_{L+1} \mid K\right] = \mu_L(K), \forall K$$

where $E\left[\mu_{L+1} \mid K\right]$ denotes the expectation over μ when an additional signal is observed, given K good signals out of L have already been observed. Then

$$\begin{aligned}
E\left[\mu_{L+1} \mid K\right] &= \frac{p_{L+1}(K+1)}{p_L(K)} \mu_{L+1}(K+1) + \frac{p_{L+1}(K)}{p_L(K)} \mu_{L+1}(K) \\
&= \frac{p_{L+1}(K+1)}{p_L(K)} \frac{\mu \overline{F}^{K+1} \left(1-\overline{F}\right)^{L-K}}{p_{L+1}(K+1)} + \frac{p_{L+1}(K)}{p_L(K)} \frac{\mu \overline{F}^{K} \left(1-\overline{F}\right)^{L-K+1}}{p_{L+1}(K)} \\
&= \overline{F} \mu_L(K) + \left(1-\overline{F}\right) \mu_L(K) = \mu_L(K)
\end{aligned}$$

Now we can immediately show that

$$E\left[v_{L+1} \mid K\right] = v_L(K), \forall K$$

since from the martingale property of $\mu_L(K)$, it follows that

$$\begin{aligned}
E\left[v_{L+1} \mid K\right] &= E\left[\mu_{L+1} \mid K\right] \overline{V} + \left(1 - E\left[\mu_{L+1} \mid K\right]\right) \underline{V} \\
&= \mu_L(K) \overline{V} + (1-\mu_L(K)) \underline{V} \\
&= v_L(K)
\end{aligned}$$

■

Proof of Lemma 4

Lemma 4. *If $l \geq m \geq 0$, then* **I** *is more informative than* **C**.

Proof. We use Blackwell's Theorem: $\mathbf{I} \supset \mathbf{C}$ iff $\mathbf{I} \succ \mathbf{C}$, i.e. $\mathbf{I}$ is more informative than $\mathbf{C}$ iff $\mathbf{I}$ is sufficient for $\mathbf{C}$. The experiment $\mathbf{I}$ is defined to be sufficient for $\mathbf{C}$ $[\mathbf{I} \succ \mathbf{C}]$, when $\mathbf{I} \cdot \mathbf{T} = \mathbf{C}$ for some $(l+1) \times (m+1)$ Markov matrix $\mathbf{T}$, i.e. $\mathbf{c}_{Vk} = \sum_{j=0}^{l} \mathbf{i}_{Vj} \mathbf{t}_{jk}$ for all $V \in \mathcal{V}$ and $k = 0, 1, ..., m$.

First, for the case where $l = m$, the sufficiency condition is trivially satisfied since $\mathbf{I} = \mathbf{C}$. As a result, the square matrix $\mathbf{T}$ is the identity matrix: $\mathbf{t}_{jk} = 1$ if $j = k$ and $\mathbf{t}_{jk} = 0$ otherwise.

For any $m = l - 1$, we can show that $\mathbf{I} \cdot \mathbf{T} = \mathbf{C}$ where $\mathbf{T}_{[l+1,l]} = [\mathbf{t}_{jk}]\,, j \in \{0, ..., l\}\,, k \in \{0, ..., m\}$ with

$$\mathbf{t}_{jk} = \begin{cases} \frac{\binom{m}{k}}{\binom{l}{j}}, & \text{if } j = k \text{ or } j = k+1 \\ 0, & \text{otherwise} \end{cases}$$

To see this, note that the left hand side $\mathbf{I} \cdot \mathbf{T}$ is equivalent, for all $V \in \mathcal{V}$ and $k \in \{0, ..., l-1\}$, to

$$\sum_{p=0}^{l} \mathbf{i}_{Vp} \cdot \mathbf{t}_{pk}$$

and because of the definition of $\mathbf{t}_{jk}$ this is

$$\mathbf{i}_{Vk} \cdot \mathbf{t}_{kk} + \mathbf{i}_{Vk+1} \cdot \mathbf{t}_{k+1k}$$

Now we can substitute for $\mathbf{i}_{Vk}, \mathbf{c}_{Vk}$ and $\mathbf{t}_{jk}$ to get

$$\begin{aligned} & \binom{l}{k} F^k (1-F)^{l-k} \cdot \frac{\binom{m}{k}}{\binom{l}{k}} + \binom{l}{k+1} F^{k+1} (1-F)^{l-k-1} \cdot \frac{\binom{m}{k}}{\binom{l}{k+1}} \\ = \quad & \binom{m}{k} F^k (1-F)^{l-k-1} \end{aligned}$$

Recall that $m = l - 1$ so this is equal to

$$\binom{m}{k} F^k (1-F)^{m-k} = \mathbf{c}_{Vk}$$

for all V and k thus establishing that $\mathbf{I} \cdot \mathbf{T} = \mathbf{C}$.

Now consider the case of any $l > m + 1$. Then we can define the matrix $\mathbf{I}^z$ to be the $2 \times (l + 1 - z)$ matrix for all $z = 0, ..., m-1$ generated by the binomial process from $l - z$ draws. Note that for $z = 0$, $\mathbf{I}^z = \mathbf{I}$. From the result in the first part of the proof, we know that $\mathbf{I} \cdot \mathbf{T}_{[l+1,l]} = \mathbf{I}^1$ and for any z, $\mathbf{I}^z \cdot \mathbf{T}_{[l+1-z,l-z]} = \mathbf{I}^{z+1}$. Now by definition of $\mathbf{I}^z$, for $z = m$, $\mathbf{I}^m = \mathbf{C}$. As a result, there exists an $(l+1) \times (m+1)$ Markov Transformation matrix $\mathbf{T}_{[l+1,m+1]}$ such that $\mathbf{I} \cdot \mathbf{T}_{[l+1,m+1]} = \mathbf{C}$ where

$$\mathbf{T}_{[l+1,m+1]} = \prod_{p=1}^{l-m} \mathbf{T}_{[l-p+1,l-p]}$$

■

Proof of Proposition 5

Proposition 5. *Let* $\frac{v_L(k+1)}{v_L(k)} > \frac{\pi^i}{\pi^{-i}} > \frac{v_L(k)}{v_L(k+1)}$ *for all* k*. Then in the second price auction, an ex-post Nash equilibrium strategy for any player* i *with signal* j *is to bid* $w^i(j) = \pi^i v_L(\min\{2j, L\})$*.*

Proof. Consider the candidate ex-post Nash equilibrium where given realizations of the signals $\widehat{k}, \widehat{j}$, C bids $b_C(\widehat{k}) = \pi^C v_L(\lambda_C \cdot \widehat{k}), \forall \widehat{k} \leq m$ and I bids $b_I(\widehat{j}) = \pi^I v_L(\min\left\{\lambda_I \cdot \widehat{j}, L\right\}), \forall \widehat{j} \leq l$. We show that $\lambda_C = \lambda_I = 2$ is an equilibrium provided π^C and π^I are not too different. Given realizations of the other bidder's signal k (or j), an ex-post Nash equilibrium requires that each bidder bids her valuation

$$
\begin{aligned}
b_C(\widehat{k}) &= \pi^C v_L(\widehat{k} + j) \\
b_I(\widehat{j}) &= \pi^I v_L(\widehat{j} + k)
\end{aligned}
$$

or

$$
\begin{aligned}
\pi^C v_L(\lambda_C \cdot \widehat{k}) &= \pi^C v_L(\widehat{k} + j) \\
\pi^I v_L(\min\left\{\lambda_I \cdot \widehat{j}, L\right\}) &= \pi^I v_L(\widehat{j} + k).
\end{aligned}
$$

A fixed point where $j = \widehat{j}$ and $k = \widehat{k}$ implies (with v_L an increasing function)

$$
\begin{aligned}
\lambda_C \cdot \widehat{k} &= \widehat{k} + \widehat{j} \\
\min\left\{\lambda_I \cdot \widehat{j}, L\right\} &= \widehat{j} + \widehat{k}.
\end{aligned}
$$

and therefore, for $\lambda_I \cdot \widehat{j} \leq L$ (for $\lambda_I \cdot \widehat{j} > L$, the equality holds trivially)

$$
\begin{aligned}
\lambda_C \cdot \widehat{k} &= \widehat{k} + \widehat{j} \\
\lambda_I \cdot \widehat{j} &= \widehat{j} + \widehat{k}.
\end{aligned}
$$

which is satisfied for $\lambda_C = \lambda_I = 2$.

Now consider the condition $\frac{v_L(k+1)}{v_L(k)} > \frac{\pi^i}{\pi^{-i}} > \frac{v_L(k)}{v_L(k+1)}$ and analyze the case where $1 \geq \frac{\pi^i}{\pi^{-i}}$ (the other case follows analogously). Then for any k $\pi^{-i} v_L(k) < \pi^i v_L(k+1)$ the high productivity firm π^{-i} does not derive higher value from hiring the worker while receiving fewer signals and conditional on the signals, no firm has an incentive to deviate from the equilibrium strategy. ■

Proof of Proposition 6

Proposition 6. *Consider a sequence* $\left\{(\overline{F}, 1 - \underline{F})\right\}_n$ *converging to* $(1, 1)$ *and a corresponding sequence* $\left\{v_L(K)\right\}_n$*. Then,* $\forall l > 3m$*, the equilibrium bid of* C *is* $v_L(2k) = \underline{V}$*, irrespective of the cardinality of good signals* k *received by* C*.*

In order to prove the result, we make use of the following Lemma

Lemma 12 *Consider a sequence* $\left\{\left(\overline{F},1-\underline{F}\right)\right\}_n$ *converging to* $(1,1)$ *and a corresponding posterior belief* $\{\mu_L(K)\}_n$ *and value* $v_L(K)$*, then*

$$v_L(K)=\begin{cases}\underline{V} & \text{, if } K<\frac{L}{2}\\ \mu\overline{V}+(1-\mu)\underline{V} & \text{, if } K=\frac{L}{2}\\ \overline{V} & \text{, if } K>\frac{L}{2}\end{cases}$$

Proof. From the definition of $\mu_L(K)$ we have that

$$\mu_L(K)=\frac{\mu\overline{F}^K\left(1-\overline{F}\right)^{L-K}}{\mu\overline{F}^K\left(1-\overline{F}\right)^{L-K}+(1-\mu)\underline{F}^K\left(1-\underline{F}\right)^{L-K}}.$$

where

$$\lim_{\left(\overline{F},1-\underline{F}\right)\to(1,1)}\mu_L(K)=\frac{0}{0}.$$

Dividing by $\overline{F}^K\left(1-\overline{F}\right)^{L-K}$, we get:

$$\mu_L(K)=\frac{\mu}{\mu+(1-\mu)\frac{\underline{F}^K\left(1-\underline{F}\right)^{L-K}}{\overline{F}^K\left(1-\overline{F}\right)^{L-K}}}.$$

Using the fact that in the limit, $r=\log\overline{F}/\log(1-\underline{F})=1$, and therefore $\overline{F}=(1-\underline{F})$ We can rewrite as

$$\mu_L(K)=\frac{\mu}{\mu+(1-\mu)\left(\frac{\overline{F}}{1-\overline{F}}\right)^{L-2K}}$$

and for $K<\frac{L}{2}$, $\lim_{\left(\overline{F},1-\underline{F}\right)\to(1,1)}\mu_L(K)=0$. And from the definition $v_L(K)=\mu_L(K)\overline{V}+(1-\mu_L(K))\underline{V}$, the limit $\lim_{\left(\overline{F},1-\underline{F}\right)\to(1,1)}v_L(K)=\underline{V}$. Likewise, if $K=\frac{L}{2}$, $\lim_{\left(\overline{F},1-\underline{F}\right)\to(1,1)}\mu_L(K)=\mu$ and $\lim_{\left(\overline{F},1-\underline{F}\right)\to(1,1)}v_L(K)=\mu\overline{V}+(1-\mu)\underline{V}$ and if $K>\frac{L}{2}$, $\lim_{\left(\overline{F},1-\underline{F}\right)\to(1,1)}\mu_L(K)=1$ and $\lim_{\left(\overline{F},1-\underline{F}\right)\to(1,1)}v_L(K)=\overline{V}$. ■

Now the proof of the Proposition follows:

Proof. From Lemma 12, $\lim_{\left(\overline{F},1-\underline{F}\right)\to(1,1)}v_L(2m)=0$ provided $2m<\frac{L}{2}$. Since $L=m+l$, it follows that $\lim_{\left(\overline{F},1-\underline{F}\right)\to(1,1)}v_L(2m)=0$ provided $l>3m$. From Lemma 1, for any $K\leq 2m$ beliefs satisfy $\mu_L(K)\leq\mu_L(2m)$ and as a result, for any signal $k\in\{0,1,...,m\}$, and given an equilibrium bid $v_L(2k)=\mu_L(2k)\overline{V}+(1-\mu_L(2k))\underline{V}$, the limit of the equilibrium bid $\lim_{\left(\overline{F},1-\underline{F}\right)\to(1,1)}v_L(2k)=\underline{V}$ for any $k\leq m$ provided $l>3m$. ■

Proof of Proposition 7

Proposition 7 *(Turnover and Differential Information) Consider any pair* l,m *such that* $l+m=L$ *and* $l\geq m$*. Then turnover decreases the more information is differential: turnover given* l,m *is greater than turnover given* l',m' $(l'+m'=L)$ *or all* $l'>l$*.*

Proof. Turnover is defined as $T(l,m)=\Pr\{k>j\mid l,m\}\cap\frac{1}{2}\Pr\{k=j\mid l,m\}$. Then

$$\begin{aligned}T(l,m) &= \sum_{j=0}^{m-1}\sum_{k=j+1}^{m}p(j,k)+\frac{1}{2}\sum_{j=0}^{m}p(j,j)\\ &= \sum_{j=0}^{m-1}\sum_{k=j+1}^{m}\binom{l}{j}\binom{m}{k}p_L(j+k)+\frac{1}{2}\sum_{j=0}^{m}\binom{l}{j}\binom{m}{j}p_L(2j)\end{aligned}$$

for any l, m. Now consider any l, m and l', m' such that $l' = l + 1$ and $m' = m - 1$. Then turnover is greater for l, m than for l', m'

$$T(l,m) - T(l',m') = \sum_{j=0}^{m-1} p_L(j,m) + \frac{1}{2} p_L(m,m) > 0.$$

since$\binom{l}{j}\binom{m}{k} \geq \binom{l'}{j}\binom{m'}{k}$ for any $k \geq j$. By induction, this holds for any l', m' where $l' > l$. ■

Proof of Proposition 8

Proposition 8 *(Wages and Differential Information) Consider any pair l, m such that $l + m = L$ and $l \geq m$. Then expected wages decrease the more information is differential: expected wages given l, m are higher than expected wages given l', m' ($l' + m' = L$) for all $l' > l$.*

Proof. The expected wage $w(l, m)$ is given by

$$w(l,m) = \sum_{j=0}^{m} \left(\sum_{k=j}^{m} p_L(j,k) v_L(j,j) + \sum_{k=0}^{j-1} p_L(j,k) v_L(k,k) \right) + \sum_{j=m+1}^{l} \sum_{k=0}^{m} p_L(j,k) v_L(k,k)$$

for any l, m. Now consider any l, m and l', m' such that $l' = l + 1$ and $m' = m - 1$. Then expected wages are higher for l, m than for l', m'

$$w(l,m) - w(l',m') = \sum_{j=0}^{m} p_L(j,m) v_L(j,j) > 0$$

■

Proof of Proposition 10

Proposition 10. *Let $l \geq m > 1$. Then under the information structure S there is no full extraction of surplus by the employee.*

Proof. First, we establish some properties on the rank of the matrix $P(l, m)$.

Define $P_1(l, m) = P(l, m \mid \mu = 1)$ with each entry in the matrix denoted by $p_1(j, k)$ and $P_2(l, m) = P(l, m \mid \mu = 0)$ with $p_2(j, k)$. Therefore we can write $P(l, m)$ as

$$\begin{aligned} P(l,m) &= \mu P_1(l,m) + (1-\mu) P_2(l,m) \\ &= \mu \left(\binom{l}{s^I} \binom{m}{s^C} \overline{F}^{s^I + s^C} \left(1 - \overline{F}\right)^{l+m-s^I-s^C} \right) + (1-\mu) \left(\binom{l}{s^I} \binom{m}{s^C} \underline{F}^{l+m-s^I-s^C} \left(1 - \underline{F}\right)^{s^I+s^C} \right). \end{aligned}$$

First, we show that the rank $r\left[P_i(l,m)\right] = 1$. Without loss of generality, we derive the rank for P_1. Consider two columns k and k'. Any element in row j and column k can be written as

$$\begin{aligned} p_i(j,k) &= \binom{l}{j} \binom{m}{k} \overline{F}^{j+k} \left(1 - \overline{F}\right)^{l+m-j-k} \\ &= \binom{m}{k} \binom{m}{k'}^{-1} \overline{F}^{k-k'} \left(1 - \overline{F}\right)^{k'-k} \left[\binom{l}{j} \binom{m}{k'} \overline{F}^{j+k'} \left(1 - \overline{F}\right)^{l+m-j-k'} \right] \\ &= C(k,k') \cdot p_i(j,k') \end{aligned}$$

where $C(k,k') = \binom{m}{k}\binom{m}{k'}^{-1}\overline{F}^{k-k'}\left(1-\overline{F}\right)^{k'-k}$ is independent of j. Therefore, columns k and k' are linearly dependent. Since this is true for any k', all columns are linearly dependent. It immediately follows that the rank $r\left[P_i(l,m)\right] = 1$.

Now we show that $r\left[P\left(l,m\right)\right] < 3$. Therefore, consider any 3×3 minor of $P(l,m)$ with $m > 1$ denoted by P^{3M} and let the rows and columns be written as j, j', j'' and k, k', k'' respectively. We show that $\det\left[P^{3M}\right] = 0$. Let $p^{3M}(k)$ denote the k-th column, and $p_1^{3M}(k) = p^{3M}(k \mid \mu = 1)$ and $p_2^{3M}(k) = p^{3M}(k \mid \mu = 0)$. Write $\det\left[P^{3M}\right]$ as

$$\det\left[P^{3M}\right] = \det\left[\mu p_1^{3M}\left(k\right) + \left(1-\mu\right)p_2^{3M}\left(k\right), \mu p_1^{3M}\left(k'\right) + \left(1-\mu\right)p_2^{3M}\left(k'\right), \mu p_1^{3M}\left(k''\right) + \left(1-\mu\right)p_2^{3M}\left(k''\right)\right].$$

Using the property that the determinant of a matrix in which each element in the column of the determinant is a sum of two summands, then the determinant expands into the sum of two determinants, we can expand $\det\left[P\left(l,l\right)\right]$ as follows

$$\begin{aligned}
\det\left[P^{3M}\right] &= \det\left[\mu p_1^{3M}\left(k\right) + \left(1-\mu\right)p_2^{3M}\left(k\right), \mu p_1^{3M}\left(k'\right) + \left(1-\mu\right)p_2^{3M}\left(k'\right), \mu p_1^{3M}\left(k''\right) + \left(1-\mu\right)p_2^{3M}\left(k''\right)\right] \\
&= \det\left[\mu p_1^{3M}\left(k\right), \mu p_1^{3M}\left(k'\right) + \left(1-\mu\right)p_2^{3M}\left(k'\right), \mu p_1^{3M}\left(k''\right) + \left(1-\mu\right)p_2^{3M}\left(k''\right)\right] \\
&\quad + \det\left[\left(1-\mu\right)p_2^{3M}\left(k\right), \mu p_1^{3M}\left(k'\right) + \left(1-\mu\right)p_2^{3M}\left(k'\right), \mu p_1^{3M}\left(k''\right) + \left(1-\mu\right)p_2^{3M}\left(k''\right)\right]
\end{aligned}$$

and continuing expanding for each of the columns, we can write $\det\left[P^{3M}\right]$ as the sum of 2^3 determinants

$$\begin{aligned}
\det\left[P^{3M}\right] &= \det\left[\mu p_1^{3M}\left(k\right), \mu p_1^{3M}\left(k'\right), \mu p_1^{3M}\left(k''\right)\right] \\
&\quad + \det\left[\mu p_1^{3M}\left(k\right), \mu p_1^{3M}\left(k'\right), \left(1-\mu\right)p_2^{3M}\left(k''\right)\right] \\
&\quad + \det\left[\mu p_1^{3M}\left(k\right), \left(1-\mu\right)\mu p_2^{3M}\left(k'\right), \mu p_1^{3M}\left(k''\right)\right] \\
&\quad + \det\left[\mu p_1^{3M}\left(k\right), \left(1-\mu\right)\mu p_2^{3M}\left(k'\right), \left(1-\mu\right)p_2\left(k''\right)\right] \\
&\quad + \det\left[(1-\mu)p_2^{3M}\left(k\right), \mu p_1^{3M}\left(k'\right), \mu p_1^{3M}\left(k''\right)\right] \\
&\quad + \det\left[(1-\mu)p_2^{3M}\left(k\right), \mu p_1^{3M}\left(k'\right), \left(1-\mu\right)p_2^{3M}\left(k''\right)\right] \\
&\quad + \det\left[(1-\mu)p_2^{3M}\left(k\right), \left(1-\mu\right)\mu p_2^{3M}\left(k'\right), \mu p_1^{3M}\left(k''\right)\right] \\
&\quad + \det\left[(1-\mu)p_2^{3M}\left(k\right), \left(1-\mu\right)\mu p_2^{3M}\left(k'\right), \left(1-\mu\right)p_2^{3M}\left(k''\right)\right].
\end{aligned}$$

All of these determinants have at least 2 columns that correspond to parts of columns of at least one of the constituent matrices P_i. Since $r\left[P_i(l,m)\right] = 1$, each of these determinants is zero. (Notice that this is not the case when for 2×2 minors as there will be some determinant with one column corresponding to P_1 and the other to P_2. For example, when $l = m = 1$ the only 2×2 minor is the entire matrix. It is easily verified that $\det\left[P\left(1,1\right)\right] = \mu\left(1-\mu\right)\left(\overline{F} - \underline{F}\right)^2 \neq 0$ for any $\overline{F} \neq \frac{1}{2}$ and $\underline{F} \neq \frac{1}{2}$.) If all 3×3 minors have a zero determinant, it follows that the rank of the matrix $r\left[P\left(l,m\right)\right] < 3$.

We now first show that there is no full extraction of the surplus by a dominant strategy auction. For that purpose, we verify that the conditions in Theorems 1 (Crémer and McLean (1988)) are violated.

For the condition in Theorem 1, we verify that there exists a $\left\{\theta^i\left(s^i\right)\right\}$ for $i \in \{I, C\}$ such that:

$$\sum_{s^i} \theta^i\left(s^i\right) p\left(s^{-i} \mid s^i\right) = 0$$

for all s^{-i}. This is equivalent to:

$$\begin{aligned} \sum_{s^i} \theta^i\left(s^i\right) p\left(s^i\right) p\left(s^i, s^{-i}\right) &= 0 \\ \sum_{s^i} \theta_2^i\left(s^i\right) p\left(s^i, s^{-i}\right) &= 0 \end{aligned} \tag{1}$$

where $\theta_1^i\left(s^i\right)=\theta\left(s^i\right) p\left(s^i\right)$. To show that condition (1) holds, it is sufficient to show that the matrix $P\left(l, m\right)=\left[p(s^I, s^C)\right]$ has rank $r[P(l, m)]<l+1$. Since $l>1$, and above we have shown that $r[P(l, m)]<3$, it immediately follows that this rank condition is always satisfied. Therefore, there is no full rent extraction using dominant strategy mechanisms. Observe that for any asymmetric matrix P (where $l \neq m$) it is automatic that rank $r\left[P\left(l, m\right)\right]<l+1$, since for any asymmetric $(l+1) \times(m+1)$ matrix, the rank satisfies $r\left[P\left(l, m\right)\right] \leq \min \{l+1, m+1\}$.

We next verify that there is no full rent extraction using a Bayesian auction. Crémer and McLean (1988) establish that there will be no full rent extraction if there does not exist lotteries with zero expected value conditional on the true type. Using Farkas's Lemma, they show that a solutions satisfying these requirements exists if and only if the dual of this system has no solution. We therefore show that the conditions in Theorems 2 (Crémer and McLean (1988)) are violated. This amounts to showing that for $i \in\{I, C\}$, there exists a signal $s^i \in S^i$ and a family $\left\{\theta^i\left(t^i\right)\right\}_{t^i \in S^i \backslash s^i}$ such that the following two requirements hold:

1. $$\theta^i\left(t^i\right) \geq 0 \tag{2}$$

   for all $t^i \in S^i \backslash s^i$; and

2. $$p\left(s^{-i} \mid s^i\right)=\sum_{t^i \neq s^i} \theta^i\left(t^i\right) p\left(s^{-i} \mid t^i\right)$$

   for all $s^{-i} \in S^{-i}$.

The second requirement is equivalent to:

$$\begin{aligned} p\left(s^i, s^{-i}\right) p\left(s^i\right) &= \sum_{t^i \neq s^i} \theta^i\left(t^i\right) p\left(t^i\right) p\left(t^i, s^{-i}\right) \\ p\left(s^i, s^{-i}\right) &= \sum_{t^i \neq s^i} \widetilde{\theta}^i\left(t^i\right) p\left(t^i, s^{-i}\right) \end{aligned} \tag{3}$$

for all $s^{-i} \in S^{-i}$, and where

$$\widetilde{\theta}^i\left(t^i\right)=\frac{\theta^i\left(t^i\right) p\left(t^i\right)}{p\left(s^i\right)} \geq 0 \tag{4}$$

for $t^i \neq s^i$. Now we can write (3) in matrix form for each firm C (the same argument applies for firm I) as

$$\underset{[l \times m]}{\left[p\left(t^i, s^{-i}\right)\right]} \cdot \underset{[m \times 1]}{\left[\widetilde{\theta}^i\left(t^i\right)\right]}=\underset{[l \times 1]}{\left[p\left(s^i, s^{-i}\right)\right]}$$

for $t^i \neq s^i$. Rewriting including $t^i = s^i$ we get

$$\underset{[(l+1)\times(m+1)]}{\left[p\left(t^i, s^{-i}\right)\right]} \cdot \underset{[(m+1)\times 1]}{\left[\widetilde{\theta}^i\left(t^i\right), -1\right]} = \underset{[(l+1)\times 1]}{[0]} \tag{5}$$

where all but one $\widetilde{\theta}^i\left(t^i\right)$ have the same sign, and the other has the opposite sign. The solution to (5) lies in the orthogonal complement of the row space. Since $r[P(l,m)] = r \leq 2$, the solution space to (5) is of dimension $m + 1 - r < m - 1$. Because p is non-negative, at least one coefficient must have the opposite sign of the other coefficients. Given the dimension of the solution space is less than $m - 1$, there will therefore always exist a solution to (5) that also satisfies (2). ■

References

[1] Acemoglu, Daron, and Robert Shimer, "Wage and Technology Dispersion", *Review of Economic Studies* **67**, 2000, 585-608.

[2] Burdett, Ken, and Melvyn G. Coles, "Equilibrium Wage-Tenure Contracts", *Econometrica* **71(5)**, 2003, 1377-1404.

[3] Burdett, Ken, and Dale Mortensen, "Wage Differentials, Employer Size, and Unemployment", *International Economic Review* **39(3)**, 1998, 257-273.

[4] Crémer, Jacques, and Richard McLean, "Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions", *Econometrica* **56(6)**, 1988, 1247-1257.

[5] Felli, Leonardo, and Christopher Harris, "Learning, Wage Dynamics, and Firm-Specific Human Capital", *Journal of Political Economy* **104(4)**, 1996, 838-868.

[6] Greenwald, Bruce, "Adverse Selection in the Labour Market", *Review of Economic Studies* **53**, 1986, 325-347.

[7] Harris, Milton, and Bengt Holmström, "A Theory of Wage Dynamics", *Review of Economics Studies* **XLIX**, 1982, 315-333.

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

[9] Kambourov, Gueorgui, and Iourii Manovskii, "Occupational Mobility and Wage Inequality", UPenn mimeo, 2004.

[10] Krueger, Dirk, and Harald Uhlig, "Competitive Risk Sharing Contracts with One-Sided Commitment", NBER Working Paper **10135**, 2003.

[11] Milgrom, Paul, and Robert J. Weber, "A Theory of Auctions and Competitive Bidding", *Econometrica* **50(5)**, 1982, 1089-1122.

[12] Moscarini, Giuseppe, "Excess Worker Reallocation", *Review of Economic Studies*, 2001, 593-612

[13] Moscarini, Giuseppe, "Job Matching and the Wage Distribution", *Econometrica* **73(2)**, March 2005, 481-516.

[14] Postel-Vinay, Fabien, and Jean-Marc Robin, "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity", *Econometrica* **70(6)**, 2002, 2295-2350.

[15] Wolinski, Asher, "Information Revelation in a Market with Pairwise Meetings", *Econometrica* **58(1)**, 1990, 1-23.