BLS WORKING PAPERS

U.S. DEPARTMENT OF LABOR
Bureau of Labor Statistics

OFFICE OF EMPLOYMENT AND UNEMPLOYMENT STATISTICS

A Model of Asymmetric Employer Learning with Testable Implications

Joshua C. Pinkston, U.S. Bureau of Labor Statistics

Working Paper 365
January 2003

A Model of Asymmetric Employer Learning With Testable Implications*

Joshua C. Pinkston
Bureau of Labor Statistics

December, 2002

Abstract: This paper develops and tests a unique model asymmetric employer learning. The previous literature on asymmetric learning assumes that a worker's employer is perfectly informed while outside firms possess only public information. This paper relaxes that assumption, allowing firms to profitably bid for employed workers under conditions that were not profitable in previous models. The model in this paper is the first in the literature to predict either wage growth without promotions or mobility between firms without firm- or match-specific productivity. The bidding through which firms compete for a worker produces a sequence of wages that converges to the current employer's conditional expectation of the worker's productivity. This convergence of wages allows the model to be tested using an extension of existing work on employer learning. Wage regressions estimated on a sample of men from the NLSY produce strong evidence of asymmetric learning.

JEL: J3, D82, D83, D44

Keywords: Asymmetric Employer Learning, Wage and Performance Relationship, Applications of Auction Theory

---

* I would like to thank Joseph Altonji, Todd Elder, Mark Loewenstein, Dale Mortensen, Jim Spletzer, Jay Stewart, Christopher Taber, Charles Zheng and seminar participants at the Bureau of Labor Statistics for helpful advice and comments. All mistakes and opinions are my own.
Mailing address: Bureau of Labor Statistics, 2 Massachusetts Ave., NE, Suite 4945, Washington, D.C. 20212. Phone: (202) 691-7403. Fax: (202) 691-6425. Email: Pinkston_J@bls.gov.

This paper develops and tests a unique model of asymmetric employer learning. Firms that encounter a worker receive a private signal of the worker's productivity, but the worker's current employer accumulates more private information than outside firms with the informational advantage increasing with the duration of the worker's current employment spell. Outside firms can profitably bid for an employed worker despite the current employer's informational advantage, even when the worker is equally productive in any firm. The bidding process, which is described using standard results from the auction literature, and the assumption that each worker is equally productive at any firm imply that the winning firm learns the private information of the firm it outbid. As a result, the current employer's private information becomes more precise with the length of the worker's current employment spell, not with tenure as one might expect. More importantly, wages in this model converge to the employer's expectation as the employment spell increases in length due to competition from less well-informed firms.

This convergence of wages to the current employer's expectation allows this model to be tested empirically with a simple extension of the work of Altonji and Pierret (2000) and Farber and Gibbons (1996) on public learning. Essentially, employer learning implies an increase in the coefficients in wage regressions on variables that are correlated with productivity but difficult for employers to observe, such as AFQT scores. The model developed in this paper implies that employer learning that occurs publicly is reflected as learning with experience in the labor market, while employer learning that occurs privately is reflected as learning over the current employment spell. This last implication results from the convergence of wages to the current employer's expectations.

Both the model and the empirical test in this paper make important contributions to the literature on asymmetric learning. The contribution of the test is more obvious since this literature is marked by a large gap between theory and evidence. The estimation results presented in this paper provide strong evidence of asymmetric employer learning using a sample of men from the NLSY.

The model makes important theoretical contributions by relaxing the informational assumptions of earlier models and overcoming some of their shortcomings as a result. The previous literature, most of which was concerned with the relationship of wages to task assignment (e.g., Waldman, 1984; Bernhardt, 1995), assumed the current employer was perfectly informed, resulting in some unrealistic implications. Gibbons and Waldman (1999) criticize this literature for producing no wage growth without promotions. Previous models also produce no mobility between jobs without firm- or match-specific productivity because it is otherwise never profitable for outside firms to bid for a worker against a perfectly informed employer. The previous literature is generally unable to produce any competition for employed workers without assumming match-specific productivity. In contrast, my model predicts both wage growth without promotions and mobility between jobs despite assuming that the worker is equally productive at all firms[1].

The next section briefly discusses the previous literature on asymmetric information. Section 2 introduces employer learning in a simpler public information setting. Section 3 then develops the model of asymmetric employer learning. In Section 4, I describe the wage regressions that provide the test of my model. Section 5 describes the data I use and Section 6 presents the estimation results. Finally, Section 7 concludes and discusses future research.

[1] As will be discussed later, this assumption is not necessary for most results of my model to follow. The one result that would change is the convergence of wages to employer expectations, which would occur over tenure under match-specific productivity. Extending the model to allow, match-specific productivity, however, will be left to future research.

1. Previous Literature on Asymmetric Information

The literature on asymmetric information between firms has been largely theoretical, focusing primarily on the relationship of wages to task assignment. The motivation for much of this work has been studies of personnel records that suggest wages are tied to jobs and increase with seniority even when productivity does not. (See Medoff and Abraham, 1980, 1981; and Baker, Gibbs and Holmstrom 1994a, 1994b.) Despite the literature's empirical motivation, only one previous paper, Gibbons and Katz (1991), has developed and tested empirical implications of an asymmetric information model, leaving a large gap between theory and evidence in this literature.

Waldman (1984) develops the basic asymmetric information model. He assumes that only the worker's current employer receives direct information about the worker's ability. Outside firms learn about the worker only by observing her task assignment at the current firm, and the current employer becomes perfectly informed after one period of tenure. Wages never rise unless a promotion signals higher ability to outside firms, resulting in wages being more closely tied to jobs than to ability. Workers are inefficiently assigned to jobs as firms determine assignment strategically.

Several other papers expand upon the basic idea of Waldman (1984). For example, Milgrom and Oster (1987) develop a model of labor market discrimination in a setting where a promotion makes a worker completely "visible" to outside firms. Bernhardt (1995) focuses on promotion "fast tracks" in which workers who are promoted quickly are more likely to be promoted again than are their peers in the same job. Scoones and Bernhardt (1998) argue that this asymmetric learning environment, combined with match-specific productivity, can lead to an inefficiently large investment in firm-specific human capital.

In general, these models assume that outside firms possess no information that the current employer does not also possess. Typically, the current employer is perfectly informed and other firms receive information only by observing the employer's actions. This implies that outside firms cannot profitably bid for workers in the absence of match-specific productivity. As a result, there is no mobility between jobs without match- (or firm-) specific productivity and no wage growth without promotions. This last point is the major criticism that Gibbons and Waldman (1999) make of this literature.

The model developed in this paper relaxes the informational assumptions of earlier models by assuming that the worker's employer only becomes perfectly informed in the limit, as the length of the current employment spell approaches infinity; and by allowing outside firms to receive noisy signals from interviews or other firm-specific evaluations. These generalizations have the significant effect of allowing outside firms to profitably bid for employed workers even when there is no match-specific productivity. As a result, wages grow to reflect the employer's private learning even when there are no promotions and workers are bid away from their employers by less well-informed firms.

Gibbons and Katz (1991) develop and test a model of layoffs under asymmetric information. In their model, employers have discretion over whom they lay off, as well as information about the worker that the market does not. As a result, layoffs signal to the market that the worker is of lower ability. Because displacement by a plant closing should not contain the same negative signal, workers who are laid off are compared to those who are displaced by a plant closing to control for the effects of displacement. Their estimation using CPS data supports all of these predictions; however, their results could simply reflect selection on unobservables determining who gets laid off.

The current paper helps bridge the gap between theory and evidence in this literature. The model of asymmetric learning that is developed in Section 3 implies that wage equations will reflect evidence of asymmetric employer learning as the time since the last period of nonemployment increases, in addition to evidence of market learning with experience. This implication is tested using data from the NLSY, and the estimation results support the existence of asymmetric learning.

2. The Public Learning Benchmark

This section introduces employer learning by considering learning when all information about the worker is public. The next section will use many of the same terms and concepts when modelling asymmetric employer learning. Let μ_i denote worker i's actual productivity and assume that productivity has a normal distribution: $\mu \sim N\left(m, \sigma_\mu^2\right)$. Assume that when a worker enters the labor market firms observe an initial signal of productivity:

$$s_{0i} = \mu_i + \varepsilon_{0i},$$

where $\varepsilon_{0i} \sim N\left(0, \sigma_\varepsilon^2\right)$.

With all firms in the market having the same information, the worker's initial wage equals initial expected productivity:

$$w_{0i} = E\left(\mu | s_{0i}\right) = \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + \sigma_\mu^2} \cdot m + \frac{\sigma_\mu^2}{\sigma_\varepsilon^2 + \sigma_\mu^2} \cdot s_{0i} \tag{1}$$

(See DeGroot, 1970.) It is easy to see that the less reliable the initial signal, s_{0i}, is (the higher σ_ε^2 is), the more initial wages are based on m, the population mean, relative to the initial signal.

Assume now that in all later periods of experience, $x = 1, 2, ..,$ the market observes an additional signal of productivity for each worker:

$$s_{xi} = \mu_i + \psi_{xi},$$

where $\psi_{xi} \sim N\left(0, \sigma_\psi^2\right)$. The variance of the error term ψ_{xi} is assumed not to vary with time. Essentially, the market collects information at the same rate for all workers once those workers enter the labor market.

Combining the initial signal, s_0, and all of the later signals, s_x, the market forms an updated signal in each period of the form (dropping the i subscripts for convenience)

$$\begin{aligned} S_x &= \frac{\sigma_\psi^2}{x\sigma_\varepsilon^2 + \sigma_\psi^2} \cdot s_0 + \sum_{\tau=1}^{x} \frac{\sigma_\varepsilon^2}{x\sigma_\varepsilon^2 + \sigma_\psi^2} \cdot s_\tau \\ &= \mu + \eta_x \end{aligned}$$

for each level of experience x, where $\eta_x \sim N\left(0, \sigma_x^2\right)$, and $\sigma_x^2 = \frac{\sigma_\varepsilon^2 \sigma_\psi^2}{x\sigma_\varepsilon^2 + \sigma_\psi^2}$.

Wages at experience level x equal the market's expectation of productivity conditional on S_x, which is simply

$$\begin{aligned} E\left(\mu | S_x\right) &= \frac{\sigma_x^2}{\sigma_x^2 + \sigma_\mu^2} \cdot m + \frac{\sigma_\mu^2}{\sigma_x^2 + \sigma_\mu^2} \cdot S_x \\ &= \alpha_{mx} m + \alpha_{sx} S_x. \end{aligned} \qquad (2)$$

It is a matter of simple algebra to show that $E(\mu|S_x) = E(\mu|s_0, s_1, s_2, .., s_x)$, but the updated market signal, S_x, is used in this paper for simplicity.

The predictions of employer learning for equation (2) are driven by the variance of the updated signal, σ_x^2. As the updated signal becomes more reliable (σ_x^2 falls), α_{sx} increases and α_{mx} decreases. The coefficient α_{sx}, therefore, increases with experience, while α_{mx} decreases, because the updated signal becomes more reliable with experience $\left(\frac{\partial \sigma_x^2}{\partial x} = \frac{-\left(\sigma_\varepsilon^2\right)^2 \sigma_\psi^2}{\left(x\sigma_\varepsilon^2 + \sigma_\psi^2\right)^2} < 0\right)$. In other words, wages become increasingly based on the market's information about an individual worker relative to the population mean as that worker's experience increases.

3. The Asymmetric Learning Model

In the previous section all information about a worker's productivity was publicly observed and summarized by the updated market signal, S_x, which became more reliable with experience. In this section, the assumptions of the public learning model are relaxed. Firms observe private signals of a worker's productivity and the worker's current employer accumulates more private information the longer the worker is continuously employed, gaining an increasing informational advantage over outside firms.

I model wages in each period using a bidding process that can be seen as a series of bidding wars (English auctions) with two bidders. One firm offers the worker a wage, another firm makes a counter-offer, and so on until one firm drops out. The remaining firm then hires the worker at the wage where the losing firm dropped out. This bidding process allows me to exploit standard results from the auction literature[2] and provides a convenient way of describing bidding under

[2] See McAffee and McMillan (1987) and Klemperer (1999) for readable surveys of this literature.

asymmetric information. Finally, after the process of wage determination is modelled, I describe the private learning of the current employer and show that the sequence of wages produced by the model converges to the expectation of the increasingly well informed employer.

3.1. Bidding for an Unemployed Worker

The sequence of events considered in this section begins in period 0, when the worker is unemployed. The period begins with two firms encountering the worker. Each of these firms receives a private initial signal of the worker's productivity (e.g., from an interview), in addition to observing the updated market signal, S_x. The two firms then engage in a bidding war for the worker's services and the winning firm hires the worker at the highest wage offered by the losing firm[3]. As a result of observing the wage at which the losing firm dropped out, the firm that hires the worker also learns the initial signal of the losing firm.

Specifically, assume that any firm f that encounters worker i for the first time receives a private signal, ν_{fi}, from an interview or other evaluation:

$$\nu_{fi} = \mu_i + e_{fi}, \tag{3}$$

where $e_{fi} \sim N\left(0, \sigma_\nu^2\right)$. For simplicity, the worker is assumed not to encounter a firm she has worked for or been interviewed by in the past.

Assume now that an unemployed worker solicits per period wage offers from two firms, each of which collects a signal, ν_{fi}, $f = 0, 1$. Assume firms compete for a worker by engaging in

[3] The results would be similar (but more complicated) if more than two firms bid for the worker at the same time. The difference would be that the firms with the highest signals would bid for the worker after gaining additional information about the worker by observing the wages at which the lower firms dropped out (Milgrom and Weber, 1982).

a bidding war or English (open, ascending bid) auction. Since the worker's productivity is assumed to be the same in all firms this is a common-value auction.

As shown by Milgrom and Weber (1982) (MW from here on), this bidding war is strategically equivalent to a second-price auction: The firm with the highest signal wins the bidding and pays the worker a wage equal to the highest bid the losing bidder was willing to make[4]. This result might initially seem to be an unusual description of wage determination, but viewing the auction as a series of offers and counter-offers should make it more intuitive. The worker still chooses employment with the firm that placed the highest bid, but that bid equals the optimal bid of the losing firm. In other words, the winner simply pays the worker's outside option[5].

MW show that an equilibrium bid for each firm in this auction is the expectation of productivity conditional on the signal that it receives, ν_{fi}, and the signal of the other firm being the same, or

$$b(\nu_{fi}) = E(\mu | S_{xi}, \nu_{fi}, \nu_{ki} = \nu_{fi})$$

where $f, k = 0, 1$ and $f \neq k$[6]. In other words, $b(\nu_{fi})$ is the highest wage firm f is willing to bid for worker i. Appendix A discusses the assumptions that are necessary for this equilibrium bidding strategy to hold, as well as how those assumptions nest mine. Finally, note that I assume firms act as though the expected value of future auctions does not impact their optimal bids[7].

[4] MW actually develop a general auction framework that includes any mix of private- and common-value auctions, allowing the results of this subsection and the next to follow even if productivity had match-specific or firm-specific elements.

[5] A recent paper by Postel-Vinay and Robin (2000) uses a somewhat similar framework in an equilibrium search model. In their case, the bidding is similar to a private-value English auction since firms differ in productivity. Their work also differs in that it involves no employer learning or asymmetric information.

[6] This is a special case of Thereom 6 in MW. See Appendix A.

[7] This assumption does not affect the results since the *ex post* expected discounted present value of the worker's services is the same regardless of which firm wins. Furthermore, it simplifies my analysis by allowing me to ignore the discounted value of future expected rents, which will vary with employment spell duration and experience.

Assume, without loss of generality, that $\nu_{0i} > \nu_{1i}$. Firm 0 wins the bidding and the resulting wage can be written as (dropping the i subscripts)

$$\begin{aligned} w_1 &= E\left(\mu|S_x, \nu_0 = \nu_1, \nu_1\right) \\ &= \frac{\sigma_\nu^2}{\sigma_\nu^2 + 2V_x} \cdot E\left(\mu|S_x\right) + 2 \cdot \frac{V_x}{\sigma_\nu^2 + 2V_x} \cdot \nu_1 \\ &< \frac{\sigma_\nu^2}{\sigma_\nu^2 + 2V_x} \cdot E\left(\mu|S_x\right) + \frac{V_x}{\sigma_\nu^2 + 2V_x} \cdot \nu_0 + \frac{V_x}{\sigma_\nu^2 + 2V_x} \cdot \nu_1 \\ &= E\left(\mu|S_x, \nu_0, \nu_1\right), \end{aligned}$$

where V_x is the variance of μ conditional on S_x. Note that paying a wage equal to the losing firm's optimal bid implies that the winning firm observes the losing firm's signal. $E\left(\mu|S_x, \nu_0, \nu_1\right)$ is firm 0's *ex post* expectation of productivity, implying that the winning firm extracts an expected first-period rent of $\frac{V_x}{\sigma_\nu^2+2V_x}\left(\nu_0 - \nu_1\right)$.

Klemperer (1999) provides a very intuitive explanation for this equilibrium bidding strategy: The case in which a bidder is tied for having the highest signal is the marginal case in which that bidder is indifferent between winning and losing. If she bids any higher, she risks paying more than her *ex post* expectation. If she bids any lower, she could have improved her chances of winning at a positive profit by increasing her bid slightly.

Finally, it is important to note that the equilibrium described above is only one possible equilibrium. This paper follows MW in only considering the strategically symmetric equilibrium (i.e., the equilibrium in which all bidders follow the same strategy) of a two-player common-value English auction. There is also a continuum of strategically asymmetric equilibria in which one bidder bids more aggressively and the other bids more timidly (Milgrom, 1981). Bikhchandani and Riley (1993) show that if any component of the item's value is private, then the symmetric equilibrium is unique, at least when the auction is limited to two bidders[8]. The likelihood of there being some degree of match-specific productivity in labor markets makes the symmetric equilibrium a likely approximation of reality. Furthermore, asymmetric equilibria in the current setting would require that firms adopt a well known strategy as far as when to bid aggressively and when not to. An asymmetric equilibrium would not be stable if firms did not know how aggressive competing bidders were, or if some firms were always aggressive and others always timid[9].

3.2. Bidding for an Employed Worker

In the previous subsection, the two firms bidding for the worker were equally well informed. When a firm bids for the services of a worker who is already employed, however, information is not symmetric. The current employer has more reliable information. Fortunately, the equilibrium in common-value English auctions does not depend on the two bidders having symmetric information[10], allowing it to apply to bidding wars that occur for employed workers.

Let t index the number of periods since the worker's last spell of nonemployment, and let $S_t = \mu + \eta_t$ denote the current employer's signal, where $\eta_t \sim N\left(0, \sigma_t^2\right)$ and $\sigma_t^2 < \sigma_\nu^2$. (This signal and its evolution will be described in greater detail in the next section.) Assume that in each period of employment a firm, g, that has not previously encountered the worker draws a signal, ν_g, of the form specified in equation (3), and observes the worker's wage, w_{xt}[11]. Also

[8] This assumes that firms will not bid in order to form reputations for being aggressive. I maintain the same assumption. Bikchandani (1988) and Klemperer (1998) show that one bidder having the advantage of such a reputation can severely influence second price or English auctions even when that advantage seems very small.

[9] One would expect firms that were never aggressive to be driven out of the market since they would retain very few workers and would extract less rent from the workers they did retain. My conjecture is that the only asymmetric equilibria that would be stable in this setting are those where either the current employer or the outside employer was always aggressive.

[10] MW point this out, and it is easily seen by examining their proof of this equilibrium (Thm 6, MW).

[11] The worker will always have an incentive to reveal her wage to outside firms. This follows from the well-known

assume that firm g observes the length of the current employment spell, t, but infers nothing new from it other than the precision of the current employer's information. Finally, assume that once the worker has been interviewed by a firm it is costless for that firm to bid for the worker.

Now assume that the worker treats her current wage, w_{xt}, as a reservation price when a new firm is encountered. This imposes some degree of downward wage rigidity, preventing the worker's wage from falling every time an outside firm receives a low signal of her productivity. When the outside firm's optimal bid is below the current wage, I assume that no bidding takes place, but the outside firm (costlessly) reveals its optimal bid. This assumption simplifies the later discussion of employer learning by avoiding issues of what the current employer would infer if no outside firm bid for the worker in a given period, but it does not qualitatively affect my results[12]. Finally, in order to simplify my discussion of the sequence of wages, I add the assumption of complete downward wage rigidity, even though none of the results that follow require such a strong assumption.

The fact that the outside firm observes the worker's wage decreases but does not eliminate the current employer's informational advantage. The information contained in the current wage, w_{xt}, is a subset of the information contained in S_t[13]. Let $\widetilde{S}_t$ denote the information in S_t that is not contained in the wage. When the outside firm observes w_{xt} the current employer's private information is summarized by $\widetilde{S}_t$ instead of S_t.

result of MW that the seller in an auction maximizes expected revenue by revealing all relevant information.

[12] If no outside firm bids for a worker in a given period, that worker's employer would infer that an outside firm recieved a signal that was too low to make bidding above w_{xt} profitable. The employer would update its expectation of the worker's productivity using this information instead of the actual signal.

[13] This will be clear once S_t is described in the next subsection. For now, realize that, if w_{xt} contained information that were not in S_t, then S_t would not represent all of the information possessed by the current employer.

As mentioned above, the equilibrium in this bidding war is unaffected by the current employer having more precise information than the outside firm. The optimal bid for each firm is still the expectation of productivity conditional on its own (private) signal and the signal of the other firm being the same. In the previous subsection, the worker was unemployed and both firms had identically distributed signals. In this case, the current employer's private signal is $\widetilde{S}_t$ and that of the outside firm is ν_g, as described by equation 3. The outside firm's optimal bid is

$$b(\nu_g) = E\left(\mu | S_x, w_{xt}, \widetilde{S}_t = \nu_g, \nu_g\right).$$

The optimal bid of the current employer is

$$\begin{aligned} b(S_t) &= E\left(\mu | S_x, w_{xt}, \widetilde{S}_t, \nu_g = \widetilde{S}_t\right) \\ &= E\left(\mu | S_x, S_t, \nu_g = \widetilde{S}_t\right). \end{aligned}$$

The result of this bidding process for employed workers is that in each period either

1. the outside firm's optimal bid is below the current wage and the worker's wage remains unchanged, but $b(\nu_g)$ is revealed to the current employer;

2. the outside firm bids above the current wage but the worker is retained at a new, higher wage equal to the bid of the outside firm, $b(\nu_g)$; or

3. the outside firm bids above the current wage, and the worker is bid away at a wage equal to the bid of the now former employer, $b(S_t)$.

Because the winning bidder always observes the signal of the losing bidder and each worker is equally productive in all firms, the information contained in S_t is passed on to the new employer when the worker is bid away. As a result, the precision of the current employer's private information depends on the amount of time that has passed since the worker's last spell of unemployment, not on the worker's tenure with that employer. If, on the other hand, there were some match-specific aspect to productivity, the information collected by one firm about a worker's productivity would be more informative for that firm than for any other and the precision of each employer's information would be increasing in tenure in addition to spell length[14].

Although allowing match-specific productivity would likely make my model more realistic, I maintain the assumption that each worker is equally productive in any firm for two reasons. The lesser reason is simplicity. Although my model could easily incorporate match-specific productivity, it would complicate notation through most of the paper. The more important reason is that developing my model in this more restrictive environment helps highlight the strengths of the model. It shows that there can be wage growth that reflects private employer learning and mobility between jobs in the face of asymmetric learning, even when there are no match- (or firm-) specific components of productivity. Previous models of asymmetric learning were unable to produce either of these predictions.

3.3. The Current Employer's Learning

The bidding between firms described above results in a worker's employer observing the signals received by outside firms as long as the worker is retained. This subsection describes the current

[14] Essentially, S_t at the previous employer would be viewed as a signal of general productive ability by the new employer, with the match-specific components of S_t being treated as an additional source of error.

employer's learning based on the accumulation of these signals. For simplicity, it is assumed that the employer receives no other signals of worker productivity, but the results are unaffected if, for example, the current employer also observes signals based on per period output[15].

Since all of the signals are identically distributed the current employer's updated signal is simply the average of all initial signals received since period 0, when the worker was unemployed:

$$\begin{aligned} S_t &= \frac{1}{t+1} \cdot \sum_{g=0}^{t} \nu_g \\ &= \mu + \eta_t, \end{aligned} \tag{4}$$

where $\eta_t \sim N\left(0, \sigma_t^2\right)$, $\sigma_t^2 = \frac{\sigma_\nu^2}{t+1}$. The reliability of S_t obviously improves (σ_t^2 falls) as the length of the current employment spell, t, increases and more initial signals are observed.

Equation 4 applies even when the worker was bid away from another employer after t' periods of continuous employment. S_t can be written to match what a firm observes in this case:

$$S_t = \frac{\sigma_{t'}^2 \nu_{t'+1} + \sigma_\nu^2 S_{t'} + \sum\limits_{g=t'+2}^{t} \sigma_{t'}^2 \nu_g}{\left(t - t'\right) \sigma_{t'}^2 + \sigma_\nu^2},$$

where $\nu_{t'+1}$ is the initial signal received by the new employer. Once the value of $S_{t'}$ is plugged in from equation 4, however, this also reduces to equation 4.

The current employer's conditional expectation of productivity for the worker at any spell

[15] An earlier version of this paper incorporated signals from per period output. Since worker productivity is general, excluding such signals simplifies notation without changing any of the results.

duration t and experience x can be written as

$$\begin{aligned} E\left(\mu|S_x,S_t\right) &= \frac{\sigma_x^2\sigma_t^2}{D}\cdot m+\frac{\sigma_t^2\sigma_\mu^2}{D}\cdot S_x+\frac{\sigma_x^2\sigma_\mu^2}{D}\cdot S_t \\ &= \beta_m m+\beta_x S_x+\beta_t S_t, \end{aligned} \tag{5}$$

where $D=\sigma_x^2\sigma_t^2+\sigma_t^2\sigma_\mu^2+\sigma_x^2\sigma_\mu^2$. The relative weight put on S_t in this expectation decreases in the variance of the error on the firms' initial signals, σ_ν^2, and increases in the number of periods since the last nonemployment spell, t[16].

3.4. The Sequence of Wages

As mentioned above, I assume that wages are downwardly rigid. This entails assuming that, if $E\left(\mu|S_x,S_t\right)$ falls below the current wage, the worker is either fired or the wage is lowered and the worker immediately quits[17]. Although this assumption could be relaxed, it simplifies the analysis of this subsection by ensuring that wages are a monotonically increasing sequence.

The sequence of bidding wars described above, combined with downward wage rigidity, implies that for any worker with an employment spell of length t, there is an increasing sequence of $n(t)$ wages: $w_1,..,w_{n(t)}$, where $n(t)$ is the number of times an outside firm bid higher than the current wage during the worker's current employment spell. In the following proposition I establish that, although wages at t do not generally equal $E\left(\mu|S_x,S_t\right)$, they converge to $E\left(\mu|S_x,S_t\right)$ as t goes to infinity. This is important because it means that, as the length of uninterrupted

[16] $\frac{\partial\beta_t}{\partial\sigma_\nu^2}=-\frac{\left(\sigma_x^2+\sigma_\mu^2\right)\sigma_x^2\sigma_\mu^2}{(t+1)D^2}<0$, $\frac{\partial\beta_t}{\partial t}=\frac{t\sigma_\nu^2\left(\sigma_x^2+\sigma_\mu^2\right)\sigma_x^2\sigma_\mu^2}{(t+1)^2D^2}>0$

[17] Assuming downward wage rigidity of this form also provides a mechanism by which workers can become unemployed in this model, which is necessary for it to be complete. Alternatively, employers could be allowed to lower wages with expected productivity and the model could, for example, include an exogenous job destruction rate.

employment increases, a worker's wage resembles her current employer's expectation more and the market's expectation less due to competition from less well-informed firms. As a result, wages reflect the private learning of current employers, despite their informational advantage.

Proposition 1. *The sequence of $n(t)$ bidding wars for a worker with t periods of uninterrupted employment creates a sequence of wages, $w_{n(t)}$, that converges to the current employer's conditional expectation of the worker's productivity, $E(\mu|S_x, S_t)$, as t goes to infinity.*

Proof. This proposition follows easily from the model of wage determination described above and the monotone convergence theorem[18]. First, note that $E(\mu|S_x, S_t)$ is a bounded sequence, converging to μ as $t \to \infty$. Next, note that the sequence of wages, $w_{n(t)}$, is increasing by construction and bounded above by the sequence of conditional expectations. (Employers do not pay wages that exceed expected productivity.) The sequence of wages must, therefore, converge to some point that is less than or equal to $\lim_{t\to\infty} E(\mu|S_x, S_t)$. The sequence of wages cannot, however, converge to any point less than $\lim_{t\to\infty} E(\mu|S_x, S_t)$ because for any $w_{n(t)} < E(\mu|S_x, S_t)$ there is a positive probability that in $t+1$ some firm g will receive a signal such that $w_{n(t)} < E\left(\mu|S_x, w_{xt}, \widetilde{S}_t = \nu_g, \nu_g\right) \leq E(\mu|S_x, S_t)$ because μ and the signals have continuous distributions. ■

This section has developed a model of asymmetric employer learning that relaxes the restrictions earlier papers placed on the learning process. As a result, the informational advantage of the current employer no longer prevents other firms from profitably bidding for an employed

[18] This proposition would still follow if wages were no longer downwardly rigid. In that case, the current employer would lower wages when $E(\mu|S_x, S_t)$ fell below what the worker was being paid, but these deviations from monotonicity would decrease in magnitude as t increased and fluctuations in $E(\mu|S_x, S_t)$ decreased in magnitude. At the same time, wages would still be driven toward the employer's expectation by bids from competing firms.

worker. Workers move between firms even when there is no match specific productivity, although the model could easily be extended to allow for match-specific productivity[19]. More importantly, the above proposition shows that competition from less well-informed firms forces current employers to raise the wages of workers toward their expectations of the workers' productivity, providing a mechanism for wage growth under asymmetric learning even in the absence of promotions.

The rest of this paper adapts the model to provide an empirical test for asymmetric learning. The test is based on the above proposition, which implies that wages will reflect evidence of employer learning with the length of continuous employment under the following conditions:

1. The current employer learns more about the worker than the market does.

2. Other firms obtain signals from interviews or other evaluations that provide some amount of private information, even if that information is very imprecise compared to the current employer's private information.

4. Estimation of Wage Equations under Asymmetric Learning

Recall that a worker's wage, w_{xt}, is always less than or equal to the current employer's conditional expectation. On average, the difference between w_{xt} and $E(\mu|S_x, S_t)$ decreases with the length of the current employment spell, but this difference does not decrease monotonically because of the random nature of the signals. At the same time, the wage moves further away from the wage earned when the employment spell began, $w_{x'1} = E(\mu|S_{x'}, \nu_0 = \nu_1, \nu_1)$, where $x' = x - t$

[19] To see this, first recall that the second-price auction results of MW that I've applied here hold generally, under any combinations of private- and common-value auctions. Then imagine that $\lim_{t\to\infty} E(\mu|S_{xj}, S_{tj}) = \mu + \varepsilon$, where ε is a match-specific error term and tenure is used in place of spell length. The rest of the proof should still hold.

is the worker's experience prior to the current employment spell and ν_1 is the signal received by the firm that was outbid at the beginning of the employment spell. More formally, the expected wage can be written as

$$E(w_{xt}) = (1 - \rho(t)) \cdot E(\mu | S_{x'}, \nu_0 = \nu_1, \nu_1) + \rho(t) \cdot E(\mu | S_x, S_t), \tag{8}$$

where $\rho(t)$ is an increasing function such that $\rho(1) = 0$ and $\rho(t) \to 1$ as $t \to \infty$. Note that $\rho(t)$ is a function of the ability of outside firms to compete for the worker: the more reliable the signals of outside firms, the faster $\rho(t)$ converges to 1.

Expanding the expectations, and recalling that $S_x = \mu + \eta_x$, $S_t = \mu + \eta_t$, and $\nu_1 = \mu + e_1$, equation (8) can be rewritten as a weighted average of the population mean and the worker's productivity:

$$E(w_{xt}) = B_m m + B_{xt}\mu + \phi'. \tag{9}$$

A more detailed derivation and description of equation (9) is left to Appendix B. The following results for equation (9) are easily shown, but doing so is also left to Appendix B:

- The weight put on actual productivity, B_{xt}, is increasing while that on the population mean, B_m, is decreasing in experience due to public learning.

- Under asymmetric learning, B_{xt} increases with t while B_m decreases.

Essentially, wages become more correlated with productivity as the length of the current employment spell increases through an interaction of the employer's private learning and the wage's convergence to the employer's expectation. As a result, there will be evidence of employer learning with both experience and the length of the current employment spell in wage equations.

The obvious problem with equation (9) is that neither the worker's productivity, μ, nor the population average productivity, m, is observed in the data. In order to deal with this, I estimate wage regressions that are a simple extension of those estimated by Altonji and Pierret (2000) (AP from here on)[20]. First, consider a variable, V, that is correlated with productivity and observed by the econometrician, but is not observed by the market. Assume the variance of V, σ_V^2, and the covariance of V and productivity, $\sigma_{V\mu}^2$, do not vary with experience or employment spell length. Also assume that V is uncorrelated with η_x and η_t, the error terms in the updated signals. When μ is replaced by V in equation (9), the expectation of the OLS estimate of B_{xt} from the resulting regression is

$$E\left(\widehat{B}_{xt}\right) = B_{xt} \cdot \frac{\sigma_{V\mu}}{\sigma_V^2}.$$

The degree of bias, $\frac{\sigma_{V\mu}}{\sigma_V^2}$, does not vary with experience or employment spell length. Although using a hard-to-observe variable in place of actual productivity produces biased results, the bias does not interfere with the model's basic predictions.

The next step in producing an estimable version of equation (9) is to replace the population mean, m, with an index of easily observed variables, $Z\delta$. Essentially, $m = Z\delta$ is now interpreted as the average productivity of workers with easy-to-observe characteristics Z. Equation (9) can be rewritten as

$$\begin{aligned} E\left(w_{xt}\right) &= B_m\left(Z\delta\right) + VB_{xt} + \phi'' \qquad (10) \\ &= Z\gamma_{xt} + VB_{xt} + \phi''. \end{aligned}$$

[20] In contrast to AP, I use wage levels instead of logs in order to better match the theory. Farber and Gibbons (1996) used wage levels for the same reason. In preliminary estimation, log wage estimates were qualitatively similar to wage level estimates.

Assuming that Z is also uncorrelated with η_x and η_t, all of the results from equation (9) that are summarized above translate directly to equation (10). The predicted behavior of γ_{xt} is also qualitatively unaffected by the bias in B_{xt}.

The estimated wage regressions approximate γ_{xt} and B_{xt} using linear experience and employment spell length interactions:

$$w_{xt} = Z\gamma_0 + Z \cdot x\gamma_x + Z \cdot t\gamma_t + VB_0 + V \cdot xB_x + V \cdot tB_t + \phi'_{xt}. \tag{11}$$

Public learning implies that γ_x should be negative and B_x should be positive. Asymmetric learning implies that γ_t should be negative and B_t should be positive. In either case, wages become more correlated with the hard-to-observe variable, V, and less correlated with the easily observed variables, Z, as they become more correlated with actual productivity.

It is important to note, however, that the predictions for the coefficients on easily observed variables only hold when a hard-to-observe variable is included in the regression[21]. Without the interactions of the hard-to-observe variable included, employer learning would predict no change over time in the coefficients on easily observed variables. On the other hand, other factors could cause the effects of race or education, for example, to vary with labor market experience. Since these other factors could swamp the effects of employer learning, I compare estimates from regressions that restrict B_x and B_t to be zero to unrestricted estimates of equation (11). Regardless of what other factors effect changes in the coefficients on easily observed variables over time, γ_x and γ_t should fall (become less positive or more negative) when the interactions of the hard-to-observe variable are included in the regression.

[21] This was the major result that distinguished the work of AP from Farber and Gibbons (1996) and allowed them to test for "rational stereotyping" based on race and education.

5. Data

The regression estimates presented in this paper use data from the 2000 release of the NLSY. The NLSY data have two key advantages for the analysis in this paper. First, the data contain variables, such as AFQT scores, that are likely to be correlated with productivity but also difficult for employers to observe. Second, the data provide a large panel that includes detailed information on worker employment histories. This work history data allows the measurement of both actual work experience and employment spell length.

Employment spell length is measured using data on weekly labor force status. An employment spell ends if the worker is not employed during a week and her last job ended with an involuntary termination (firing, etc.), or if the worker is not working for at least two weeks in a row and neither returns to work at her last job nor reports making a job-to-job transition. Each employment spell then begins counting weeks worked after the previous spell ended. Employment spells are thought of as continuing through periods of nonwork after which the worker returns to the same employer, since it is unlikely that an employer would lose information gained about a worker when the worker, for example, takes a few weeks of leave. I also experimented with other definitions of an employment spell, but the estimation results were always qualitatively similar[22].

The data used for estimation are restricted to produce a sample of workers who are both committed to the labor market and likely to be paid based on their performance. Attention is limited to men who have left school for the final time by the beginning of the job in question,

[22] In preliminary estimation I defined a spell as ending every time the worker went at least two weeks without working and obtained qualitatively similar results to those presented here. I also tried defining employment spells as ending when the worker had longer spells of nonemployment without noticing qualitatively different results. In all cases, weeks of uncertain labor force status were treated as periods of nonemployment.

are not in the military and have completed at least 12 years of schooling. All jobs reported in a year are used in the analysis as long as their hourly wage is greater than $2 and less than $200, they involve between 30 and 100 hours per week, and they are not in public administration (SIC 907-937). Finally, I exclude observations of workers who have been out of the labor market at least 25% of their career up to that interview. Eligible observations are drawn from all years of the survey (1979-2000). The resulting sample has 39,885 valid observations for 3,881 men.

All wages are converted to 1984 dollars and AFQT scores are adjusted by the age at which the test was taken. Following AP, I subtract the average percentile score for the individual's age group from the individual's score and divide the difference by the standard deviation of AFQT for that age group. This results in an AFQT measure with a standard normal distribution that adjusts for AFQT scores being higher on average for individuals who were tested at an older age.

Table 1 presents basic summary statistics for my sample. No sample weighting is used for these or any other estimates in this paper. The average hourly wage, in 1984 dollars, is $8.80. Almost 70% of the sample is white and just over 75% resides in an urban area. The average worker has completed 13.3 years of schooling, and the average employment spell has lasted for 4.3 years. The average worker has 18.9 years of potential experience and 13.5 years of actual experience at the 2000 interview. Over all years in the sample, the average potential experience is 10.3 years, while the average actual experience is 7.1 years.

6. Estimation Results

The estimation results presented in this section provide strong evidence of asymmetric employer learning. For the sake of comparison, this section also presents results from regressions estimated under the assumptions of a pure public learning model. The results suggest that a significant portion of the effects of learning observed in tests of public learning may actually be due to private learning on the part of current employers. Despite the current employers' informational advantage, their private learning appears to affect the wages of workers at a faster rate than the market's public learning. It is important to note, however, that the employer's private information is lost when the employment spell ends, while public information is not.

All of the results presented below are from regressions that include dummy variables for urban residence and year. Following Farber and Gibbons (1996), interactions of the year dummies and years of schooling are included to allow the return to education to vary by year. I use quartic polynomials in the experience measure and (in the asymmetric learning equations) spell length to control for the influence of x and t on wages[23]. Years of schooling and a dummy variable for being white are the easily observed (Z) variables, and the adjusted AFQT score is the hard-to-observe (V) variable[24]. When AFQT is missing, it is coded as zero and a dummy variable indicating missing values is added. When appropriate, interactions of the missing value dummy with experience and employment spell duration are also added.

Table 2 presents results from wage regressions estimated under the assumptions of a public

[23] In equation (11), the effects of x and t are incorporated in the error term, ϕ'_{xt}, which is a nonlinear function of both.

[24] AP use father's education in addition to adjusted AFQT as their hard-to-observe variables. Intuitively, it seems that father's education is likely to have effects on the child that are observable to employers. In preliminary estimation, I find (results not shown) that father's education, as well as sibling's education, may act as an observable variable. AP also estimate some specifications with sibling's wage, which I have not yet looked at.

learning model. The two columns on the left present OLS results with experience measured as potential experience. The two columns on the right present IV results with experience measured as actual experience and potential experience used as an instrument since actual experience likely contains information about worker productivity.

Both the OLS and IV estimates in Table 2 support the existence of public learning. Most of the evidence comes from the interactions of experience (either potential or actual) with AFQT scores. AFQT has a large effect on wages when experience interactions are not included [0.94 (0.08) for OLS, 0.90 (0.08) for IV], but most of this effect is due to wages becoming more correlated with AFQT over time. When experience interactions are added, the initial effect of AFQT falls to 0.45 (0.12) in the OLS and 0.37 (0.11) in the IV regressions. The coefficient on AFQT $\times$ the experience measure is significantly positive at 0.048 (0.011) in the OLS regressions and even larger [0.074 (0.015)] in the IV regressions. Coefficients on the easily observed variables interacted with experience always become more negative (or less positive) when AFQT $\times$ the experience measure is added, but the change is never significant.

Moving to the test of the asymmetric learning model, Table 3 presents results from OLS wage regressions that use potential experience as the experience measure. The results provide strong support for my model, with most of the evidence again coming from AFQT scores. The results for grade completed and the white dummy variable are always consistent with public and private employer learning, but are never significant. As before, AFQT scores have a significant influence on wages, but most of that influence is due to wages becoming more correlated with ability over time. When the interactions of AFQT with x and t are added the coefficient on AFQT itself falls from 0.84 (0.08) to 0.36 (0.12). More importantly, the coefficient on AFQT $\times$ employment spell length in column III is 0.055 (0.020). The coefficient on AFQT $\times$ potential experience is 0.024 (0.013).

To put these coefficients on AFQT in more concrete terms, a one standard deviation increase in the adjusted AFQT score increases hourly wages by $0.27 more after five years of continuous employment than at the beginning of an employment spell. By comparison, an extra 5 years of potential experience raises the effect of the same change in AFQT by only $0.12. Despite the informational asymmetry, the current employer's private learning appears to affect wages faster than the market's public learning.

The IV regressions presented in Table 4 use instruments for employment spell length and its interactions, in addition to instruments for actual experience and it's interactions. The length of an employment spell could contain information about worker productivity, just as actual experience could[25], and evidence of employer learning would be biased if it relied on measures of time that were correlated with information learned by employers. In an attempt to create an instrument that is not correlated with information about labor market history, I first regress the length of the current employment spell on the worker's career average spell length, the average number of employment spells the worker has had per year of potential experience, and the ratio of actual to potential experience. Assuming these variables adequately control for information on labor market history, this residual is a valid instrument for employment spell duration. The interactions of this residual can also be used as instruments for the interactions of spell duration.

The results from this IV estimation again support my model. Looking at column III in Table 4, the coefficient on AFQT $\times$ employment spell duration is 0.072 (0.027), and that on AFQT $\times$ actual experience is 0.042 (0.021). According to these estimates, a one standard deviation

[25] For example, a worker who has been continuously employed for a long time is almost certainly more desirable than an otherwise similar worker whose labor market experience consists of a series of short employment spells.

increase in the adjusted AFQT score increases hourly wages by \$0.36 more after five years of continuous employment than at the beginning of an employment spell. During those five years of continuous employment, the market's public learning would increase the effect of that change in AFQT by an additional \$0.21.

Finally, note that all of my estimation results are consistent with the "rational stereotyping" of workers by both race and years of schooling, although the observed effects could easily be due to other factors. As mentioned above, coefficients on the interactions of grade and the white dummy variable should fall when analogous interactions of AFQT are added to the regression, if these easily observed variables are (perhaps illegally) used as signals. In Table 4, as in the earlier tables, the coefficients on grade or the white dummy variable interacted with experience or employment spell length decrease, but never significantly, when the analogous interactions of AFQT are added.

7. Conclusions and Directions for Future Research

This paper has developed and tested a model of asymmetric employer learning that relaxes the restrictions earlier papers placed on the learning process. As a result, firms can profitably bid for an employed worker, despite the current employer's informational advantage. In contrast to earlier work in this literature, workers in this model can be bid away from their current employer by less well-informed firms even though there is no match- or firm-specific productivity. Furthermore, competition from less well-informed firms forces the current employer to raise the worker's wages toward the employer's expectation of that worker's productivity, resulting in wage growth even in the absence of promotions.

This convergence of wages to the current employer's expectation allows the model to be tested empirically. The model implies that wages reflect evidence of employer learning as both experience and employment spell length increase. The test I derive is an extension of the employer learning work of Altonji and Pierret (2000), requiring only the estimation of basic wage regressions using data from the NLSY. The results of this estimation suggest that both public learning and the private learning of current employers affect wages. In fact, it appears as though competition from outside firms is strong enough to cause the employer's private learning to affect wages during an employment spell at least as much as the market's learning.

This paper opens multiple avenues for future research. First, the model could easily be extended to allow match- or firm-specific productivity. I have assumed throughout the paper that a given worker is equally productive at any firm to show that my model produces interesting results without allowing worker productivity to vary across firms; however, that assumption is not necessary. If a worker were no longer equally productive at any firm and there were asymmetric learning, the bidding process in this model would no longer result in the new employer learning everything about the worker that the previous employer knew. The match-specific elements of one employer's signal would always act as an additional error term when the signal was observed by another employer, regardless of how long the worker had been employed. Both the current employer's information about its own match to the worker and the worker's general productivity would both become more precise with tenure.

The model also has more implications for wage and employment dynamics than I have developed here. Developing these implications more completely could lead to interesting comparisons with other models, and possibly produce additional empirical tests. This effort could be aided by embedding this model into a basic search framework[26]. If nothing else, adding a fuller model of job search would provide a more realistic description of unemployment than the model does in its current form.

Finally, the model developed in this paper could provide an interesting framework for further work on labor market discrimination. Both Oettinger (1996) and Milgrom and Oster (1987) develop models of statistical discrimination that exploit asymmetric learning, while Altonji and Pierret (2000) and Pinkston (2002) examine empirical evidence of different types of statistical discrimination in public learning frameworks[27]. Not only do the results in this paper provide weak support for a "rational stereotyping" form of statistical discrimination, but preliminary evidence (not presented) suggests that black and white men differ in the effect asymmetric learning has on their wages relative to public learning. Interpreting such a result, however, would require theoretical work that is outside of the scope of this paper.

[26] Postel-Vinay and Robin (2000) incorporate a bidding framework that resembles a private-value English auction between firms that differ in productivity. The bidding and informational structure of this model should not be much more difficult to incorporate into a search model than theirs was, and the model could have interesting implications for both wage dispersion and employment dynamics.

[27] The term "statistical discrimination" is used to refer to two different types of discrimination. "Rational stereotyping" assumes that employers illegally use race or gender as a signal of worker ability. "Screening discrimination" assumes that employers are less able to evaluate the ability of workers from one group than another. Altonji and Pierret (2000) look at rational stereotyping. Pinkston (2002), Oettinger (1996) and Milgrom and Oster (1987) all consider screening discrimination.

References

[1] Altonji, Joseph G. and Charles R. Pierret. "Employer Learning and Statistical Discrimination," *Quarterly Journal of Economics*, 2000.

[2] Baker, G. P., M. Gibbs and B. Holmstrom. "The Internal Economics of the Firm: Evidence from Personnel Data," *Quarterly Journal of Economics*, 109, 1994.

[3] Baker, G. P., M. Gibbs and B. Holmstrom. "The Wage Policy of a Firm," *Quarterly Journal of Economics*, 109, 1994.

[4] Bernhardt, Dan. "Strategic Promotion and Compensation." *Review of Economic Studies*, 62, 1995, pp. 315-339.

[5] Bernhardt, Dan and Scoones, David. "Promotion, Turnover and Preemptive Wage Offers," *American Economic Review*, 83, 1990, pp. 771-791.

[6] Bikhchandani, Sushil. "Reputation in Repeated Second-Price Auctions," *Journal of Economic Theory* 46, 1988, pp. 97-119.

[7] Bikhchandani, Sushil, and John G. Riley. "Equilibrium in Open Auctions," UCLA, 1993.

[8] Farber, Henry S. and Robert Gibbons. "Learning and Wage Dynamics," *Quarterly Journal of Economics*, CXI (1996), pp. 1007-47.

[9] Gibbons, Robert, and Lawrence F. Katz. "Layoffs and Lemons," *Journal of Labor Economics*, 9(4), 1991.

[10] Gibbons, Robert, and Michael Waldman. "Careers in Organizations: Theory and Evidence," Chapter 36, *Handbook of Labor Economics*, Vol. 3, 1999.

[11] Klemperer, Paul. "Auctions with Almost Common Values: The "Wallet Game" and its Applications," *European Economic Review* 42, 1998, pp. 757-69.

[12] Klemperer, Paul. "Auction Theory: A Guide to the Literature," *Journal of Economic Surveys*, 13(3), 1999, pp. 227-86.

[13] McAfee, R. Preston and John McMillan. "Auctions and Bidding," *Journal of Economic Literature*, vol. XXV, 1987, pp. 699-738.

[14] Medoff, J. and K. Abraham. "Experience, Performance, and Earnings," *Quarterly Journal of Economics*, 95, 1980.

[15] Medoff, J. and K. Abraham. "Are Those Paid More Really More Productive?," *Journal of Human Resources*, 16, 1981.

[16] Milgrom, Paul. "Rational Expectations, Information Acquisition, and Competitive Bidding," *Econometrica*, 49(4), 1981, pp. 921-43.

[17] Milgrom, Paul, and Sharon Oster. "Job Discrimination, Market Forces, and the Invisibility Hypothesis," *Quarterly Journal of Economics*, CII (August 1987), pp. 453-76.

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

[19] Oettinger, Gerald S. "Statistical Discrimination and the Early Career Evolution of the Black-White Wage Gap," *Journal of Labor Economics*, 14(1) (1996), pp. 52-77.

[20] Pinkston, Joshua C. "A Test of Screening Discrimination with Employer Learning." Bureau of Labor Statistics, 2002.

[21] Postel-Vinay, Fabien, and Jean-Marc Robin. "The Distribution of Earnings in an Equilibrium Search Model with State-Dependent Offers and Counter-Offers," INRA-LEA, 2000.

[22] Scoones, David, and Dan Bernhardt. "Promotion, Turnover, and Discretionary Human Capital Acquisition," *Journal of Labor Economics*, 16, 1998, pp. 122-141.

[23] Waldman, Michael. "Job Assignments, Signaling and Efficiency," *Rand Journal of Economics*, 15, 1984, pp. 255-267.

A. The English Auction Equilibrium

This appendix discusses the optimal bid in a two-player English auction, as developed in MW. The focus will be on the assumptions they make and how those assumptions nest the assumptions of this paper.

The relevant assumptions from MW are:

1. (Assumption 2) *The value of the object to each bidder is nonnegative, continuous in its variables, and nondecreasing in its variables.* As long as the distribution of μ is such that firms never believe that a worker will have negative productivity, this assumption is very easily met by the distributional assumptions of my model.

2. (Assumption 3) *The expectation of each bidder's value for the object is finite.* This is obviously true in my model

3. (Assumption 5) *The signals received by bidders and the other variables that influence the value of the object (μ_i in my model) are "affiliated".* Roughly speaking, two variables are affiliated if a high value of one makes it more likely that the other has a high value. Klemperer (1999) explains affiliation as being equivalent to local correlation everywhere. This is guaranteed by my assumption that all signals equal μ_i plus a mean-zero standard error. MW show that variables are affiliated if the distribution of μ_i conditional on those variables satisfies the monotone likelihood ratio property.

The other main assumptions that MW make (1 and 4) relate to the symmetry of the auction and are not required for this second-price or English auction equilibrium to hold. (These assumptions are only used for their discussion of first-price auctions.)

As mentioned earlier in this paper, Theorem 6 in MW describes an equilibrium in the second-price case. (When there are two players, English and second-price auctions are equivalent.) The theorem states that the (strategically symmetric) equilibrium is the point where every bidder bids her value of the object conditional on her signal and the signal of the next highest bidder being the same. The proof of this theorem is a fairly straightforward maximization problem, with the bidder in question maximizing over the value to be substituted in place of the (unobserved) signal of the next highest bidder. The proof requires that the distribution of one bidder's signal conditional on another's is continuous, but it does not require the distributions to be symmetric.

B. Equation 9 and its Coefficients.

Recall that the expected wage can be written as

$$E\left(w_{xt}\right)=\left(1-\rho\left(t\right)\right)\cdot E\left(\mu|S_{x'},\nu_0=\nu_1,\nu_1\right)+\rho\left(t\right)\cdot E\left(\mu|S_x,S_t\right). \tag{8}$$

Expanding the expectations, and recalling that $S_x=\mu+\eta_x$, $S_t=\mu+\eta_t$, and $\nu_1=\mu+e_1$, equation (8) can be rewritten as

$$\begin{aligned}E\left(w_{xt}\right) &= \left(1-\rho\left(t\right)\right)\cdot\left[\frac{\sigma_{x'}^2\sigma_\nu^2}{D'}m+\frac{\sigma_\mu^2\sigma_\nu^2}{D'}S_{x'}+2\cdot\frac{\sigma_{x'}^2\sigma_\mu^2}{D'}\nu_1\right]+\rho\left(t\right)\cdot\left(\beta_m m+\beta_x S_x+\beta_t S_t\right)\\ &= \left(1-\rho\left(t\right)\right)\cdot\left[\frac{\sigma_{x'}^2\sigma_\nu^2}{D'}m+\frac{\sigma_\mu^2\sigma_\nu^2+2\sigma_{x'}^2\sigma_\mu^2}{D'}\mu+\phi_{x'0}\right]+\rho\left(t\right)\cdot\left[\beta_m m+\left(\beta_x+\beta_t\right)\mu+\phi_{xt}\right],\end{aligned}$$

where $D'=\sigma_{x'}^2\sigma_\nu^2+\sigma_\nu^2\sigma_\mu^2+2\sigma_{x'}^2\sigma_\mu^2$, $\phi_{xt}=\beta_x\eta_x+\beta_t\eta_t$, and $\phi_{x'0}=\frac{\sigma_\mu^2\sigma_\nu^2}{D'}\eta_{x'}+2\cdot\frac{\sigma_{x'}^2\sigma_\mu^2}{D'}e_1$. Further simplifying the wage equation yields

$$\begin{aligned}E\left(w_{xt}\right) &= B_m m+B_{xt}\mu+\phi',\ \text{where} \\ B_m &= \left(1-\rho\left(t\right)\right)\cdot\frac{\sigma_{x'}^2\sigma_\nu^2}{D'}+\rho\left(t\right)\cdot\beta_m, \\ B_{xt} &= \left(1-\rho\left(t\right)\right)\cdot\left(\frac{\sigma_\mu^2\sigma_\nu^2}{D'}+2\cdot\frac{\sigma_{x'}^2\sigma_\mu^2}{D'}\right)+\rho\left(t\right)\cdot\left(\beta_x+\beta_t\right),\ \text{and} \\ \phi' &= \left(1-\rho\left(t\right)\right)\cdot\phi_{x'0}+\rho\left(t\right)\cdot\phi_{xt}.\end{aligned} \tag{9}$$

The rest of this section will show the derivatives of B_m and B_{xt} in equation 9 in order to support the conclusions of Section 4.

B.1. Derivatives with Respect to Experience

The weight put on individual productivity, B_{xt}, is increasing and that on the population mean, B_m, is decreasing in experience due to public learning:

$$\begin{aligned}\frac{\partial B_m}{\partial x} &= \left(1-\rho\left(t\right)\right)\left(\frac{\frac{\partial\sigma_{x'}^2}{\partial x}\sigma_\nu^2 D'-\frac{\partial\sigma_{x'}^2}{\partial x}\left(\sigma_\nu^2+\sigma_\mu^2\right)\sigma_{x'}^2\sigma_\nu^2}{\left(D'\right)^2}\right)+\rho\left(t\right)\frac{\partial\beta_m}{\partial x}\\ &= \left(1-\rho\left(t\right)\right)\left(\frac{\frac{\partial\sigma_{x'}^2}{\partial x}\left(\sigma_\nu^2\right)^2\sigma_\mu^2}{\left(D'\right)^2}\right)+\rho\left(t\right)\frac{\partial\beta_m}{\partial x}\\ &< 0\end{aligned}$$

because $\frac{\partial \sigma_{x'}^2}{\partial x} < 0$ and $\frac{\partial \beta_m}{\partial x} < 0$.

$$\begin{aligned}
\frac{\partial B_{xt}}{\partial x} &= (1-\rho(t)) \left(\frac{2\sigma_\mu^2 \frac{\partial \sigma_{x'}^2}{\partial x} D' - \frac{\partial \sigma_{x'}^2}{\partial x} \left(\sigma_\nu^2 + \sigma_\mu^2\right) \left(\sigma_\mu^2 \sigma_\nu^2 + 2\sigma_{x'}^2 \sigma_\mu^2\right)}{(D')^2} \right) + \rho(t) \left(\frac{\partial \beta_x}{\partial x} + \frac{\partial \beta_t}{\partial x} \right) \\
&= (1-\rho(t)) \sigma_\mu^2 \frac{\partial \sigma_{x'}^2}{\partial x} \cdot \left(\frac{2D' - \left(\sigma_\nu^2 + 2\sigma_\mu^2\right)\left(\sigma_\nu^2 + 2\sigma_{x'}^2\right)}{(D')^2} \right) + \rho(t) \left(\frac{\partial \beta_x}{\partial x} + \frac{\partial \beta_t}{\partial x} \right) \\
&= (1-\rho(t)) \sigma_\mu^2 \frac{\partial \sigma_{x'}^2}{\partial x} \cdot \left(\frac{-\left(\sigma_\nu^2\right)^2}{(D')^2} \right) + \rho(t) \left(\frac{\partial \beta_x}{\partial x} + \frac{\partial \beta_t}{\partial x} \right) \\
\frac{\partial \beta_x}{\partial x} &= -\frac{\partial \sigma_x^2}{\partial x} \frac{\left(\sigma_t^2 + \sigma_\mu^2\right) \sigma_\mu^2 \sigma_t^2}{D^2} > 0, \frac{\partial \beta_t}{\partial x} = \frac{\partial \sigma_x^2}{\partial x} \frac{\left(\sigma_\mu^2\right)^2 \sigma_t^2}{D^2} < 0
\end{aligned}$$

The first term in $\frac{\partial B_{xt}}{\partial x}$ is clearly positive. The second term is positive as well because $\left|\frac{\partial \beta_x}{\partial x}\right| > \left|\frac{\partial \beta_t}{\partial x}\right|$. Therefore, $\frac{\partial B_{xt}}{\partial x} > 0$.

B.2. Derivatives with Respect to Employment Spell Length

Under asymmetric learning with some level of competition from outside firms, B_{xt} increases with spell length while B_m decreases:

$$\begin{aligned}
\frac{\partial B_m}{\partial t} &= \frac{\partial \rho(t)}{\partial t} \left(\beta_m - \frac{\sigma_{x'}^2 \sigma_\mu^2}{D'} \right) + \rho(t) \frac{\partial \beta_m}{\partial t} \\
&< 0
\end{aligned}$$

because $\beta_m < \frac{\sigma_{x'}^2 \sigma_\mu^2}{D'}$ (i.e., less weight is put on m in the employer's expectation) and $\frac{\partial \beta_m}{\partial t} < 0$.

$$\begin{aligned}
\frac{\partial B_{xt}}{\partial t} &= \frac{\partial \rho(t)}{\partial t} \left(\beta_x + \beta_t - \frac{\sigma_\mu^2 \sigma_\nu^2}{D'} - 2 \cdot \frac{\sigma_{x'}^2 \sigma_\mu^2}{D'} \right) + \rho(t) \left(\frac{\partial \beta_m}{\partial t} + \frac{\partial \beta_t}{\partial t} \right) \\
&> 0
\end{aligned}$$

Intuitively, the first term is positive because the current employer's expectation is more correlated with the worker's productivity than was her initial wage in the current employment spell; i.e., $(\beta_x + \beta_t) > \left(\frac{\sigma_\mu^2 \sigma_\nu^2}{D'} + 2 \cdot \frac{\sigma_{x'}^2 \sigma_\mu^2}{D'} \right)$.

Table 1. Summary Statistics*.

	Mean	Std. Dev.	Min	Max
Real Hourly Wage	8.795	6.116	2.000	193.355
Highest Grade Completed	13.263	1.906	12	20
White	0.696	0.460	0	1
Normalized AFQT	0.000	1.000	-1.871	2.544
AFQT Missing	0.044	0.204	0	1
Employment Spell Dur.	4.295	4.185	0.019	23.019
Tenure	2.423	3.038	0.019	23.038
Potential Experience	10.259	5.333	0	25
Experience	7.103	4.852	0.038	23.019
Pot. Exp in 2000	18.860	3.166	9	25
Experience in 2000	13.501	5.342	0.500	23.019
Urban	0.761	0.427	0	1

*Notes: Wages are in 1984 dollars. There are 39885 observations except for AFQT (38143 nonmissing), tenure (29821), and the experience measures in 2000 (3336).

Table 2. Coefficient Estimates under Public Learning.
OLS and IV Estimates of Wage Regressions*.

	OLS		IV	
	I	II	I	II
Grade	-0.221	-0.179	-0.564	-0.502
	(0.338)	(0.298)	(0.353)	(0.318)
Grade x Experience	0.003	-0.008	0.038	0.018
	(0.014)	(0.015)	(0.023)	(0.023)
White	-0.235	0.169	-0.319	0.117
	(0.189)	(0.201)	(0.184)	(0.193)
White x Experience	0.114	0.074	0.149	0.087
	(0.016)	(0.018)	(0.023)	(0.025)
AFQT	0.940	0.452	0.901	0.374
	(0.083)	(0.120)	(0.084)	(0.114)
AFQT x Experience	…..	0.048	…..	0.074
	…..	(0.011)	…..	(0.015)

*Note: Standard errors (in parentheses) are Huber/White, accounting for multiple observations per person.
All regressions include dummy variables for year, urban residence and missing values of AFQT, as well as the experience measure and interactions of grade with the year dummies. OLS regressions use years of potential experience and IV estimates use years of actual experience with potential experience as an instrument.

Table 3. Coefficient Estimates under Asymmetric Learning.
OLS Estimates of Wage Regressions Using Potential Experience*.

	I	II	III
Grade	-0.318	-0.329	-0.293
	(0.313)	(0.349)	(0.307)
Grade x Experience	-0.021	-0.024	-0.029
	(0.015)	(0.015)	(0.015)
Grade x Spell Duration	0.069	0.066	0.052
	(0.011)	(0.011)	(0.012)
White	0.517	-0.184	0.206
	(0.169)	(0.183)	(0.195)
White x Experience	0.096	0.101	0.081
	(0.018)	(0.018)	(0.021)
White x Spell Duration	-0.008	-0.018	-0.064
	(0.031)	(0.030)	(0.035)
AFQT	…..	0.843	0.358
	…..	(0.080)	(0.115)
AFQT x Experience	…..	…..	0.024
	…..	…..	(0.013)
AFQT x Spell Duration	…..	…..	0.055
	…..	…..	(0.020)

*Note: Standard errors (in parentheses) are Huber/White, accounting for multiple observations per person.
All regressions also include dummy variables for year, urban residence and (when appropriate) missing values of AFQT, as well as the experience measure, interactions of grade with the year dummies, and (when appropriate) interactions of experience and spell duration with the missing value dummy for AFQT.

Table 4. Coefficient Estimates under Asymmetric Learning.
IV Estimates of Wage Regressions Using Actual Experience*.

	I	II	III
Grade	-0.566	-0.600	-0.530
	(0.382)	(0.447)	(0.370)
Grade x Experience	0.029	0.025	0.018
	(0.026)	(0.026)	(0.026)
Grade x Spell Duration	0.007	0.009	-0.014
	(0.023)	(0.022)	(0.022)
White	0.322	-0.398	0.120
	(0.226)	(0.211)	(0.209)
White x Experience	0.169	0.164	0.132
	(0.029)	(0.029)	(0.033)
White x Spell Duration	-0.042	-0.032	-0.104
	(0.043)	(0.042)	(0.047)
AFQT	…..	0.834	0.227
	…..	(0.089)	(0.135)
AFQT x Experience	…..	…..	0.042
	…..	…..	(0.021)
AFQT x Spell Duration	…..	…..	0.072
	…..	…..	(0.027)

*Note: Standard errors (in parentheses) are Huber/White, accounting for multiple observations per person.

All regressions also include dummy variables for year, urban residence and (when appropriate) missing values of AFQT, as well as the experience measure, interactions of grade with the year dummies, and (when appropriate) interactions of experience and spell duration with the missing value dummy for AFQT. Potential experience and its interactions are used as instruments for actual experience and it's interactions. The residual of spell duration regressed on the worker's average spell duration and ratios of number of spells to potential experience and actual experience to potential experience (see text) is used as an instrument for spell duration, and the residual's interactions as instruments for spell duration's interactions.