BLS WORKING PAPERS

U.S. Department of Labor
U.S. 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 390
January 2006

A Model of Asymmetric Employer Learning With Testable Implications*

Joshua C. Pinkston
Bureau of Labor Statistics

May, 2005

Abstract: This paper develops and tests a unique model of asymmetric employer learning. The model relaxes the informational assumptions used in most of the previous literature and assumes firms compete for workers through bidding wars. As a result, outside firms can profitably compete for an employed worker who is equally productive in any firm, despite the current employer's informational advantage. The model in this paper is the first in the literature to predict either wage growth without changes in publicly observed information (e.g., 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 previous work on employer learning. Wage regressions estimated on a sample of men from the NLSY produce evidence consistent with the model's predictions.

JEL: J3, M5, 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, Michael Waldman, Charles Zheng and seminar participants at the Bureau of Labor Statistics, the Census Bureau, and the Federal Reserve Board. All mistakes and opinions are my own. No statement in this paper is intended to reflect opinions or policies of the Bureau of Labor Statistics or the Department of Labor.
A previous version of this paper is available as BLS Working Paper #365.
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.Josh@bls.gov.

This paper develops and tests a unique model of asymmetric employer learning. The model allows all firms that encounter a worker to receive a private signal of the worker's productivity; however, the worker's current employer is able to accumulate more private information than outside firms. Firms are also assumed to compete for workers through bidding wars. Despite the current employer's informational advantage, this assumption allows less well-informed firms to profitably bid for an employed worker even when the worker is equally productive in any firm. This competition causes wages in the model to converge to the employer's expectation as the worker's employment spell increases in length, and results in workers with the same publicly observable signals having different wages.

The convergence of wages to the employer's expectation also allows the 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. They show that as wages become more correlated with productivity due to employer learning, the coefficients in wage regressions on variables that are correlated with productivity but difficult for employers to observe will increase. The model 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.

Most of the previous literature on asymmetric employer learning has been theoretical, focusing primarily on the relationship of wages to task assignment. Waldman (1984) develops a basic model of task assignment under asymmetric information. He assumes that the current employer becomes perfectly informed after one period of tenure, and outside firms learn about the worker only by observing her task assignment at the current firm. Wages do not rise unless a promotion signals higher ability to outside firms, implying that workers in the same job (with the same publicly observable characteristics) have the same wage. Workers are inefficiently assigned to jobs as firms determine assignment strategically. Several later papers, including Milgrom and Oster (1987), Bernhardt (1995) and Scoones and Bernhardt (1998), expand on this basic model[1].

Most of these previous 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 or some other change in publicly observed information. This last point is the major criticism that Gibbons and Waldman (1999) make of the literature, and is one of the shortcomings this paper addresses.

The only previous paper in this literature that does not assume the current employer knows everything competing firms know is Waldman (1990). In his model of "up-or-out" contracts, the market observes a signal of the worker's productivity that the employer does not; however, when outside firms compete for the worker their bid reveals their information to the current employer. Because the current employer can then counter knowing the value of both the market's signal and its own, a severe winner's curse problem causes the market to condition its bid on the minimum employer signal sufficient to generate a promotion. As a result, the worker is always retained. In essence, the bidding mechanism in Waldman (1990) prevents outside firms from

[1] 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". Scoones and Bernhardt (1998) include match-specific productivity and show that workers may invest in an inefficiently large amount of firm-specific human capital in order to increase their chances of being promoted and possibly bid away.

competing for workers in a way that could generate wage growth without promotions or mobility between jobs.

The model developed in this paper shows that there can be wage growth that reflects private employer learning and mobility between jobs in the face of asymmetric employer learning, even when the worker is equally productive in any firm. I assume that the current employer only becomes perfectly informed in the limit as the length of the current employment spell approaches infinity, and I allow outside firms to receive noisy private signals from interviews or other firm-specific evaluations. I also allow outside firms to compete for workers through bidding wars (ascending bid auctions). Because the winning bidder pays the highest losing bid, bidding wars avoid the winner's curse that affects Waldman (1990) and others. As a result, outside firms can profitably bid for employed workers as long as they have some private information. This bidding causes wages to converge to the employer's expectation even when there are no promotions or other signals of the employer's private information, and workers who otherwise appear identical to the market have different wages.

Although this literature was originally motivated by studies of personnel records[2], Gibbons and Katz (1991) is the only previous paper to develop and test a model of asymmetric employer learning. In their model, 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 their predictions.

The current paper is one of a few recent papers that attempt to further bridge the gap

[2] See, for example, Medoff and Abraham (1980, 1981), and Baker, Gibbs and Holmstrom (1994a, 1994b).

between theory and evidence in this literature[3]. The model of asymmetric learning that is developed in the next section implies that wage regressions will reflect evidence of asymmetric employer learning as the length of the worker's current employment spell increases, in addition to evidence of public learning with experience. This implication is tested using data from the NLSY, and the estimation results suggest that wages during an employment spell are affected at least as much by asymmetric learning as they are by public learning.

The next section develops the model of asymmetric employer learning. In Section 2, I discuss the empirical test of my model. Section 3 describes the data I use and Section 4 presents the estimation results. Section 5 concludes and discusses avenues for future research.

1. The Asymmetric Learning Model

Turning to the theoretical contribution of this paper, I first describe the basic assumptions of my model. I then describe the equilibrium bidding strategies of firms that encounter the worker. In Section 1.3, I describe the current employers' private information and the information contained in a worker's wage. This section then concludes with a discussion of the sequence of wages produced by the model, which shows that wages converge to the current employer's expectation as the length of the employment spell increases.

[3] The other two are DeVaro and Waldman (2004), who empirically investigate the value of promotions as signals; and Schoenberg (2004), who performs tests that also extend earlier work on public learning. Schoenberg's work will be discussed later in the paper as the empirical results are presented.

1.1. Basic Assumptions

I assume workers are known to be heterogeneous with worker-specific productivity, μ_i, having a normal distribution:

$$\mu \sim N\left(m, \sigma_\mu^2\right).$$

In what follows, $x = 1, 2, ...$ indexes periods of the worker's labor market experience. Periods of time during a spell in which the worker is continuously employed, regardless of whether he changes employers, are indexed by t. Periods of tenure with the current employer are indexed by τ.

Workers are assumed to transition from employment to unemployment due to an exogenous rate of job destruction. This ensures that the length of the current employment spell and experience are not generally the same, while avoiding the added complication of modelling transitions into unemployment[4].

1.1.1. Timing and Informational Assumptions

Market Learning over Experience I assume that some learning about a worker's productivity is public so that my model will nest the learning models of Altonji and Pierret (2000) and Farber and Gibbons (1996) (AP and FG in what follows) in which all learning is public. At any level of experience x let S_x summarize the market's information. I assume S_x is an unbiased estimate of the worker's true productivity:

$$S_x = \mu + \eta_x,$$

[4] Previous versions of this paper assumed that wages were downwardly rigid and workers were fired if their expected productivity fell below the wage.

where $\eta_x \sim N\left(0, \sigma_x^2\right)$ and σ_x^2 is decreasing in experience. In other words, the market's information about a worker grows more precise with experience[5].

Intuitively, the market's information can be thought of as the information contained in a worker's resume (e.g., what jobs he has held, the duration of each and the duration of past unemployment spells). It will be important later in the paper to distinguish the market's information from information that will appear public from the viewpoint of the two firms involved in a bidding war in a given period. For example, both of those firms will observe the worker's current wage in my model; but the market as a whole will not observe that wage or the worker's wage history, and the wage will not enter into S_x. Although this assumption may at first seem unusual, I believe it is realistic: A firm that wants to hire a worker is likely to know that worker's current wage and something of her experience, but is not likely to have observed every single wage innovation over a worker's career.

Time During an Employment Spell First, I assume that any firm f that encounters worker i for the first time in any period receives a private signal, ν_{fi}, from an interview or other evaluation:

$$\nu_{fi} = \mu_i + e_{fi}, \tag{1}$$

where $e_{fi} \sim N\left(0, \sigma_\nu^2\right)$. The distribution of ν_{fi} is common knowledge. For simplicity, the worker is assumed never to encounter a firm she has worked for or been interviewed by in the past: Firms are sampled at random without replacement.

At $t = 0$, the worker is unemployed and randomly encounters two firms. Each firm interviews

[5] S_x is easily derived using a standard Bayesian updating argument; i.e., S_x is a variance-weighted sum of all signals the market has received up to period x.

the worker and receives a private signal, ν_{fi}. The firms then engage in a bidding war for the worker and the worker agrees to become employed by the winner[6].

In each later period, $t \geq 1$, of the employment spell, the worker is assumed to encounter one new firm. The worker reveals her current wage to the new firm, and that firm draws its own private signal described by equation (1). If the new firm is willing to pay more than the worker's current wage, it engages in a bidding war with the current employer. As before, the worker becomes employed by the winning bidder.

I assume there are never more than two firms bidding for the worker in any period for simplicity. I discuss relaxing that assumption to allow more bidders below.

As will be discussed shortly, bidding wars result in the winner observing any private information possessed by a losing bidder. This implies that the current employer's information at $t = 1$ includes its own signal and that of the losing firm in $t = 0$. At $t = 2$, it includes the signals of the losing bidders in $t = 0$ and $t = 1$. In any period, the current employer always knows the values of all signals received by losing firms in all previous periods of the current employment spell in addition to its own. (The evolution of S_t will be discussed in greater detail later in this section.)

Finally, I assume that the current employer lowers the wage so that $w_{xt} = E(\mu|S_x, S_t)$ if S_x or S_t fall enough that the worker's expected productivity would otherwise fall below the current wage.

[6] I'm assumming that any reservation wages unemployed individuals might have are low enough not to matter in order to simplify the presentation.

1.1.2. The Assumption of Bidding Wars

As mentioned above, I model wages as the outcomes of a series of bidding wars (English auctions), each of which has two bidders. One firm offers the worker a wage, another firm makes a counteroffer, and so on until one firm drops out. The remaining firm then hires the worker at the wage where the losing firm dropped out, which will be less than the remaining firm's *ex post* expectation of the worker's productivity.

This bidding process allows me to exploit standard results from the auction literature[7], making it very tractable, while also providing intuitively appealing results that are useful when modelling bidding under asymmetric information. Bidding wars result in winning firms observing the signals of losing firms, allowing employers to learn about a worker by observing the intensity with which other firms compete for that workers' services as in Lazear (1986). Bidding wars also result in the employer paying the worker her outside option, a result that is key, for example, in Scoones and Bernhardt (1998).

More importantly, bidding wars allow outside firms to bid profitably for workers under a wider range of assumptions than other bidding mechanisms do[8]. The winning firm in a bidding war pays the worker the highest wage that the losing firm was willing to pay. This means that the winning firm pays a wage that is less than its *ex post* expectation of the worker's productivity as long as it possesses some private information about the worker, even if the losing bidder had private information that was more precise. In contrast, assuming that the current employer makes one final counteroffer after an outside firm reveals its bid results in such a severe winner's curse for the outside firm that it can never do better than earn zero profits, even under symmetric

[7] See McAffee and McMillan (1987) and Klemperer (1999) for readable surveys.

[8] Although it is difficult to justify any specific bidding assumption empirically, we at least know that bidding wars for workers do take place. Examples can be found in academic labor markets.

information[9].

Milgrom and Weber (1982) (MW from here on) show that an equilibrium bid for each firm in this auction is the expectation of productivity conditional on the signal it receives and the signal of the other firm being the same. Abstracting for a moment from what information bidders in my model have in a certain period, if I_f and I_g denote the information of firms f and g, respectively, the optimal bid of firm f is

$$b\left(I_{fi}\right) = E\left(\mu_i | I_{fi}, I_{gi} = I_{fi}\right). \tag{2}$$

In other words, $b\left(I_{fi}\right)$ is the highest wage firm f is willing to bid for worker i when it has the information in I_{fi}[10]. The specific form this optimal bid takes for each firm in each period of my model will be described in the next subsection. 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[11].

With the bidding strategy described above, firm f will win the bidding war if $I_{fi} > I_{gi}$ (i.e.,

[9] Whether the outside bid is the bid of one other firm or the market's expectation [as in Waldman (1984) and others], any outside bid that is based on the bidder's information essentially reveals that information to the current employer. The outside bidder will only outbid the current employer by bidding above the expectation of productivity conditional on both its own information and that of the current employer. As a result, it can do no better than zero profits.

The outside firm would obviously face less of a winner's curse in models where the current employer does not make a counter-offer; but the winner's curse in those cases, which resemble first-price auctions, would still be larger than with a bidding war.

[10] This is Thereom 6 in MW. See Appendix A.

[11] This greatly simplifies my analysis by allowing me to ignore the discounted value of future expected rents, which will vary with employment spell duration and experience. It should not, however, qualitatively affect my results because the *ex post* expected discounted present value of the worker's services is the same regardless of which firm wins. Any option value associated with future auctions would vary with the *ex post* conditional variance of μ, which is the same for both firms, not with the variance conditional on the information each firm possesses prior to the bidding war.

if its signal is higher than firm g's), and will pay a wage equal to firm g's bid, $b(I_{gi})$. Firm f's *ex post* expectation would then be $E(\mu_i|I_{fi}, I_{gi})$, which is greater than $b(I_{gi}) = E(\mu_i|I_{fi} = I_{gi}, I_{gi})$. This equilibrium strategy, therefore, results in positive expected profits for the winning firm.

This equilibrium result is well known in the auction literature and is formally presented in work cited above; however, the intuition is worth repeating. As discussed in Klemperer (1999), 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 wins the bidding above that point, she pays more than her *ex post* expectation. At lower bids, she could further improve her chances of winning at a positive profit by continuing to counter offers. Any firm f, therefore, will continue to counter bids up to $b(I_{fi})$, and drop out of the bidding above that point.

Note that this argument does not depend on information being symmetric. The analysis of this paper is simplified by the fact that the equilibrium in a bidding war is essentially the same under either symmetric or asymmetric information. This result is well known in the auction literature; however, I formalize the argument in the context of my model as Proposition 1 in the next subsection to provide further intuition.

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), but there are good reasons to only consider the symmetric equilibrium in this case. First, 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[12]. The likelihood of there being

[12] This assumes that firms will not bid in order to form reputations for being aggressive. I maintain the same

some degree of match-specific productivity in labor markets makes the symmetric equilibrium a likely approximation of reality[13]. Furthermore, asymmetric equilibria in the current setting would require that firms adopt well known strategies 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[14].

Finally, I should point out that relaxing my assumption of two bidders in each bidding war would only have a qualitative effect on my results as the number of bidders approached infinity. As MW show, an English auction with more than two bidders reduces to an auction between the two bidders with the highest signals in which the signals of all other bidders are common knowledge. As the number of bidders approaches infinity, the bidding war will converge to Bertrand competition with the final bidders having perfect information and the worker being paid a wage equal to her actual productivity {see Klemperer (1999)}. Otherwise, since my model already incorporates some information that is common knowledge between bidders, S_x and (when $t \geq 1$) the wage, allowing more than two bidders in a given bidding war would mostly serve to complicate my presentation.

---

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.

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

[14]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.

Intuitively, it seems the equilibrium in which the current employer always bids aggressively would be similar to the one-shot counter-offer bidding structure used more commonly in this literature. Confirming this intution and examining the question of when one method of bidding for workers would prevail in a market would be interesting, but are well outside of the scope of this paper.

1.2. Bidding for Workers in Equilibrium

1.2.1. Unemployed Workers

Recall that in $t = 0$ an unemployed worker solicits per period wage offers from two firms, each of which collects a signal, ν_{fi}, $f = 0, 1$. Firms are assumed to compete for a worker by engaging in a bidding war.

Applying the result of MW described by equation (2) to this setting, the optimal bid of firm 0 for worker i with experience x is

$$b(\nu_{0i}) = E(\mu | S_{xi}, \nu_{0i}, \nu_{1i} = \nu_{0i})$$

where S_x is the market's information about worker i. In other words, $b(\nu_{0i})$ is the highest wage firm 0 is willing to bid for worker i. The optimal bid of firm 1 follows the same form.

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)

$$\begin{aligned} w_{x1} &= b(\nu_1) = E(\mu | S_x, \nu_0 = \nu_1, \nu_1) \\ &= \frac{\sigma_\nu^2}{\sigma_\nu^2 + 2V_x} \cdot E(\mu | S_x) + 2 \cdot \frac{V_x}{\sigma_\nu^2 + 2V_x} \cdot \nu_1 \\ &< \frac{\sigma_\nu^2}{\sigma_\nu^2 + 2V_x} \cdot E(\mu | S_x) + \frac{V_x}{\sigma_\nu^2 + 2V_x} \cdot \nu_0 + \frac{V_x}{\sigma_\nu^2 + 2V_x} \cdot \nu_1 \\ &= E(\mu | S_x, \nu_0, \nu_1), \end{aligned}$$

where V_x is the variance of μ conditional on S_x. In other words, firm 0 hires the worker at a wage equal to the optimal bid of firm 1, which is less than firm 0's *ex post* expectation of

μ_i, $E(\mu|S_x, \nu_0, \nu_1)$. Firm 0 extracts a positive expected first-period rent of $\frac{V_x}{\sigma_\nu^2 + 2V_x}(\nu_0 - \nu_1)$. Furthermore, note that paying a wage equal to firm 1's optimal bid results in firm 0 observing ν_1.

1.2.2. Employed Workers

Recall that in each period t of an employment spell a new firm, $f = t+1$, that has not previously encountered the worker draws a signal, ν_{t+1}, of the form specified in equation (1), and observes the worker's wage, w_{xt}[15]. Also assume that firm $t+1$ observes the length of the current employment spell, t, and the worker's tenure with her current employer, denoted by τ; but that these variables contain no information that is not already contained in S_x other than the precision of the current employer's information, S_t, and the value of w_{xt} as a signal. 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 current employer from lowering the worker's wage every time an outside firm receives a low signal of her productivity[16]. 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

[15] The worker will always have an incentive to reveal her wage to outside firms. This follows from the well-known result of MW that the seller in an auction maximizes expected revenue by revealing all relevant information.

[16] A worker could easily give her employer a reason to respect this reservation wage policy by committing to quit if the employer deviates. Such a commitment could be motivated either as a punishment mechanism that is optimal in a repeated game, or out of a sense of fairness.

does not qualitatively affect my results[17].

The fact that the outside firm observes the worker's wage decreases but does not eliminate the current employer's informational advantage if $\tau > 1$; however, when $\tau = 1$, w_{xt} is the only bid the current employer has observed. When the outside firm in any period where $\tau = 1$ observes the wage, therefore, information is again symmetric between bidders[18].

In later periods of tenure ($\tau > 1$) w_{xt} could equal any one of multiple bids, each of which was made with the employer possessing a different amount of information, or it could even equal the current employer's expectation (if the expectation fell enough to require the employer to lower the wage). This is discussed in greater detail in the next subsection where I show that the information contained in the current wage, w_{xt}, is a subset of the information contained in S_t. For now, note that bids will be conditioned on w_{xt} and let $\tilde{S}_t$ denote the information in S_t that is not also contained in w_{xt}. Finally, assume that $E\left(\mu | w_{xt}, \tilde{S}_t\right)$ is continuously increasing in both w_{xt} and $\tilde{S}_t$. In the rest of this subsection I will limit my attention to the later periods of tenure in which information is asymmetric.

As mentioned above, the equilibrium strategies in this bidding war are 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, as described by equation (2). The private signal of the

[17] 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.

[18] If $t = 1$ and $\tau = 1$ the current employer's updated signal, S_1, will be a weighted average of ν_0 and ν_1; however, the wage will reveal ν_1 to firm 2. The employer's optimal bid will be

$$b(S_1) = E(\mu | S_x, \nu_1, \nu_0, \nu_2 = \nu_0)$$

and firm 2's optimal bid will be

$$b(\nu_2) = E(\mu | S_x, \nu_1, \nu_0 = \nu_2, \nu_2).$$

If $t > 1$ and $\tau = 1$, then the bids would take a similar form, but ν_1 would be replaced with S_{t-1}.

outside firm is ν_{t+1}, as described by equation (1), and that of the current employer is $\tilde{S}_t$, which is more precise than ν_{t+1}. The following proposition formalizes this equilibrium result.

Proposition 1. *The following strategies form an equilibrium of the bidding war in any period* $t > 1$ *under the assumptions made above:*

1. Firm $t+1$ *will bid for the worker up to its optimal bid,*

$$b\left(\nu_{t+1}\right) = E\left(\mu | S_x, w_{xt}, \tilde{S}_t = \nu_{t+1}, \nu_{t+1}\right),$$

if $b\left(\nu_{t+1}\right) > w_{xt}$. *Otherwise, firm* $t+1$ *simply reveals* ν_{t+1} *to the worker.*

2. The current employer will bid for the worker up to its optimal bid,

$$\begin{aligned} b\left(S_t\right) &= E\left(\mu | S_x, w_{xt}, \tilde{S}_t, \nu_{t+1} = \tilde{S}_t\right) \\ &= E\left(\mu | S_x, S_t, \nu_{t+1} = \tilde{S}_t\right). \end{aligned}$$

Although this proposition is a simple application of a well-known result from MW, Appendix B contains a proof of Proposition 1 that does not refer to the result of MW. The proof takes a simple approach, showing that each bidder's strategy is a best response to the other bidder's strategy, and is intended to give readers more intuition about why this result holds even under asymmetric information.

The result of this bidding process for employed workers is that in each period one of the following occurs:

1. The outside firm's optimal bid is below the current wage, and $b\left(\nu_{t+1}\right)$ is simply revealed to the current employer. The worker's wage remains unchanged unless $w_{xt} >$

$E(\mu|S_{x+1}, S_t, \nu_{t+1})$, in which case $w_{xt+1} = E(\mu|S_{x+1}, S_t, \nu_{t+1})$.

2. The outside firm bids above the current wage, w_{xt}, but the worker is retained at a higher wage, w_{xt+1}, equal to the bid of the outside firm, $b(\nu_{t+1})$.

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

It is important to note that the equilibrium strategies described in Proposition 1 will result in positive expected profits, even for the less well-informed firm. The *ex post* expectation of the worker's productivity is $E(\mu|S_x, S_t, \nu_{t+1})$ regardless of which firm wins. Because firm $t+1$ will only win the bidding war when $\nu_{t+1} > \tilde{S}_t$ this *ex post* expectation will always be greater than $b(S_t)$, the wage firm $t+1$ pays. As S_t increases in precision, however, the weight put on ν_{t+1} in the firms' expectations will decrease, causing the expected profit of firm $t+1$ to decrease as well.

Because the winning bidder always observes the signal of the losing bidder and each worker is equally productive in all firms, all information contained in S_t is passed on to the new employer when the worker is bid away. The precision of the current employer's information, therefore, depends on the length of the worker's current employment spell, not on the worker's tenure with that firm.

Allowing match-specific productivity would likely make the model more realistic, but I maintain the assumption that each worker is equally productive in any firm for two reasons. The more obvious reason is simplicity. Although my model should be able to incorporate match-specific productivity, it would add another level of complexity to employer learning and further complicate notation in the paper[19].

The more important reason for assuming that all productivity is general is that it highlights the strengths of the model. It shows that there can be wage growth that reflects private employer learning and mobility between jobs in a setting where previous models of asymmetric employer learning could not produce either prediction.

1.3. Employer Learning and the Asymmetry of Information

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. In this subsection I first describe the current employer's learning based on the accumulation of these signals. I then discuss the information contained in the worker's wage and show that information is asymmetric between the current employer and an outside firm whenever $\tau > 1$.

1.3.1. The Current Employer's Learning

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[20]. 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

[19] The basic result of MW that I apply allows the value of the good being sold (in this case labor) to include any combination of common and private values. The main complication that allowing some amount of match-specific productivity would add is that 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. This will be discussed along with empirical evidence in Section 4.

[20] The first version of this paper incorporated signals from per period output. Since worker productivity is general and the market does not observe the signals of all bidders, excluding such signals simplifies notation and the timing of events without changing any of the results.

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 $\eta_t \sim N\left(0, \sigma_t^2\right)$, and $\sigma_t^2 = \frac{\sigma_\nu^2}{t+1}$. The reliability of S_t obviously improves (σ_t^2 falls) as the length of the current employment spell, t, increases and more 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. Because information is transmitted to a new employer by the bidding process, the precision of the current employer's information depends on t, not on tenure (τ).

The current employer's conditional expectation of productivity for the worker at any spell 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[21].

1.3.2. Information Provided by a Worker's Wage.

In the first period of tenure ($\tau = 1$) the outside firm knows that the worker's wage was the optimal bid of the previous employer at $t-1$ ($w_{xt} = b(S_{t-1})$ if $\tau = 1$) and information is symmetric between the bidders. If an outside firm observes the worker in her second period of tenure at a job, it learns less from the wage. At $\tau = 2$, w_{xt} would equal $b(\nu_t)$ if the wage was bid up in the previous period. If the outside firm in the previous period did not bid up the worker's wage, w_{xt} would equal either $w_{xt-1} = b(S_{t-2})$ or $E(\mu|S_x, S_t)$, depending on whether $w_{xt-1} < E(\mu|S_x, S_t)$ or not. There is no way for firm $t+1$ to tell which of these three values w_{xt} takes on when $\tau = 2$ because outside firms do not observe a worker's history of wage innovations, or signals received when no bidding occurs. The outside firm, therefore, cannot infer either S_{t-2} or ν_t from w_{xt} at $\tau = 2$[22].

In general, a worker's wage is the minimum of the employer's expectation and an increasing sequence resulting from the bids of outside firms:

$$w_{xt} = \min\left[E(\mu|S_x, S_t), \bar{b}\right], \tag{6}$$

where $\bar{b} = \max[b(S_{t-\tau+1}), ..., b(\nu_t)]$ (or $\max[b(\nu_1), ..., b(\nu_t)]$) if the wage was never lowered by the current employer, or $\bar{b} = \max[E(\mu|S_{x-\psi}, S_{t-\psi}), b(\nu_{t-\psi+1}), ..., b(\nu_t)]$ if the current em-

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

[22] Knowing that the wage was an unknown firm's bid does not allow that firm's bid to be inferred unless it is known when that bid was made because the weight a bid puts on a signal changes as t increases and S_t becomes more precise.

ployer last lowered the worker's wage in period $t - \psi$. Needless to say, it is difficult at best to describe the information that an outside firm $t+1$ could infer from w_{xt}; however, it clearly contains less information than S_t whenever $\tau > 1$[23]. While the current employer observes all of the signals contained in S_t, the outside firm cannot infer any of them if $\tau > 1$.

Fortunately, my model does not require a precise characterization of the information contained in w_{xt}. All it requires is that w_{xt} provides information about productivity such that $E\left(\mu | w_{xt}, \tilde{S}_t\right)$ is continuously increasing in w_{xt} and $\tilde{S}_t$. This condition would be met regardless of whether outside firms could interpret equation (6) precisely or just followed a "rule of thumb" when inferring information about a worker, as long as that rule of thumb was common knowledge[24].

1.4. The Sequence of Wages

As a result of the sequence of bidding wars described above, the worker's wage, described by equation (6), increases toward the current employer's expectation unless the wage has to be lowered so that employing the worker is not unprofitable, at which point $w_{xt} = E(\mu | S_x, S_t)$. Once the wage is lowered, it can increase again only if the difference between $E(\mu | S_x, S_t)$ and w_{xt} temporarily increases. Despite such deviations from a monotonic convergence, I establish in the following proposition that, although wages at t do not generally equal $E(\mu | S_x, S_t)$, they converge to $E(\mu | S_x, S_t)$ as t goes to infinity. This is important because it means that as the length of uninterrupted employment increases competition from less well-informed firms

[23] Even if outside firms observed the entire history of wage innovations on the current job, they would still have less information than is in S_t because it would not observe the signals that were too low to result in a wage innovation.

[24] For example, bounded rationality could keep outside firms from inferring anything from w_{xt} other than a lower bound for S_t.

causes a worker's wage to resemble her current employer's expectation more and the market's expectation less. As a result, wages reflect the private learning of current employers, despite their informational advantage.

Proposition 2. *The sequence of bidding wars for a worker with t periods of uninterrupted employment creates a sequence of wages, w_{xt}, 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.*

Appendix B contains the proof of this proposition. The intuition behind the proof is as follows: If the worker has not experienced a wage cut during an employment spell, or the wage cuts took place early in the employment spell, the result follows from a monotonic convergence argument. If, on the other hand, wage cuts continue to occur later in the employment spell, the difference between the wage when the last decrease took place, $w_{xt-\gamma} = E(\mu|S_{x-\gamma}, S_{t-\gamma})$, and the current expectation, $E(\mu|S_x, S_t)$, converges to zero.

Proposition 2 shows that competition from less well-informed firms forces employers in my model to raise the wages of workers toward their expectations of the workers' productivity. Since an employer's expectation converges to the worker's productivity, the worker's wage will also converge to her productivity. In the next section, I show that this implication is testable using an extension of the test of public learning developed by AP. In short, my model predicts that wages will reflect evidence of employer learning as the length of the employment spell increases. Section 4 presents the estimation results from the empirical test of this prediction.

Proposition 2 also makes an important contribution to the theoretical literature on asymmetric employer learning by showing that wages can grow toward the employer's expectation even in the absence of promotions or other publicly observable signals. Two workers with exactly the same value of S_x in this model will not generally have the same wage. The inability of previous asymmetric learning models to produce this result is one of the main criticisms of the literature made by Gibbons and Waldman (1999).

The model in this paper is also consistent with some of the important and well known empirical findings that have motivated much of the previous work on asymmetric employer learning and careers in organizations. For example, wages will rise on average with seniority in my model, even though productivity does not, as noted by the studies of Medoff and Abraham (1980, 1981) and Baker, Gibbs and Holmstrom (1994a, 1994b). This happens because wages on a job start out being lower than the employer's expectation of productivity but are bid up toward the employer's expectation. Since the employer's information is based on a sequence of randomly drawn, unbiased signals and the worker's ability is fixed in the model, the employer's expectations will appear not to increase on average as tenure (or experience or employment spell duration) increases. The wage, on the other hand, will be bid up toward that expectation as long as the worker remains continuously employed[25].

The model also incorporates real wage cuts, but predicts that they will not occur as often as wage increases, as documented by BGH. Because wages are typically lower than the worker's expected productivity and are only reduced when it is necessary to keep them from exceeding the worker's expected productivity, wage cuts will occur less frequently in my model than in the public employer learning model of FG or other learning models in which wage equals expected productivity. Wage cuts will occur, but not every time an employer receives a negative signal

[25]Cases in which tenure (as opposed to spell length) is unusually long might be exceptions to this case. Without the addition of match-specific productivity or an increase in human capital exceptionally long tenure in this model would likely be due to high positive errors on initial signals that would then be followed by falling wages. At average levels of tenure, however, the affect of wages being bid up toward the current employer's unbiased expectation would still cause wages to rise with tenure.

about the worker.

Finally, this model has implications for worker mobility that could be compared to other literatures in labor economics, and these implications deserve a more formal treatment in future research. One of the more intuitive implications for mobility is that the average wage increase associated with job-to-job transitions should decrease as the spell length increases. Jovanovic (1984) produces a similar result with increases in tenure instead of spell length, but the reasoning behind the results is different. In his model, workers become more willing to leave for a lower wage as tenure increases due to higher relative option values of new jobs. In my model, the wage increases are due to the difference between the employer's expectation of the worker's productivity and the wage the worker is paid, which is decreasing on average as spell length increases[26].

2. Estimation of Wage Equations under Asymmetric Learning

The prediction that wages converge to the current employer's expectation of the worker's productivity, despite the employer's private information, allows this model to be tested using a simple extension of the work of AP. When all learning is public, as in AP, wages equal expected productivity and become more correlated with the worker's actual productivity as experience accumulates and the market's expectation becomes more accurate. In my model of asymmetric employer learning, wages become more closely related to actual productivity as the length of the current employment spell increases due to both the wage converging to the current employer's expectation (see Proposition 2) and that expectation becoming more accurate as the employer

[26] Eeckhout (2005) develops a model that is very similar to mine with the express intent of comparing its results under general human capital to the results other models achieve under match-specific productivity.

accumulates more private information. In either case, when the wage becomes more correlated with the worker's actual productivity, it becomes more correlated with variables that are correlated with productivity but difficult for employers to observe and less correlated with easily observed variables.

Because the model developed in the previous section includes a term that captures the market's learning, the wage regression developed in this section tests for both public learning and asymmetric learning. Any learning that is public should result in evidence of learning over experience. On the other hand, evidence of learning over the length of the current employment spell suggests asymmetric learning[27].

In what follows, Z denotes a vector of easily observed variables, such as education or race, that are correlated with productivity. I assume that Z is related to productivity through a linear function, $f(Z) = Z\delta$, such that

$$Z\delta = \mu + \zeta;$$

where $\zeta \sim N\left(0, \sigma_Z^2\right)$ and neither $f(Z)$ nor σ_Z^2 varies by labor market experience, the length of the current employment spell, or tenure. I assume that Z, δ, and σ_z^2 are all common knowledge in the labor market. Adding these easily observed variables to my model, therefore, in no way affects the results discussed above. They were left out of the discussion of the theoretical model for the sake of notational simplicity.

To see how wages evolve, first recall that a worker's wage, w_{xt}, is always less than or equal

[27] Recall that experience and the length of the current employment spell differ in this model due to the assumption of exogenous job destruction; however, any assumption that resulted in periods of unemployment would produce the same result.

to the current employer's conditional expectation. On average, the difference between w_{xt} and $E(\mu|Z, S_x, S_t)$ 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, $w_{x'1} = E(\mu|Z, S_{x'}, \nu_0 = \nu_1, \nu_1)$, where $x' = x - t$ 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

$$E(w_{xt}) = [1 - \rho(t)] \cdot E(\mu|Z, S_{x'}, \nu_0 = \nu_1, \nu_1) + \rho(t) \cdot E(\mu|Z, S_x, S_t), \tag{7}$$

where $\rho(t)$ is a monotonically increasing, differentiable 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. By expanding the expectations and recalling that $S_x = \mu + \eta_x$, $S_t = \mu + \eta_t$, and $\nu_1 = \mu + e_1$, the expected wage can be written as a weighted average of the population mean, m; an easily observed variable, Z; actual productivity, μ; and an error term:

$$E(w_{xt}) = B_m m + B_Z(Z\delta) + B_{xt}\mu + \phi'. \tag{8}$$

(See Appendix C for a more complete description of equation 8.)

Due to the public component of learning, B_m and B_z are decreasing in experience while B_{xt} is increasing in experience. More important in the current context, my model implies that B_{xt} increases with t and B_m and B_Z decrease with t. (The relevant derivatives are presented in Appendix C.) This follows because the current employer's expectation of the worker's productivity converges to μ as t increases, while competition from less well-informed firms in the market causes the worker's wage to converge to that expectation as t increases.

Since the worker's actual productivity obviously can't be observed, my estimation instead uses a variable that is correlated with productivity but is unlikely to be observed by employers, as do AP and FG. 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 ζ, η_x and η_t, the error terms in $f(Z)$ and the updated signals. When μ is replaced by V in equation (8), the expectation of the OLS estimate of B_{xt} from the resulting regression is

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

Since the degree of bias, $\frac{\sigma_{V\mu}}{\sigma_V^2}$, does not vary with experience or employment spell length, it does not interfere with the model's basic predictions.

In the actual estimation, I approximate the coefficients in equation (8) with linear interactions of x and t, resulting in a wage equation of the form

$$w_{xt} = C + Z\gamma_0 + Z \cdot x\gamma_x + Z \cdot t\gamma_t + VB_0 + V \cdot xB_x + V \cdot tB_t + \phi'_{xt}, \tag{9}$$

where C is a vector containing a constant and the experience and spell length terms[28]. Any learning that is public results in γ_x being negative and B_x being positive. If there is asymmetric

[28]The effects of the interactions of m with t and x obviously can't be seperated from any effects t and x have on the wage directly.

learning, γ_t will be negative and B_t will be positive.

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[29]. Without the interactions of the hard-to-observe variable, 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 (9). Although other factors could determine the sign of γ_x and γ_t, these variables should nonetheless fall (become less positive or more negative) when the interactions of the hard-to-observe variable are added to the regression.

3. 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

[29] 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.

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 experimented with other definitions of an employment spell, but the estimation results were always qualitatively similar[30].

I also create a measure of tenure that is consistent with my measure of employment spell length. The tenure variable that is included in the NLSY counts all weeks between the start of the job and either the date the job ended or the interview date, regardless of whether the worker was employed or not. Since my measure of spell length only counts weeks working, I define tenure at a job as weeks worked between the start date and either the end date or the interview date. Even though this could count some weeks in which the worker was employed by other firms, the resulting tenure variable never exceeds the spell length, unlike the standard tenure variable in the NLSY[31].

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,

[30] 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.

[31] The tenure measure I create is highly correlated with the standard tenure measure, with a highly significant Pearson correlation coefficient of 0.977. Furthermore, my tenure measure and the standard NLSY measure are similarly correlated with spell length, with significant correlation coefficients of tenure and spell length in either case falling between 0.77 and 0.78.

are not in the military and have completed at least 12 years of schooling[32]. I use observations from the CPS job in each year if its hourly wage (in 1987 dollars) is between $2 and $200, it involves between 35 and 100 hours worked per week, and there are no missing values for any of the key variables (not counting tenure). I exclude 2537 observations of workers who had been out of the labor market at least 25% of their career up to that interview. Finally, I impose two restrictions intended to improve the reliability of my experience measures: I drop 12 observations from people who had more than 4 years of potential experience in 1979, and I drop 2825 observations in which actual experience is calculated to exceed potential experience by more than one year[33]. Eligible observations are drawn from all years of the survey (1979-2000). The resulting sample has 30,374 valid observations for 3,677 men.

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 in the population of workers in my sample (but not the full panel) 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 1987 dollars, is $9.88. Almost 70% of the sample is white and just over 76% resides in an urban area. The average worker has completed 13.25 years of schooling. The average worker has a tenure on the CPS

[32] I experimented with using the definitions of labor market entry used by FG and AP, as well as extending the sample to include men who had complete 8 years or more of education and found qualitatively similar results in each case.

[33] This last group appears to consist mostly of people who had a missing interview as they transitioned from school, or who reported an average of more than 52 weeks worked per year for multiple consecutive years.

job of 3.6 years, but has been continuously employed for 5.1 years. The average worker has 18.8 years of potential experience and 15.2 years of actual experience at the 2000 interview. Over all years in the sample, the average potential experience is 10 years, while the average actual experience is 8 years.

4. Estimation Results

The estimation results presented in this section are consistent with my model of asymmetric employer learning. For the sake of comparison, this section first presents results from regressions estimated under the assumptions of a pure public learning model before presenting the tests of asymmetric employer learning. The results suggest that a significant portion of the evidence of learning observed in tests of public learning may actually be due to asymmetric employer learning. Despite the current employers' informational advantage, the effects of private learning are reflected in the wage regressions at least as strongly as the effects of public learning.

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[34]. 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[35].

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

[35] The use of testing by firms does not suggest that AFQT is not itself difficult to observe. Even employers that give their own tests will not actually observe the AFQT score. Their test will be another noisy signal of worker productivity that will be included in S_{tj}. The results presented in this and other papers suggest that

Table 2 presents results from wage regressions estimated under the assumptions of a public 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. It is important to use an instrument in this case because actual experience is likely correlated with ability, which could cause bias in the coefficients on experience and its interactions. If AFQT only captures part of a worker's ability or is a noisy measure of ability, then part of the effect of ability on wages could be picked up by actual experience. Furthermore, actual experience could be used by employers to learn about a worker's ability. In any case, potential experience is correlated with actual experience, but should not be correlated with ability, making it a valid instrument.

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 [1.05 (0.09) for OLS, 0.96 (0.09) 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 0.39 (0.12) in the IV regressions, a statistically significant decrease in both cases. The coefficient on AFQT $\times$ the experience measure is significantly positive at 0.060 (0.012) in the OLS regressions and even larger [0.070 (0.015)] in the IV regressions. As predicted, coefficients on the easily observed variables interacted with experience always become more negative (or less positive) when AFQT $\times$ the experience

---

employer-provided testing does not allow employers to observe AFQT scores. If it did, AFQT would not behave like a hard-to-observe variable in wage regressions.

AP also use father's education as a hard-to-observe variable. Because parental education might affect productivity in ways that are observable, like language development, I do not use it in this paper. Preliminary estimation (not shown) supports this decision.

measure is added, as predicted, 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 support 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 both 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 significantly from 0.988 (0.090) to 0.397 (0.120). More importantly, the coefficient on AFQT $\times$ employment spell length in column III is 0.054 (0.022). The coefficient on AFQT $\times$ potential experience is slightly smaller but also significant at 0.031 (0.015).

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 \$0.15. Despite the informational asymmetry, the current employer's private learning appears to affect wages at least as much as the market's public learning as long as the worker remains employed.

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[36], and the coefficients on spell length and its interactions could be

[36] For example, a worker who has been continuously employed for a long time could be more able or disciplined than an otherwise similar worker whose labor market experience consists of a series of short employment spells.

biased for the same reason as those on actual experience and it's interactions. The length of an employment spell could also be correlated with a match-specific component that affects the wage on the current job for the same reasons that higher tenure on a job could be[37].

In an attempt to create an instrument that is not correlated with information about the worker's productivity or a match-specific component of the wage, I regress the length of the current employment spell on the worker's career average spell length, actual experience, the total duration of the current job, and a dummy variable for missing values of duration. As long as these variables control for all components of employment spell length that are correlated with productivity[38], and duration of the current job controls for any match-specific components that are correlated with the residual in the wage equation, the residual from this regression is a valid instrument for employment spell length[39]. 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 0.084 (0.034), while that on AFQT $\times$ actual experience is only 0.027 (0.025). According to these estimates, a one standard deviation increase in the adjusted AFQT score increases hourly wages by \$0.41 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

[37]A high match value would make the job more valuable, resulting in both higher wages and a higher expected job duration. Since longer employment spells are often the result of greater tenure on the job, employment spell length could also be reflecting match-specific components of the wage. See, for example, Altonji and Shakotko (1987) or Abraham and Farber (1987) for more detailed discussions endogenous tenure.

[38]This should include anything that the market learns about productivity from observing spell duration, including things like "discipline." If more disciplined workers have longer spell lengths, then the career average spell length should capture that.

[39]The R^2 for this first-stage regression is 0.78. The estimated coefficients are as follows: Constant, -1.93 (0.06); Average Spell Duration, 0.48 (0.01); Actual Experience, 0.35 (0.01); Duration of Current Job, 0.51 (0.01); and Job Duration Missing, 1.69 (0.16).

in AFQT by a statistically insignificant \$0.13. Furthermore, the coefficient on AFQT again falls significantly, from 0.934 (0.094) to 0.276 (0.133), when the interactions of AFQT with experience and spell duration are added. As in all of my estimation, the interactions of grade and race with experience and spell length have effects that are consistent with both public and private learning; however, the change in their coefficients when the interactions with AFQT are added is never significant.

The model presented in the current paper implies that the private information of one employer is passed on to the next whenever a worker is bid away by a new employer. This transmission of information is the result of firms engaging in bidding wars for the worker's services combined with the assumption that worker productivity is not match-specific. If employers do not bid for workers through bidding wars, or some other mechanism that allows the winning bidder to observe the bid of the losing bidder, then little or no information will be transmitted between employers. If match-specific productivity is important, then the previous employer's ultimate bid will be less meaningful for the new employer because it serves only as a signal of general productivity, with the match-specific component acting as an additional error term.

To consider this issue further, I divide time in the current spell into tenure with the current employer and time worked in the spell with other employers. If only part of the previous employers' information is transmitted to the new employer, our intuition suggests there will be more evidence of employer learning with tenure on the current job than with time worked in the spell with other employers, because part of the information accumulated by the previous employer(s) was lost when the worker took the current job. Of course, if no information is transmitted from one employer to the next, the model would predict only evidence of learning over experience and tenure, as in the recent paper by Schoenberg (2004)[40]. In either of these cases, the regression estimates I've presented so far might be providing evidence of employer learning with the length of the employment spell simply because employment spell length and tenure are highly correlated.

I examine this possibility in Table 5 and find no evidence that would lead one to reject my model's implication that private information accumulated by one employer is transmitted to the next when workers move from one job to another. Table 5 repeats the analysis of Table 3 with spell length divided into time in the spell before the current job (Spell-Tenure) and tenure. The coefficient on AFQT x (Spell-Tenure) is 0.071 (0.032), providing significant evidence that is consistent with information being transmitted between employers in a job-to-job transition. The coefficient on AFQT x Tenure is also significantly positive [0.047 (0.026)]; however, the fact that it is not larger than the coefficient on AFQT interacted with (Spell-Tenure) give us no reason to believe that the transmission of information between employers is limited by match-specific productivity or different bidding mechanisms[41]. In an alternative specification (not shown) I included interactions with both the full spell length and tenure, and found no statistically significant evidence of learning over tenure but still found significant evidence consistent with learning over spell duration[42].

Finally, it is important to note that my estimation, like that in AP or FG, could be affected

[40] Schoenberg (2004) presents a two-period model in which outside firms bid against the perfectly-informed current employer, who makes a single counter-offer in the second period. This bidding is profitable for the outside firm, and results in the current employer offering a wage that incorporates some of its private information, under the assumption that there is a transitory nonpecuniary value associated with each job.

[41] Results from IV specifications were qualitatively similar.

[42] I also estimated specifications that only include interactions with experience and tenure but found no statistically significant evidence of learning over tenure, although the coefficients had the predicted signs. Schoenberg (2004) reports evidence of learning over tenure from wage regressions, but her only significant coefficient on AFQT x Tenure comes from a specification that does not allow the effect of education or race to vary with tenure, making it unclear whether or not AFQT x Tenure is picking up part of the effects of Grade or Race x Tenure.

by human capital accumulation if AFQT scores are correlated with the returns to training. Following AP, I repeated my OLS and IV estimation with measures of both current and accumulated employer-provided training added to the regressions. Although the measure of training in the NLSY is far from perfect[43], one would at least expect controlling for training to reduce the evidence of learning if that evidence is biased upward by a correlation between training and ability. The results of the regressions with training added (not shown) indicate that controlling for training further reduces the evidence of public learning but has no noticeable effect on the evidence of asymmetric learning.

5. Conclusions and Directions for Future Research

This paper develops and tests a model of asymmetric employer learning that relaxes the informational assumptions of most papers in the literature and allows firms to compete for workers through bidding wars. As a result, outside firms can profitably compete 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 employer by less well-informed firms even though there is no match- or firm-specific productivity. Furthermore, competition from outside firms forces the current employer to raise the worker's wages toward the employer's expectation of that worker's productivity, resulting in different wages for workers who have the same publicly observable characteristics.

This convergence of wages to the current employer's expectation allows the model to be

[43] The NLSY does not ask about training that went on for a month or less until 1988. I use observations from 1988 on to predict training in earlier years, as do AP. This prediction is based on a probit estimate using a flexible function of grade, AFQT, experience, spell length and tenure, as well as controls for urban residence and the first occupation after leaving school.

tested empirically. The model implies that wages reflect evidence of public learning as experience increases and asymmetric learning as the length of the employment spell length increases. Extending the work of Altonji and Pierret (2000), I develop a simple test of the model that requires only the estimation of basic wage regressions using data from the NLSY. The resulting estimation suggests that both public learning and private, asymmetric learning affect wages. In fact, competition from outside firms appears to be 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. Finally, I estimate a version of these wage regressions that tests the transmission of information between employers that is implied by the model, and find no evidence that contradicts that implication.

This paper opens multiple avenues for future research. First, the basic framework of this model could be used to revisit issues discussed in the literature on task assignment. The model demonstrates how competition from less well-informed firms can reduce the current employer's ability to pay the worker less than her expected productivity. Such competition would also reduce the incentive for employers to keep workers in lower paying jobs in order to avoid signalling the worker's ability to outside firms. Many of the results of previous models would be preserved; however, promotions would have different effects at different points in a worker's career. Such a combined model could provide new testable implications as well as a means of explaining various empirical observations made previously about workers' careers in firms.

The model also has more implications for wage and especially 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 might be aided by embedding this model into a basic search framework[44]. Adding a fuller model of job search would provide a more realistic description of unemployment than the model has in its current form, and would allow an analysis of how employer learning affected search intensity, for example.

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 (2003) examine empirical evidence of different types of statistical discrimination in public learning frameworks[45]. Although the results in this paper provide weak support (if any) for a "rational stereotyping" form of statistical discrimination, it is possible that black and white men differ in the effect asymmetric learning has on their wages relative to public learning. Exploring that possibility could provide new insights into how discrimination arises and why it persists.

[44] 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.

[45] 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 (2003), Oettinger (1996) and Milgrom and Oster (1987) consider screening discrimination.

References

[1] Abraham, Katherine G. and Henry S. Farber. “Job Duration, Seniority, and Earnings,” *The American Economic Review*, 77(3), 1987.

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

[3] Altonji, Joseph G. and Robert A. Shakotko. “Do Wages Rise with Job Seniority?” *Review of Economic Studies*, 1987.

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

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

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

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

[8] Bikhchandani, Sushil, and John G. Riley. “Equilibrium in Open Auctions,” UCLA, 1993.

[9] DeVaro, Jed and Michael Waldman. "The Signalling Role of Promotions: Further Theory and Empirical Evidence," Cornell University, 2004.

[10] Eeckhout, Jan. "Employer Learning and General Human Capital," University of Pennsylvania, 2005.

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

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

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

[14] Jovanovic, Boyan. “Matching, Turnover, and Unemployment,” *Journal of Political Economy*, vol. 92(1), 1984.

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

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

[17] Lazear, Edward P. "Raids and Offer Matching," in R. G. Ehrenberg, ed., *Research in Labor Economics,* vol. 8, 1986, pp. 141-165.

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

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

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

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

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

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

[24] 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.

[25] Pinkston, Joshua C. "Screening Discrimination and the Determinants of Wages." *Labour Economics*, December, 2003.

[26] 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.

[27] Schoenberg, Uta. "Testing for Asymmetric Employer Learning", University of Rochester, 2004.

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

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

[30] Waldman, Michael. "Up-or-out Contracts: A Signalling Perspective," *Journal of Labor Economics,* 8, 1990, pp. 230-250.

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. Proofs of Propositions 1 and 2

B.1. Proof of Proposition 1

I first show that the strategy of firm $t+1$ described in the proposition is a best response to the current employer bidding up to $b(S_t)$. If the current employer bids up to $b(S_t)$ and firm $t+1$ outbids it, firm $t+1$ will observe S_t through $b(S_t)$ and it's profit in period $t+1$will be its *ex post* expectation minus $b(S_t)$:

$$\pi_{t+1} = E\left(\mu | S_x, w_{xt}, \tilde{S}_t, \nu_{t+1}\right) - E\left(\mu | S_x, w_{xt}, \tilde{S}_t, \nu_{t+1} = \tilde{S}_t\right).$$

Obviously, $\pi_{t+1} > 0$ for every $\nu_{t+1} > \tilde{S}_t$, and $\pi_{t+1} < 0$ for every $\nu_{t+1} < \tilde{S}_t$. Recall that all of these expectations, including $b(\nu_{t+1})$, are continuous and increasing in all of their terms.

For any final bid $b < b(\nu_{t+1})$, there are values of $\tilde{S}_t$ such that $\nu_{t+1} > \tilde{S}_t$ ($\pi_{t+1} > 0$) and $b(S_t) > b$; thus, setting the final bid below $b(\nu_{t+1})$ lowers the probability that firm $t+1$ will win the bidding at a positive profit. For any final bid $b > b(\nu_{t+1})$, there are values of $\tilde{S}_t$ such that $\nu_{t+1} < \tilde{S}_t$ ($\pi_{t+1} < 0$) and $b(S_t) < b$; thus, if firm $t+1$ decides to bid above $b(\nu_{t+1})$ it introduces the possibility of winning at a negative profit without increasing it's chances of earning a positive profit. Finally, note that when $b(\nu_{t+1}) < w_{xt}$ firm $t+1$ can never earn a positive profit from bidding. No course of action will make it better offer than accepting zero profit and costlessly revealing its signal. Therefore, firm $t+1$'s best response to $b(S_t)$ is to bid for the worker up to $b(\nu_{t+1})$, if $b(\nu_{t+1}) > w_{xt}$; and simply reveal ν_{t+1} to the worker otherwise.

I now need to show that bidding up to $b(S_t)$ is the current employer's best response to firm $t+1$ bidding up to $b(v_{t+1})$ when $b(\nu_{t+1}) > w_{xt}$. If the current employer wins the bidding war in this case, it will observe ν_{t+1} and it's profit will be

$$\pi_{t+1} = E\left(\mu | S_x, w_{xt}, \tilde{S}_t, \nu_{t+1}\right) - E\left(\mu | S_x, w_{xt}, \tilde{S}_t = \nu_{t+1}, \nu_{t+1}\right).$$

The argument here follows the same logic that was used above: If the current employer sets it's ultimate bid below $b(S_t)$ it will fail to win the worker's services in some cases that would have yielded positive profits. If the current employer decides that it will stop bidding at some point above $b(S_t)$, it introduces the possibility of winning at a negative profit. ■

B.2. Proof of Proposition 2

First note that the wage, w_{xt}, as described in equation (6), is bounded above by $E(\mu|S_x, S_t)$, and $\lim_{t\to\infty} E(\mu|S_x, S_t) = \mu$.

Consider the case in which there are no wage cuts first. If the worker is at her first and only job of the employment spell,

$$\bar{b} = \max\left[b(\nu_1), ..., b(\nu_t)\right].$$

The wage then is a monotonically increasing sequence that is bounded above by $E(\mu|S_x, S_t)$. Since $E(\mu|S_x, S_t)$ is itself bounded, the monotone convergence theorem implies that the wage converges to $E(\mu|S_x, S_t)$. Furthermore, it is easy to see that the wage cannot converge to anything less than $E(\mu|S_x, S_t)$ because whenever $w_{xt} < E(\mu|S_x, S_t)$ there is a positive probability that another firm will place a bid between w_{xt} and $E(\mu|S_x, S_t)$. Therefore, in this case w_{xt} converges to $E(\mu|S_x, S_t)$.

If the worker has had previous employers during this employment spell,

$$\bar{b} = \max\left[b(S_{t-\tau+1}), ..., b(\nu_t)\right].$$

The same monotonic convergence argument used above applies here. Furthermore, for any τ,

$$|E(\mu|S_x, S_t) - b(S_{t-\tau+1})| \to 0 \text{ as } t \to \infty$$

because $b(S_{t-\tau+1}) = E\left(\mu | S_{x-\tau+1}, S_{t-\tau+1}, \nu_{t-\tau+2} = \tilde{S}_{t-\tau+1}\right)$ and $E(\mu|S_x, S_t)$ converges to μ; thus, as job changes occur later in the employment spell, the starting wages on those jobs also converge to $E(\mu|S_x, S_t)$.

The more interesting case is when the wage is not monotonically increasing; however, even when there are wage cuts there are still two effects ensuring that w_{xt} converges to $E(\mu|S_x, S_t)$. First, if the wage cut took place in a fixed period in the past and the wage is not lowered again, then the monotonic convergence argument used above again applies to the sequence of wages that follows that period. Second, if these wage cuts continue to occur as t increases, convergence still occurs because

$$|E(\mu|S_x, S_t) - E(\mu|S_{x-\psi}, S_{t-\psi})| \to 0 \text{ as } t \to \infty$$

for any fixed ψ. (This follows from the fact that $E(\mu|S_x, S_t)$ is a convergent sequence.) ■

C. Equation 8 and its Coefficients.

With Z added to the model the current employer's expectation of the worker's productivity is simply

$$E(\mu|Z, S_x, S_t) = \beta_m m + \beta_Z (Z\delta) + \beta_x S_x + \beta_t S_t,$$

where $\beta_m = \frac{\sigma_x^2\sigma_t^2\sigma_Z^2}{D}$, $\beta_Z = \frac{\sigma_x^2\sigma_t^2\sigma_\mu^2}{D}$, $\beta_x = \frac{\sigma_t^2\sigma_Z^2\sigma_\mu^2}{D}$, $\beta_t = \frac{\sigma_x^2\sigma_Z^2\sigma_\mu^2}{D}$, and $D = \sigma_x^2\sigma_t^2\sigma_Z^2 + \sigma_x^2\sigma_t^2\sigma_\mu^2 + \sigma_\mu^2\sigma_t^2\sigma_Z^2 + \sigma_\mu^2\sigma_x^2\sigma_Z^2$. The first wage the worker was paid in the current employment spell is

$$w_{x'1} = E(\mu|Z, S_{x'}, \nu_0 = \nu_1, \nu_1) = \frac{\sigma_{x'}^2\sigma_\nu^2\sigma_Z^2}{D'}m + \frac{\sigma_{x'}^2\sigma_\nu^2\sigma_\mu^2}{D'}(Z\delta) + \frac{\sigma_\mu^2\sigma_\nu^2\sigma_Z^2}{D'}S_{x'} + \frac{2\sigma_\mu^2\sigma_{x'}^2\sigma_Z^2}{D'}\nu_1,$$

where $D' = \sigma_{x'}^2\sigma_\nu^2\sigma_Z^2 + \sigma_{x'}^2\sigma_\nu^2\sigma_\mu^2 + \sigma_\mu^2\sigma_\nu^2\sigma_Z^2 + 2\sigma_\mu^2\sigma_{x'}^2\sigma_Z^2$.

Inserting these expectations, equation (7) can be rewritten as

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

Recalling that $S_x = \mu + \eta_x$, $S_t = \mu + \eta_t$, and $\nu_1 = \mu + e_1$, this equation is easily simplified as

$$\begin{aligned} E(w_{xt}) &= B_m m + B_Z (Z\delta) + B_{xt}\mu + \phi', \text{ where} \qquad (8) \\ B_m &= (1-\rho(t)) \cdot \frac{\sigma_{x'}^2\sigma_\nu^2\sigma_Z^2}{D'} + \rho(t) \cdot \beta_m, \\ B_Z &= (1-\rho(t)) \cdot \frac{\sigma_{x'}^2\sigma_\nu^2\sigma_\mu^2}{D'} + \rho(t) \cdot \beta_Z \\ B_{xt} &= (1-\rho(t)) \cdot \left(\frac{\sigma_\mu^2\sigma_\nu^2\sigma_Z^2}{D'} + \frac{2\sigma_{x'}^2\sigma_\mu^2\sigma_Z^2}{D'}\right) + \rho(t) \cdot (\beta_x + \beta_t), \text{ and} \\ \phi' &= (1-\rho(t)) \cdot \left(\frac{\sigma_\mu^2\sigma_\nu^2\sigma_Z^2}{D'}\eta_{x'} + \frac{2\sigma_{x'}^2\sigma_\mu^2\sigma_Z^2}{D'}e_1\right) + \rho(t) \cdot (\beta_x\eta_x + \beta_t\eta_t). \end{aligned}$$

The rest of this section will show the derivatives of B_m, B_Z and B_{xt} in equation (8) in order to support the conclusions of Section 1.

C.1. Derivatives with Respect to Experience

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

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

because $\frac{\partial \sigma_{x'}^2}{\partial x}<0$ and $\frac{\partial \beta_m}{\partial x}=\frac{\partial \sigma_x^2}{\partial x}\frac{\left(\sigma_t^2\sigma_Z^2\right)^2\sigma_\mu^2}{D^2}<0$.

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

because $\frac{\partial \sigma_{x'}^2}{\partial x}<0$ and $\frac{\partial \beta_Z}{\partial x}=\frac{\partial \sigma_x^2}{\partial x}\frac{\left(\sigma_t^2\sigma_\mu^2\right)^2\sigma_Z^2}{D^2}<0$.

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

The first term in $\frac{\partial B_{xt}}{\partial x}$ is clearly positive. The second term is positive as well: $\frac{\partial \beta_x}{\partial x}+\frac{\partial \beta_t}{\partial x}=-\frac{\partial \sigma_x^2}{\partial x}\frac{\left(\sigma_t^2\sigma_Z^2+\sigma_t^2\sigma_\mu^2\right)\sigma_\mu^2\sigma_t^2\sigma_Z^2}{D^2}>0$. Therefore, $\frac{\partial B_{xt}}{\partial x}>0$.

C.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 and B_Z decrease:

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

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

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

because $\beta_Z < \frac{\sigma_{x'}^2 \sigma_\nu^2 \sigma_\mu^2}{D'}$ and $\frac{\partial \beta_Z}{\partial t} < 0$.

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

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., $(\beta_x + \beta_t) > \left(\frac{\sigma_\mu^2 \sigma_\nu^2}{D'} + 2 \cdot \frac{\sigma_{x'}^2 \sigma_\mu^2}{D'}\right)$. The second term is positive because $\left|\frac{\partial \beta_x}{\partial t}\right| < \left|\frac{\partial \beta_t}{\partial t}\right|$. (The algebra is analogous to the case of $\left(\frac{\partial \beta_x}{\partial x} + \frac{\partial \beta_t}{\partial x}\right)$ above.)

Table 1. Summary Statistics*.

	Mean	Std. Dev.	Min	Max
Real Hourly Wage	9.881	6.517	2.002	192.054
Highest Grade Completed	13.253	1.873	12	20
White	0.699	0.459	0	1
Black	0.258	0.437	0	1
Normalized AFQT	-0.104	0.982	-2.018	2.361
Employment Spell Length	5.134	4.464	0	23.019
Tenure	3.608	3.760	0	22.846
Potential Experience	10.037	5.311	0	25
Experience	8.029	4.985	0.019	24.615
Pot. Experience in 2000	18.827	3.167	9	25
Experience in 2000	15.158	5.055	0.231	24.615
Urban	0.762	0.426	0	1

*Notes: Wages are in 1987 dollars. Spell length, tenure and the experience variables are measured in years. There are 30374 observations except for Tenure (29639 nonmissing), and the experience measures in 2000 (2150).

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

	OLS		IV	
	I	II	I	II
Grade	0.930	1.079	0.963	1.106
	(0.135)	(0.136)	(0.141)	(0.142)
Grade x Experience	-0.002	-0.017	0.003	-0.015
	(0.018)	(0.018)	(0.023)	(0.023)
White	-0.534	-0.006	-0.611	-0.106
	(0.193)	(0.199)	(0.191)	(0.197)
White x Experience	0.137	0.085	0.153	0.089
	(0.017)	(0.019)	(0.022)	(0.025)
AFQT	1.049	0.446	0.956	0.387
	(0.093)	(0.122)	(0.093)	(0.122)
AFQT x Experience	…..	0.060	…..	0.070
	…..	(0.012)	…..	(0.015)

*Note: Standard errors (in parentheses) are Huber/White, accounting for multiple observations per person.
All regressions also include a quartic time trend, dummy variables for urban residence, as a quartic polynomial in the experience measure and interactions of grade with the time trend. OLS regressions use years of potential experience and IV estimates use years of actual experience with potential experience as an instrument.

Table 3. Asymmetric Learning.
OLS Wage Regressions Using Potential Experience*.

	I	II	III
Grade	1.0064	0.8341	0.9783
	(0.1363)	(0.1337)	(0.1345)
Grade x Experience	-0.0283	-0.0361	-0.0441
	(0.0186)	(0.0181)	(0.0183)
Grade x Spell Length	0.0621	0.0625	0.0497
	(0.0138)	(0.0135)	(0.0146)
White	0.3408	-0.5131	-0.0030
	(0.1717)	(0.1891)	(0.1945)
White x Experience	0.1190	0.1262	0.1004
	(0.0207)	(0.0204)	(0.0231)
White x Spell Length	-0.0201	-0.0339	-0.0838
	(0.0364)	(0.0354)	(0.0390)
AFQT	…..	0.9878	0.3967
	…..	(0.0901)	(0.1201)
AFQT x Experience	…..	…..	0.0307
	…..	…..	(0.0150)
AFQT x Spell Length	…..	…..	0.0540
	…..	…..	(0.0219)

*Note: Standard errors (in parentheses) are Huber/White, accounting for multiple observations per person.
All regressions also include a quartic time trend, a dummy variable for urban residence, quartic polynomials in the experience measure and spell duration and interactions of grade with the time trend.

Table 4. Asymmetric Learning.
IV Wage Regressions Using Actual Experience*.

	I	II	III
Grade	0.9510	0.7640	0.9249
	(0.1636)	(0.1602)	(0.1587)
Grade x Experience	-0.0004	-0.0104	-0.0181
	(0.0274)	(0.0271)	(0.0274)
Grade x Spell Length	0.0236	0.0277	0.0084
	(0.0252)	(0.0250)	(0.0265)
White	0.0240	-0.8447	-0.2710
	(0.2700)	(0.2735)	(0.2643)
White x Experience	0.1647	0.1666	0.1456
	(0.0365)	(0.0361)	(0.0387)
White x Spell Length	-0.0094	-0.0087	-0.0905
	(0.0575)	(0.0571)	(0.0639)
AFQT	…..	0.9336	0.2758
	…..	(0.0939)	(0.1327)
AFQT x Experience	…..	…..	0.0274
	…..	…..	(0.0246)
AFQT x Spell Length	…..	…..	0.0837
	…..	…..	(0.0337)

*Note: Standard errors (in parentheses) are Huber/White, accounting for multiple observations per person.
All regressions also include a quartic time trend, a dummy variable for urban residence, quartic polynomials in the experience measure and spell duration and interactions of grade with the time trend. 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, actual experience, duration of the current job and a dummy variable for missing values of duration is used as an instrument for spell duration; and the residual's interactions are used as instruments for spell duration's interactions.

Table 5. Separating Tenure from Earlier Periods in the Spell.
OLS Estimates of Wage Regressions*.

	I	II	III
Grade	1.023	0.852	0.990
	(0.137)	(0.135)	(0.136)
Grade x Experience	-0.027	-0.035	-0.042
	(0.019)	(0.018)	(0.019)
Grade x (Spell-Tenure)	0.061	0.062	0.045
	(0.018)	(0.018)	(0.019)
Grade x Tenure	0.062	0.061	0.050
	(0.017)	(0.016)	(0.018)
White	0.454	-0.382	0.102
	(0.175)	(0.191)	(0.198)
White x Experience	0.112	0.118	0.094
	(0.021)	(0.021)	(0.023)
White x (Spell-Tenure)	-0.052	-0.070	-0.140
	(0.059)	(0.057)	(0.064)
White x Tenure	-0.015	-0.026	-0.068
	(0.040)	(0.039)	(0.043)
AFQT	….	0.982	0.423
	….	(0.091)	(0.123)
AFQT x Experience	…..	….	0.028
	…..	….	(0.015)
AFQT x (Spell-Tenure)	…..	….	0.071
	…..	….	(0.032)
AFQT x Tenure	…..	….	0.047
	…..	….	(0.026)

*Note: Standard errors (in parentheses) are Huber/White, accounting for multiple observations per person.
All regressions include a quartic time trend; a dummy variable for urban residence; quartic polynomials in the experience measure, (spell length - tenure) and tenure; interactions of grade with the time trend; and a missing value dummy variable for tenure along with the appropriate interactions.