A Model of Asymmetric Employer Learning with Testable Implications*

Joshua C. Pinkston
Bureau of Labor Statistics

March 2008

Abstract: This paper helps close the gap between theory and empirical evidence in the literature on asymmetric employer learning. If an employer's private learning is reflected in a worker's wage and one employer's private information is transmitted to the next when the worker makes a job-to-job transition, then asymmetric employer learning will appear in wage regressions as learning over an employment spell. Extending previous work that assumes all learning takes place publicly, this paper develops wage regressions that test for both asymmetric employer learning and public learning. The empirical results, including tests of alternative explanations, are consistent with asymmetric employer learning's having at least as much of an effect on wages during an employment spell as does public learning. The model developed in this paper illustrates how the story suggested by the empirical work might unfold. It shows that outside firms can profitably compete with a better-informed employer through bidding wars, even when the worker is equally productive in all firms. Furthermore, this competition results in different wages for workers with the same publicly observable characteristics, a result that previous models of asymmetric learning have not produced.

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, Andrew Cohen, Todd Elder, Harley Frazis, Mark Loewenstein, Dale Mortensen, Uta Schoenberg, 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.

Despite a relatively large number of theoretical papers that involve an employer accumulating more information about a worker than other firms do, there is very little empirical evidence of such asymmetric employer learning. This paper helps close the gap between theory and evidence in this literature by extending previous empirical work that assumes all learning about a worker’s productivity takes place publicly. If an employer’s private learning is reflected in a worker’s wage and one employer’s private information is transmitted to the next when the worker makes a job-to-job transition, then wage regressions will reflect evidence of employer learning over spells of continuous employment. At the same time, any learning that does take place publicly will appear as learning over experience in the labor market, as in Altonji and Pierret (2000) and Farber and Gibbons (1996).

The empirical results suggest that employer learning over a spell of continuous employment has at least as much of an effect on a worker’s wage during that spell as learning over labor market experience does. For the sake of comparison, I present results from regressions estimated under the assumptions of a pure public-learning model, as well as specifications that nest public and asymmetric employer learning. The results suggest that much of what is seen as evidence of learning over labor market experience in previous work on public learning may actually be due to employer learning within spells of continuous employment.

This paper argues that evidence of employer learning over spells of continuous employment is consistent with an asymmetric-learning model in which an employer’s private learning is reflected in a worker’s wage and one employer’s private information is transmitted to the next when the worker makes a job-to-job transition. Of course, other stories might also explain the basic empirical results of this paper. For example, information might not be transmitted from one employer to the next, and what appears in wage regressions to be learning over spells of
uninterrupted employment might actually be due to learning over tenure on the current job. In addition, the same results could arise if occupation changes were associated with breaks in employment and followed by the accumulation of occupation-specific human capital. I carefully consider these and other alternative explanations and find no evidence that is inconsistent with the story of asymmetric employer learning this paper tells.

The theoretical model presented in this paper illustrates both how the private learning of employers might influence wages and how that private information might be transmitted from one employer to the next. The model shows that outside firms can profitably compete with a better-informed employer through bidding wars even when a worker is equally productive in all firms. Furthermore, bidding wars allow a firm that outbids the worker’s previous employer to observe that employer’s private information by observing its final bid. I show that this competition and the transmission of information between employers causes wages to converge to employers’ private expectations as employment spells increase in length, for reasons similar to those behind the well-known result that an expected winning bid in an auction increases in the number of bidders. This convergence, along with the employers’ private learning, then implies that wages become more correlated with ability as the employment spell increases in length, which is precisely what the empirical analysis suggests.

Although the main contribution of this paper is empirical, the model also makes a noteworthy contribution. One of the main criticisms of previous models of asymmetric employer learning is that they do not allow outside firms to compete for workers in a way that produces wage growth without promotions or other changes in publicly observed information. In addition to illustrating the story told by the empirical results, the model applies widely known results from

---

1 See Gibbons and Waldman (1999) for a discussion of this criticism.
the auction literature to overcome this criticism. It is the first model of asymmetric employer learning in which competition from outside firms results in different wages for workers who have the same publicly observable characteristics.  

Most previous models of asymmetric employer learning focus on the relationship of wages to task assignment and make assumptions that severely restrict the ability of outside firms to compete for workers. Waldman (1984) develops a basic model of task assignment under asymmetric information that several later papers expand on. Outside firms in this basic framework possess no information that the current employer does not also possess. In later work, Waldman (1990) relaxes this assumption, but assumes that the current employer can make a final counteroffer after observing the outside firms' information. The result in either case is that the outside firms cannot compete for workers in a way that produces wage growth without promotions or changes in other observable characteristics. The model used in the current paper avoids this problem by assuming that outside firms possess some private information and that firms compete for workers through bidding wars.

The current paper is one of a few recent papers that attempt to bridge the gap between theory and evidence in this literature. Although this literature was originally motivated by studies of personnel records, Gibbons and Katz (1991) is the only previous paper to develop and test a model of asymmetric employer learning. In their model, layoffs signal that the worker

---

2 See Golan (2005) and Schoenberg (2007) for two recent models that also overcome this criticism.

3 Exceptions that do not focus on task assignment include the layoff models of Gibbons and Katz (1991) and Laing (1994) and the adverse-selection model of Greenwald (1986).


5 The others are DeVaro and Waldman (2007), who empirically investigate the value of promotions as signals; Schoenberg (2007), who performs tests that also extend earlier work on public learning; and Kahn (2007), who considers the variance of wage changes for movers and stayers. Schoenberg’s work will be discussed later in the paper as the empirical results are presented.

is of lower ability. Because displacement by a plant closing should not contain the same signal of low ability, workers who are laid off are compared with those who are displaced by a plant closing in order to control for the effects of displacement. Their estimation using CPS data supports their predictions.

The first section of this paper presents the model of asymmetric employer learning described previously. Section 2 describes how this model relates to the empirical test of asymmetric employer learning. Section 3 describes the data I use from the National Longitudinal Survey of Youth 1979. Section 4 presents estimation results and considers alternative explanations for the main results. Section 5 concludes the paper and discusses avenues for future research.

1. The Asymmetric-Learning Model

The model described in this section demonstrates how the private information of employers might become reflected in workers’ wages and how information might be passed from one employer to the next when the worker makes a job-to-job transition. Deviations from the assumptions made in the model are used later in the paper as a starting point for examining the robustness of the empirical results.

Although the model does make noteworthy theoretical contributions, it also makes assumptions that may not be realistic for all workers in the labor market. For example, recent work suggests that the assumption that firms compete for workers through bidding wars may be realistic in some settings, but not in others. Postel-Vinay and Robin (2004) develop a theoretical model in which the ability of employers to commit not to match outside offers can lead to a dual labor market in which low-productivity firms don’t match outside offers, but high-productivity...
firms do. Barron, Berger and Black (2006) present a model in which firms adopt policies of countering some outside offers, but not others. More importantly, they present evidence from a survey of employers that suggests employers are willing to match offers for roughly 41% of workers.\footnote{As far as I know, Barron, Berger and Black. (2006) is the only paper to present empirical evidence on employers’ willingness to match the offers of outside firms.} Of course, employers’ refusal to counter offers implies not only that they won’t engage in bidding wars, but also that they won’t engage in the type of bidding more commonly assumed in asymmetric-learning models in which the employer makes one final counteroffer.\footnote{Models of asymmetric employer learning based on a mechanism with no counteroffers may be able to produce the same implications as the model presented here; however, that issue must be left to future research.}

This section proceeds as follows: In Section 1.1, I present the basic assumptions of the model. I then describe the equilibrium bidding strategies of firms that encounter the worker in Section 1.2. In Section 1.3, I describe a current employer’s private information and the information contained in a worker’s wage. The section then concludes in Section 1.4 with a discussion of the sequence of wages produced by the model and shows that wages converge to the current employer’s expectation as the length of the employment spell increases. It is this convergence result, combined with the employer’s learning, that allows the model to fit the story described by the paper’s empirical results.

1.1. Basic Assumptions

I assume workers are known to be heterogeneous with worker-specific productivity $\mu_i$, which has a normal distribution\footnote{The analysis in this paper relies on the normal distribution for the simplicity and regularity it provides.}:

$$\mu \sim N\left(m, \sigma^2_\mu\right).$$
In what follows, \( x = 1, 2, \ldots \) indexes periods of a worker’s labor market experience. Periods of time during which the worker is continuously employed, regardless of whether he changes employers, are indexed by \( t \).

Employed workers are assumed to become unemployed at an exogenous rate of job destruction. This assumption ensures that the length of the current employment spell and experience are not generally the same, while avoiding the complication of modeling transitions into unemployment.

### 1.1.1. Market Learning over Experience

I assume some learning about a worker’s productivity is public so that the model nests the models of Altonji and Pierret (2000) and Farber and Gibbons (1996) (AP and FG, respectively, in what follows), in which all learning is public. At any level of experience \( x \), let \( S_x \) summarize the market’s information, and assume \( S_x \) is an unbiased estimate of the worker’s productivity, defined as

\[
S_x = \mu + \eta_x,
\]

where \( \eta_x \sim N(0, \sigma_x^2) \) and \( \sigma_x^2 \) is decreasing in experience. In other words, the market’s information about a worker grows more precise with experience.\(^{10}\)

Intuitively, the market’s information can be thought of as the information contained in a worker’s resume. Later, it will be important to distinguish the market’s information from information that appears public from the perspective of the two firms involved in a bidding war.

\(^{10}\)\( S_x \) can be modeled using a standard Bayesian updating argument; i.e., during each period, the market observes a noisy signal of the worker’s productivity, and \( S_x \) is a variance-weighted sum of all signals the market has received up to period \( x \).
in a given period. For example, both of those firms observe the worker’s current wage in my model, but the market as a whole does not observe that wage or the worker’s wage history, and the wage does 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 the worker’s career.

1.1.2. Time During an Employment Spell

I assume any firm $f$ that encounters worker $i$ for the first time in any period receives a private signal $\nu_{fi}$ from an interview or other evaluation, given as

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

(1)

where $e_{fi} \sim N(0, \sigma^2_v)$. The distribution of $\nu_{fi}$ is common knowledge. For simplicity, the worker is assumed never to encounter a firm she has been interviewed by in the past: Firms are sampled at random without replacement. Once a firm evaluates a worker, it is assumed to be costless for that firm to bid for the worker.

At $t = 0$, the worker is unemployed and randomly encounters two firms. Each firm receives a private signal $\nu_{fi}$. The firms then engage in a bidding war for the worker’s services and the worker agrees to become employed by the winner.\footnote{I'm assuming that any reservation wages unemployed individuals might have are low enough not to matter, in order to simplify the presentation.}

In each later period $t \geq 1$, 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 \( \nu_f \). 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. The current employer’s information in \( t \) consists of a weighted average of all signals observed up to that point, \( S_t \). If \( E(\mu|S_x, S_t) \) falls enough that it would otherwise be below the current wage, I assume that the current employer lowers the wage so that it equals \( E(\mu|S_x, S_t) \).\(^{13}\)

A number of assumptions about the worker’s behavior simplify the analysis that follows. I assume all workers search at the same rate, avoiding the issues of adverse selection discussed in Greenwald (1986). I also assume that moving from one employer to another does not influence the market’s learning; i.e., I don’t model situations in which workers base mobility decisions on how those decisions affect their resume.\(^{14}\) Furthermore, since workers are assumed to be equally productive everywhere, they do not consider the option values of offers and simply accept the job that offers the highest wage in each period.\(^{15}\)

Finally, I assume that firms do not allow workers to bargain with the winning bidder after a bidding war, so that the worker accepts the wage determined by the bidding war; however,

\(^{12}\)The fact that the current employer has already evaluated the worker to obtain its own private signal does not imply that the interviews or evaluations of later firms are not private information. Even if every employer gave standardized tests, any variance in the worker’s performance from one test to the next would be sufficient to make the outcome of each test the private information of the firm that gave that test. Such signals will be "affiliated," meaning that a high value of one makes a high value of another more likely, which is one of the key assumptions that allows the results of Milgrom and Weber (1982) to apply to this setting.

\(^{13}\)Workers could give employers an incentive not to abuse this ability by committing to quit if the employer lowers wages for any other reason. Motivated purely as a punishment mechanism in a repeated game, such a commitment has the problem of not being subgame perfect; however, such a commitment is easily (and perhaps more realistically) motivated by a sense of fairness. Especially in a setting where workers observe the same signals as employers, workers might prefer searching from unemployment to tolerating treatment as "unfair" as arbitrary wage cuts.

\(^{14}\)I know of no theoretical model that allows the market’s learning to be manipulated in this way. I address this issue to some extent in the empirical section by using instruments for both actual experience and current spell length.

\(^{15}\)Adding a degree of match-specific productivity to the model would introduce cases in which workers would accept offers that include lower wages, but have higher option values, as in Jovanovic (1984).
workers are able to use this wage as a reservation price in the next period’s bidding war. This condition imposes a degree of downward wage rigidity, preventing the employer from lowering the worker’s wage every time an outside firm receives a low signal of her productivity; however, this wage rigidity is limited by employers’ ability to lower wages when a worker’s expected productivity would otherwise fall below the wage.

1.2. Bidding for Workers in Equilibrium

1.2.1. Unemployed Workers

As mentioned previously, 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 drops out. The remaining firm then hires the worker at the wage at which the losing firm dropped out.

This bidding process allows me to exploit standard results from the auction literature and has intuitively appealing implications that allow the model to fit the empirical results of this paper. For example, bidding wars result in winning firms observing the signals of losing firms. This condition implies that employers learn about a worker by observing the intensity with which other firms compete for that worker’s services, as in Lazear (1986), and new employers observe the information of the previous employer in a job-to-job transition.

\[16\text{The most realistic way for this reservation policy to be enforced would probably be for workers simply not to report low outside offers; however, that assumption would complicate (but not qualitatively affect) my description of the employers’ information.}

If one assumes that employers do not allow workers to set a reservation wage that is higher than the current wage, which is in line with the assumption that firms do not allow workers to bargain for higher wages after a bidding war is completed, it is straightforward to show that the worker would want to set the reservation wage equal to the current wage, because doing so would prevent the wage from being lowered in some cases and would cause it to be lowered less in others.

\[17\text{See McAfee and McMillan (1987) and Klemperer (1999) for excellent surveys.}\]
Bidding wars also allow firms to bid profitably for workers even when they are less well informed. The winning firm pays 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 more precise private information.

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 other firm’s signal being the same. Recalling that in \( t = 0 \), each firm \( f = 0, 1 \) collects a signal \( \nu_{fi} \), 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 \( S_{xi} \) is the market’s information about worker \( i \).\(^{18}\) 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 pays a wage equal to firm 1’s bid \( b(\nu_{1}) \). Dropping the \( i \) subscripts, the resulting wage can be written as

\[
w_{x1} = b(\nu_{1}) = E(\mu|S_{x}, \nu_{0} = \nu_{1}, \nu_{1})
\]

\[
= \sigma_{\nu}^{2} \cdot E(\mu|S_{x}) + 2 \cdot \frac{V_{x}}{\sigma_{\nu}^{2} + 2V_{x}} \cdot \nu_{1}
\]

\[
< \frac{\sigma_{\nu}^{2}}{\sigma_{\nu}^{2} + 2V_{x}} \cdot E(\mu|S_{x}) + \frac{V_{x}}{\sigma_{\nu}^{2} + 2V_{x}} \cdot \nu_{0} + \frac{V_{x}}{\sigma_{\nu}^{2} + 2V_{x}} \cdot \nu_{1}
\]

\[
= E(\mu|S_{x}, \nu_{0}, \nu_{1}),
\]

\(^{18}\)This is a simple application of Theorem 6 in MW. In this paper, \( b() \) is also a function of \( x \) and \( t \), since they affect the relative precision of the signals and thus the weight each signal is given in \( b() \). In what follows, I write \( b() \) as a function of only the bidder’s signal in order to simplify notation.
where \( V_x \) is the variance of \( \mu \) conditional on \( S_x \). Firm 0 extracts a positive expected first-period rent of \( \frac{V_x}{\sigma_x^2 + 2V_x} (\nu_0 - \nu_1) \). Furthermore, paying a wage equal to firm 1’s optimal bid results in firm 0 observing \( \nu_1 \).

This result of MW is well known in the auction literature; however, the intuition behind it 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 \( b(\nu_f) \) and drop out of the bidding above that point.

This paper follows MW in that it only considers the symmetric equilibrium (i.e., the equilibrium in which all bidders follow the same strategy) of a two-player common-value English auction. Both bidders in this equilibrium bid the expectation of the worker’s productivity, conditional on their own information and the other bidder’s information being the same. There is also a continuum of strategically asymmetric equilibria in which one bidder bids more aggressively and the other bids more timidly (Milgrom, 1981).\(^{19}\)

On a related point, I assume that firms do not bid in order to establish a reputation that might influence future auctions or to collect information.\(^{20}\) Combined with other assumptions

\(^{19}\)There are good reasons to consider only 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. 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. (Timid firms would be driven out of the market.)

\(^{20}\)Bikchandani (1988) and Klemperer (1998) show that one bidder’s having the advantage of such a reputation can severely influence second-price or English auctions even when that advantage seems very small. Although it is not clear how it would apply to a model with second-price auctions, Virag (2006) develops a model in which bidders in each of multiple periods compete through a first-price auction and shows that bidders in his model are more aggressive than they would be in a static model because they learn more from higher bids.
detailed earlier, this implies that the assumption that firms ignore the expected value of future auctions is innocuous in this model. To understand why the optimal bid in this specific setting is the same as the MW bid, consider the alternatives:

1. An ultimate bid that is below the MW equilibrium bid lowers the probability that the firm will hire the worker and earn positive profits in the current and future periods. Any protection from future losses that might be provided by stopping at a lower bid is already provided by the assumption that firms can lower the wage to prevent it from exceeding expected productivity.

2. A bid above the MW equilibrium bid introduces the possibility that the firm will hire the worker at a wage above the employer’s ex post expectation. In the next period’s auction, that wage would be above the employer’s expectation of the worker’s productivity and used by the worker as a reservation bid. In order to avoid a loss, the employer would either not counter an outside offer and let the worker go or lower the wage after the outside firm did not bid. Any loss from paying a wage above the ex post expectation of the worker’s productivity in the current period would, therefore, not be balanced by higher expected profits in future periods.21

A firm in this model is indifferent between winning and losing at the same ultimate bid described in the one-period setting of MW. If firm \( f \) wins the bidding at a wage equal to \( b(\nu_f) \), it earns zero expected profits in the current period. In the next period, any outside firm willing to bid above the worker’s wage will outbid firm \( f \), yielding zero profits. If the outside firm

---

21 If workers did not use their current wage as a reservation price in the next period, firms would benefit from bidding above the myopic wage in one period because the wage in the next period would be determined by an auction without a reservation price in which they had an information advantage. See, for example, Eeckhout (2006).
doesn’t bid, firm f’s expectation will be updated downward and the wage will be lowered to equal that new expectation, again resulting in zero profits.

Finally, relaxing the assumption of two bidders in each period would only have a qualitative effect on my results as the number of bidders approaches 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. Since my model already incorporates information that is common knowledge between bidders, allowing more than two bidders in each period would simply complicate my presentation.22

1.2.2. Employed Workers

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 \(w_{xt} \).23 Assume that firm f = t + 1 observes the length of the current employment spell \(t\), but that \(t\) contains 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.

Because the worker treats \(w_{xt}\) as a reservation price, no bidding takes place when the outside firm’s optimal bid is below the current wage; however, I assume the outside firm (costlessly) reveals its optimal bid. This assumption avoids issues of what the current employer would infer if no outside firm bid for the worker in a given period, but it does not qualitatively affect my

---

22 A more realistic setting might endogenize the number of bidders, allowing it to vary based on the worker’s search intensity. This type of setting would further complicate the bidding of firms in each period, since outside firms would not necessarily know the number of firms that previously observed the worker. As a result, they would not know the precision of the employer’s information and, if they won the bidding war, would not be able to precisely infer \(S_t\).

23 The worker always has an incentive to reveal her wage to outside firms. This condition follows from the well-known result of MW that the seller in an auction maximizes expected revenue by revealing all relevant information.
results. Similarly, if the outside firm’s optimal bid is above the worker’s current wage, but the current employer’s bid is not, the worker changes employers and the (now) previous employer reveals its signal $S_t$.

The fact that the outside firm observes the worker’s wage decreases, but does not eliminate, the current employer’s informational advantage if $t > 1$; however, when $t = 1$, $w_{xt}$ is the only bid the current employer has observed, and information is symmetric between bidders.\(^{24}\) When $t > 1$, $w_{xt}$ could equal any one of multiple bids, each from an auction involving a different amount of information, or it could equal the current employer’s expectation. This situation is discussed in greater detail in the next subsection. For now, note that bids are conditioned on $w_{xt}$, and let $\tilde{S}_t$ summarize the information in $S_t$ that is not also contained in $w_{xt}$.\(^{25}\)

As mentioned earlier, the equilibrium strategies in this bidding war are unaffected by the fact that the current employer has 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 other firm’s signal of the being the same. The private signal of the outside firm is $\nu_{t+1}$, and that of the current employer is $\tilde{S}_t$. The following proposition formalizes this result:

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

1. Firm $f = t + 1$ bids for the worker up to its optimal bid,

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

\(^{24}\)If $t = 1$, the current employer’s updated signal $S_1$ is a weighted average of $\nu_0$ and $\nu_1$; however, the wage reveals $\nu_1$ to firm 2. The employer’s optimal bid is $b(S_1) = E(\mu | S_x, \nu_1, \nu_0, \nu_0 = \nu_0)$, and firm 2’s optimal bid is $b(\nu_2) = E(\mu | S_x, \nu_1, \nu_0 = \nu_2, \nu_2)$.\(^{25}\)Like $S_t$, $\tilde{S}_t$ is a real-valued variable. E.g., at $t = 1$, $\tilde{S}_1 = \nu_0$.\n
14
if \( b(\nu_{t+1}) > w_{xt} \). Otherwise, firm \( f = t + 1 \) simply reveals \( \nu_{t+1} \).

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

\[
\begin{align*}
\quad b(S_t) &= E \left( \mu | S_x, w_{xt}, \tilde{S}_t, \nu_{t+1} = \tilde{S}_t \right) \\
&= E \left( \mu | S_x, S_t, \nu_{t+1} = \tilde{S}_t \right),
\end{align*}
\]

if \( b(S_t) > w_{xt} \). Otherwise, if \( b(\nu_{t+1}) > w_{xt} \), the employer reveals \( S_t \).

This proposition is a simple application of a well-known result from MW. Even without referring to MW, its proof is a straightforward matter of showing that each bidder’s strategy is the best response to the other bidder’s strategy.

Proposition 1 implies that in each period \( t \geq 1 \), one of the following occurs:

1. The outside firm’s optimal bid is below \( w_{xt} \), and \( \nu_{t+1} \) is revealed to the current employer.

The worker’s wage is 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’s optimal bid is above \( w_{xt} \), but the employer’s bid is not. The worker becomes employed by the outside firm, and \( S_t \) is revealed. The wage is unchanged unless \( w_{xt} > E(\mu | S_{x+1}, S_t, \nu_{t+1}) \), in which case \( w_{xt+1} = E(\mu | S_{x+1}, S_t, \nu_{t+1}) \).

3. The outside firm bids above \( w_{xt} \), but the worker is retained at a higher wage \( w_{xt+1} = b(\nu_{t+1}) \).

4. The outside firm bids above \( w_{xt} \), and the worker is bid away at a wage equal to the final bid of the now former employer, \( b(S_t) \).
The cases in which the outside firm is willing to pay more than \( w_{xt} \), but the current employer is not, result in job-to-job transitions with either no wage increase or a wage decrease.\textsuperscript{26} Although more work would be needed to demonstrate whether the model predicts this situation occurs as often as it does in the data, wage decreases with job changes are known to happen.

The equilibrium strategies described in Proposition 1 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 \( f = t + 1 \) wins the bidding war only when \( \nu_{t+1} > \tilde{S}_t \), this ex post expectation is always greater than \( b(S_t) \), the wage firm \( f = t + 1 \) pays. As \( S_t \) increases in precision, however, the weight put on \( \nu_{t+1} \) in the firms’ expectations decreases, causing the expected profit of firm \( f = t + 1 \) to decrease as well.

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.\textsuperscript{27} The first is that it highlights the strengths of the model by showing 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.

Secondly, assuming that all productivity is general allows the model to fit the data better. The empirical results are consistent with a story in which one employer’s information is transmitted to the next in a job-to-job transition. The more important match-specific productivity is, the less information is transmitted in job-to-job transitions. Section 4 discusses this issue further to help motivate testing of the robustness of the result that information is transmitted.

\textsuperscript{26}Even though \( E(\mu|S_x, S_t) > w_{xt} \), it is possible that \( \tilde{S}_t \) can be low enough relative to \( S_x \) and \( w_{xt} \) that \( b(S_t) \), which puts more weight on \( \tilde{S}_t \) than \( E(\mu|S_x, S_t) \) does, can fall below \( w_{xt} \) when \( E(\mu|S_x, S_t) \) and \( w_{xt} \) are close enough.

\textsuperscript{27}The 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.
between employers in an employment spell.

1.3. Employer Learning and the Asymmetry of Information

The bidding between firms described earlier results in a worker’s employer observing the signals received by outside firms as long as the worker is retained. In this subsection, I describe the 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 when \( t > 1 \).

1.3.1. The Current Employer’s Learning

For simplicity, I assume the employer receives no other signals of worker productivity, but the results are unaffected if the employer also observes signals based on per-period output. Since the initial signals are identically distributed, the employer’s updated signal is simply the average of all initial signals received since \( t = 0 \), or

\[
S_t = \frac{1}{t+1} \sum_{g=0}^{t} \nu_g 
\]

\[
= \mu + \eta_t, \tag{3}
\]

where \( \eta_t \sim N (0, \sigma_t^2) \) and \( \sigma_t^2 = \frac{\sigma_t^2}{t+1} \). The reliability of \( S_t \) improves (\( \sigma_t^2 \) falls) as \( t \) increases and more signals are observed. As a result, the relative weight put on \( S_t \) in \( E(\mu|S_x, S_t) \) increases in \( t \), and \( E(\mu|S_x, S_t) \) converges to \( \mu \) as \( t \to \infty \).

Equation (3) 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, specifically

\[ S_t = \frac{\sigma^2_t \nu_{t+1} + \sigma^2_{0t} S_{0t} + \sum_{g'=t'+2}^t \sigma^2_{gt} \nu_{g'}}{(t-t')\sigma^2_t + \sigma^2_{0t}}, \]

where \( \nu_{t+1} \) is the initial signal received by the new employer. Once \( S_{t'} \) is plugged in, this expression reduces to Equation (3). Because information is transmitted to a new employer by the bidding process, the precision of \( S_t \) depends on \( t \), as the empirical results suggest, not on tenure.

### 1.3.2. Information Provided by a Worker’s Wage

In \( t = 1 \), the outside firm knows that the worker’s wage was the optimal bid of the losing firm in \( t = 0 \) \((w_{x1} = b(\nu_1))\), and information is symmetric between the bidders.\(^{28}\) If an outside firm observes the worker at \( t = 2 \), it learns less from the wage. If the wage was bid up in the previous period, \( w_{x2} \) would equal \( b(\nu_2) \). If the outside firm in the previous period did not bid up the worker’s wage, \( w_{x2} \) would equal either \( w_{x1} = b(\nu_1) \) or \( E(\mu|S_x, S_2) \), depending on whether \( w_{x1} < E(\mu|S_x, S_2) \) or not. There is no way for firm \( f = t + 1 \) to tell which of these values \( w_{xt} \) takes on, because outside firms observe neither a worker’s history of wage innovations nor signals received when no bidding occurs. Firm \( f = t + 1 \), therefore, makes an inference which considers the probabilities that \( w_{xt} \) equals each possibility.\(^{29}\)

In general, a worker’s wage is the minimum of the employer’s expectation and an increasing

---

\(^{28}\)From the perspective of the outside firm, the employer’s initial signal is truncated by the fact that it was willing to pay the current wage; however, if bidding actually ensues, each bidder knows that the other’s signal is high enough to support bidding more than \( w_{xt} \).

\(^{29}\)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.
sequence resulting from the bids of losing firms, or

\[ w_{xt} = \min \left[ E(\mu | S_x, S_t), \bar{b} \right], \tag{4} \]

where \( \bar{b} \) is either the maximum of all losing bids during the employment spell if the wage was never lowered, or the maximum of the employer’s expectation in the period the wage was last lowered and all bids since that period if the wage was lowered at any point during the spell. It is difficult at best to explicitly describe the information that an outside firm could infer from \( w_{xt} \); however, this information is clearly less than that in \( S_t \).\(^{30}\) While the current employer observes all of the signals contained in \( S_t \), the outside firm cannot infer any of them if \( t > 1 \).

Fortunately, this model does not require a more precise characterization of the information contained in \( w_{xt} \). All it requires is that \( w_{xt} \) provide information about productivity such that \( E(\mu | w_{xt}, \tilde{S}_t) \) is continuously increasing in \( w_{xt} \) and \( \tilde{S}_t \), which follows from the normality of all the signals involved.

1.4. The Sequence of Wages

The worker’s wage, described by Equation (4), increases toward the employer’s expectation unless it 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 \( w_{xt} \) converges to \( E(\mu | S_x, S_t) \) as \( t \) goes to infinity:

\(^{30}\)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 they would not observe the signals that did not result in a wage innovation.
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 A 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 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 is crucial to this model’s fitting the story the empirical results tell. It shows that competition from less-well-informed firms drives the wages of workers toward their employers’ expectations of the workers’ productivity. In the next section, I use this implication to develop an extension of the test of public learning in AP. In short, the model predicts that wages will reflect evidence of employer learning as the length of the employment spell increases, which is consistent with the results in Section 4.

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).
2. The Model’s Implications for Wage Regressions

The model developed in Section 1 illustrates how asymmetric employer learning can be reflected in workers’ wages. This section uses that result to develop a test for asymmetric employer learning by extending the work of AP. In public-learning models like AP, wages equal expected productivity and become more correlated with the worker’s actual productivity as the market’s expectation becomes more accurate. In the model of asymmetric employer learning developed in Section 1, wages become more closely related to actual productivity as the length of the employment spell increases, due to both the wage converging to the employer’s expectation (Proposition 2) and that expectation becoming more accurate as the employer accumulates more private information. In either case, as 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.

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

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

where $\zeta \sim N(0, \sigma_Z^2)$ and neither $f(Z)$ nor $\sigma_Z^2$ varies by $x$ or $t$. I assume that $Z$, $\delta$, and $\sigma_Z^2$ are common knowledge in the labor market. Adding these variables to the model, therefore, does not affect the results. They were omitted only to simplify notation.

In the model described earlier, the difference between $w_{xt}$ and $E(\mu|Z, S_x, S_t)$ decreases on average with the length of the current employment spell, but this difference does not de-
crease 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, \( w_{x'} = E (\mu|Z, S_{x'}, \nu_0 = \nu_1, \nu_1) \), where \( x' = x - t \). More formally, the expected wage can be written as

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

where \( \rho (t) \) is a monotonically increasing, differentiable function such that \( \rho (1) = 0 \) and \( \rho (t) \to 1 \) as \( t \to \infty \).

By expanding expectations, the expected wage can be written as a weighted average of the population mean \( m \), the easily observed variables \( Z \), actual productivity \( \mu \), and an error term:

\[
E (w_{xt}) = B_{mx} m + B_{zxt} (Z\delta) + B_{\mu xt} \mu + \phi'.
\]

(See Appendix B for a more complete description of Equation (6) and its relevant derivatives.)

Asymmetric employer learning in my model implies that \( B_{\mu xt} \) increases with \( t \), and \( B_{mx} \) and \( B_{zxt} \) decrease with \( t \). Furthermore, \( B_{mx} \) and \( B_{zxt} \) are decreasing in experience, while \( B_{\mu xt} \) is increasing in experience, due to the public component of learning.

Since the worker’s actual productivity cannot be observed, the estimation uses a variable that is correlated with productivity, but is unlikely to be observed by employers, as do AP and FG. Let \( V \) be a variable that is correlated with productivity and observed by the econometrician, but not by the market. Assume the variance of \( V \) and the covariance of \( V \) and \( \mu \) do not vary with \( x \) or \( t \). Also assume that \( V \) is uncorrelated with the error terms in \( f (Z) \), \( S_x \), and \( S_t \). Then the bias caused by using \( V \) in place of \( \mu \) does not vary with experience or employment.
spell length and does not interfere with the model’s basic predictions.³¹

In the actual estimation, I approximate the coefficients in Equation (6) with linear interactions of $x$ and $t$, resulting in a wage equation of the form

$$w_{xt} = C + Z\gamma_0 + Z \cdot x\gamma_x + Z \cdot t\gamma_t + V B_0 + V \cdot xB_x + V \cdot tB_t + \phi'_xt,$$

(7)

where $C$ is a vector containing a constant and the experience and spell length terms.³²,³³ If there is asymmetric learning, $\gamma_t$ will be negative and $B_t$ will be positive. Any learning that is public results in $\gamma_x$ being negative and $B_x$ being positive. Regressions based on Equation (7), therefore, simultaneously test for both public and asymmetric learning.

Finally, the predictions for coefficients on easily observed variables hold only when a hard-to-observe variable is included in the regression.³⁴ Without the interactions of the hard-to-observe variable, employer learning predicts no change over time in the coefficients on easily observed variables; however, other factors might cause the effects of race or education, for example, to vary over time. 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 (7). Although other factors could determine the sign of $\gamma_x$ and $\gamma_t$, these variables should nonetheless fall (become less positive or more negative) when the interactions of the

---

³¹ Let $\sigma^2_V$ denote the variance of $V$ and $\sigma^2_{\mu x}$ denote the covariance of $V$ and $\mu$. When $\mu$ is replaced by $V$ in Equation (6), the expectation of the OLS estimate of $B_{\mu xt}$ from the resulting regression is $E(\hat{B}_{\mu xt}) = B_{\mu xt} \cdot \frac{\sigma^2_{\mu x}}{\sigma^2_V}$.

³² The effects of the interactions of $m$ with $t$ and $x$ cannot be separated from any effects $t$ and $x$ have on the wage directly.

³³ The results presented later in this paper are robust to simple changes in this specification. Adding quadratic or quartic interactions of $t$ and $x$ (not shown) produces qualitatively similar results. Regressions that use log wages instead of wage levels (not shown) also produce similar results. I use wage levels because doing so is more consistent with the theory. FG do the same thing for the same reason.

³⁴ This was the main result that distinguished the work of AP from FG and allowed AP to test for “rational stereotyping” based on race and education.
hard-to-observe variable are added to the regression.

3. Data

The estimation results presented in this paper use data from the 2000 release of the National Longitudinal Survey of Youth 1979 (NLSY). The NLSY follows a nationally representative sample of 12,686 men and women who were 14–22 years old in 1979. Respondents were surveyed in each year between 1979 and 1994 and then every other year after 1994.

The NLSY data have two key advantages for the analysis in this paper. First, the data contain variables, such as Armed Forces Qualifying Test (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 information allows the measurement of both actual work experience and employment spell length.

Employment spell length is measured using data on weekly labor force status. An employment spell ends if the worker is not employed during a week and her last job ended with an involuntary termination (e.g., firing) 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. 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.
I also create a measure of tenure that is consistent with the measure of employment spell length. The tenure variable 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 counts only weeks the worker is working, I define tenure at a job as weeks worked between the start date and either the end date or the interview date. The resulting tenure variable never exceeds the spell length, unlike the standard tenure variable in the NLSY.36

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. Eligible observations are drawn from all years of the survey (1979–2000). Attention is limited to men who have left school for the final time by the beginning of the job in question, are not in the military, and have completed at least 12 years of schooling.37 These restrictions yield a sample of 40,205 observations of the CPS job (the current or most recent job at the time of interview) with nonmissing wage and experience variables for 4,209 (out of a possible 6,403) men. I exclude 481 observations with hourly wages (in 1987 dollars) below $2 or above $200; 2,847 observations with fewer than 35, or more than 100, hours worked per week; and 1,404 observations with missing

---

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

37 It is possible that my tenure variable contains weeks worked for other employers, since there are cases in the weekly work history data in which the job reported as the main job switches from one job to another and then back to the first, but most of these cases involve the respondent holding both jobs at once. In any case, the tenure variable I create is closer to the actual number of weeks worked for the employer than the standard variable is.

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 completed 8 years or more of education, and found qualitatively similar results in each case. Dropping men with years of schooling less than 12, but greater than or equal to 8, eliminates 12,204 observations for 1,115 men.

In addition, I compared results for people who entered the labor market in managerial, professional, or technical jobs with those who did not. I did not find a statistically significant difference, but the sample of people starting their careers in these occupations is relatively small.
AFQT scores. I also exclude 2,537 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 the experience measures: I drop 12 observations from men who had more than four years of potential experience in 1979, and I drop 2,825 observations in which actual experience is calculated to exceed potential experience by more than one year. 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 Caucasian, and just over 76% resides in an urban area. The average worker has completed 13.25 years of schooling; has a tenure on the CPS job of 3.6 years, but has been continuously employed for 5.1 years; and had 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 amount of potential experience is 10 years, while the average amount of actual experience is 8 years.

---

38 Although this restriction does not have a large effect on the basic results, eliminating men who are not well attached to the labor market makes all evidence of employer learning more significant.

39 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.
4. Estimation Results

The empirical results presented in this section are consistent with the asymmetric-employer-learning story told so far. They suggest that the private learning of employers is reflected in workers’ wages and that the private information of one employer is passed on to the next when the worker is bid away by an outside firm.

The first subsection presents the main empirical results of the paper. For the sake of comparison, I present results from regressions estimated under the assumptions of a pure public-learning model before presenting tests of asymmetric employer learning. The results suggest that much of the evidence of learning observed in tests of public learning may actually be due to asymmetric employer learning.

The second subsection looks at the robustness of these results and examines alternative explanations. The first such test considers the possibility that asymmetric employer learning does affect wages, but that information is not transmitted from one employer to the next, which is motivated both by the implications of relaxing assumptions in my model and by other recent work in the literature. The other tests in this subsection are concerned with various scenarios in which human-capital accumulation is confused with employer learning.

All of the results presented in this section are from regressions that include a dummy variable for urban residence and quartic time trend. Interactions of the time trend and years of schooling are also included to allow the return to education to vary by year. I use quartic polynomials in the experience measure and (in asymmetric-learning equations) spell length to control for the influence of \( x \) and \( t \) on wages.\(^{40}\) Years of schooling and a dummy variable for being caucasian

\(^{40}\)In Equation (7), the effects of \( x \) and \( t \) are incorporated in \( C \) and the error term \( \phi_{xt} \), which is a nonlinear function of both.
are the easily observed \( (Z) \) variables, and the adjusted AFQT score is the hard-to-observe \( (V) \) variable.\(^{41}\)

### 4.1. Main Results

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 captures only 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 interactions of the experience measure 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

\(^{41}\)The AFQT is typically administered by the U.S. Department of Defense only to people who intend to enter the enlisted ranks of the military and therefore is not something firms could reliably observe. Furthermore, the use of testing by firms does not suggest that AFQT is not itself difficult to observe. Any test administered by a firm will simply be another noisy signal of worker productivity that will be included in \( S_t \).

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.
the OLS regressions and 0.39 (0.12) in the IV regressions, a statistically significant decrease in both cases. The coefficient on $\text{AFQT} \times \text{experience}$ is significantly positive [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 $\text{AFQT} \times \text{experience}$ is added, but the change is never significant.

Moving to the test of asymmetric employer learning, 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 falls from 0.988 (0.090) to 0.397 (0.120). More importantly, the coefficient on $\text{AFQT} \times \text{employment spell length}$ in column III is 0.054 (0.022). The coefficient on $\text{AFQT} \times \text{potential experience}$ is slightly smaller, but also statistically significant, at 0.031 (0.015).

To put these coefficients 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 five years of potential experience raise the effect of the same change in AFQT by $0.15. This outcome is consistent with employers’ private learning having at least as much of an effect on wages as the market’s 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 its interactions. The
length of an employment spell could contain information about worker productivity, just as actual experience could, and the coefficients on spell length and its interactions could be biased for the same reason as those on actual experience and its 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.

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 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, 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. The interactions of this residual can also be used as instruments for the interactions of spell duration.

The results from the IV estimation are again consistent with asymmetric employer learning. The coefficient on AFQT × employment spell length in column III of Table 4 is 0.084

---

42 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. Furthermore, if a desire to affect what the market learns influences a worker’s mobility decisions, actual experience or length of an employment spell might be endogenous.

43 A high match value would make the job more valuable, resulting in both higher wages and a higher expected job duration. Since longer tenure is often associated with longer employment spells, employment spell length could also reflect match-specific components. See, for example, Altonji and Shakotko (1987) or Abraham and Farber (1987) for more detailed discussions of endogenous tenure.

44 These variables should control for anything that the market learns about the worker’s 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.

45 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).
(0.034), while that on AFQT × 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 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 not significant.

4.2. Robustness and Alternative Explanations

The model presented in this 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 follows from the assumptions that firms engage in bidding wars for workers’ services and that worker productivity is not match specific. Other mechanisms through which firms might compete for workers would transmit little or no information between employers. The recent paper by Schoenberg (2007) develops such a model.46 Furthermore, 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.

46Schoenberg (2007) presents a two-period model in which outside firms bid against the perfectly informed current employer, who makes a single counteroffer 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.
To investigate the possibility that information is not transmitted between employers, I consider evidence of employer learning over tenure as well as the current employment spell. If only part of the previous employer’s information is transmitted to the new employer, our intuition suggests there will be evidence of employer learning with tenure on the current job on top of any learning over the employment spell. Of course, if no information were transmitted from one employer to the next, the model would predict only evidence of learning over experience and tenure, as in Schoenberg (2007). In either case, the estimates I presented earlier might provide evidence of employer learning with the length of the employment spell in part 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 the idea that private information accumulated by one employer is transmitted to the next in a job-to-job transition. Table 5 repeats the analysis of Tables 3 and 4, with the addition of tenure and its interactions. Both the OLS and IV specifications confirm the previous evidence of learning over time in the employment spell. The coefficient on AFQT × spell length is 0.071 (0.032) in the OLS specification and 0.115 (0.054) in the IV specification. On the other hand, neither the OLS nor the IV results provide evidence of learning over tenure [the OLS and IV coefficients on AFQT × tenure are −0.024 (0.035) and 0.046 (0.136), respectively], giving us no reason to believe that the transmission of information between employers is limited by match-specific productivity or different bidding mechanisms.

---

47 I also estimated an alternative specification in which I divided spell length into time before the current job (spell length minus tenure) and tenure on the current job. The results were in line with those in Table 5; however, IV estimation in this case produced very large standard errors for coefficients on tenure and its interactions.

48 I use the residual of tenure regressed on current job duration as an instrument for tenure. Observations used in the regressions presented in Table 5 are limited to those with nonmissing values of tenure and job duration.

49 I also estimated specifications that include only interactions with experience and tenure, but found no statistically significant evidence of learning over tenure, although the coefficients had the predicted signs. Schoenberg (2007) reports evidence of learning over tenure from wage regressions for college graduates; however, the only
Another cause for concern is that my estimation, like the estimations of AP and FG, could be affected by human-capital accumulation. If on-the-job training has a larger impact on the productivity of more able workers, an increasing effect of AFQT on wages over time could be due to increasing productivity. Following AP, I repeat my OLS and IV estimations 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, one would at least expect controlling for training to reduce evidence of learning if that evidence is biased upward by a correlation between training and ability. The results of these regressions (not shown) indicate that controlling for training reduces evidence of public learning, but has no effect on evidence of asymmetric learning.

The possibility of confusing human-capital accumulation and employer learning could be more of a problem if that human capital is occupation or industry specific and changes in occupation or industry tend to occur between employment spells. In that case, what now appears to be evidence of employer learning over the length of an employment spell could actually be due to the accumulation of occupation-specific human capital over time in an occupation (or industry).

A simple comparison of the percentage of employment spells in my data that are preceded by an occupation or industry change with the percentage in which the same type of change occurs easily observed variable she allows to vary over time is a dummy for the individual’s being a college graduate (compared with high school graduates). It is unclear, therefore, whether her result is due to asymmetric learning being more important for the more educated workers in her sample or if AFQT \times tenure is picking up part of the effects of grade or race \times tenure in her analysis.

The NLSY did 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.

I owe special thanks to an anonymous referee for suggesting this alternative.
within the spell does not suggest that occupation or industry changes are any more common
between spells of employment than within them. More employment spells (37%) contain an
industry change than are proceeded by an industry change (35%), and the difference is larger
when considering occupation changes (56% versus 23%, respectively).\textsuperscript{52,53} Unfortunately, it is
well known [see Neal (1999), for example] that occupation changes in the NLSY are frequently
due to errors in reporting instead of actual changes in occupation, making it difficult to conclude
much from the frequency of occupation changes.

As a more direct test of the possibility that my results are driven by occupation or industry
changes, I compare my basic results with results from regressions that restrict observations to
cases in which industry and occupation changes are less likely.\textsuperscript{54} First, I restrict observations to
employment spells that are not preceded by an occupation change. I then restrict observations
to spells that are not preceded by an industry change. If my results are driven by observations
that involve occupation or industry changes between spells, then excluding those observations
should change the results. Finally, because the occupation data are not necessarily reliable, I
estimate regressions that restrict attention to older workers, who should be more stable in their
careers and less likely to make industry or occupation changes.

The results of this analysis, presented in Table 6, suggest that the main results of the paper
should not be attributed to any correlation between nonemployment spells and occupation or

\textsuperscript{52}This analysis uses changes in two-digit SOC or SIC codes. Using one- or three-digit changes produced
similar results.

\textsuperscript{53}These differences are larger when comparing the percentage of observations affected by an occupation or
industry change between spells with those affected by a change within spells, because spells that are preceded
by industry or occupation changes tend to be shorter. For example, 28% of observations are from spells that
are preceded by an industry change, while 39% are from spells that contain an industry change.

\textsuperscript{54}Perhaps a better way to do this would be to develop implications for a regression that included both endoge-
nous occupation (or industry) changes and employer learning; however, such a task is very complicated. See
industry changes. The first two columns of Table 6 repeat the OLS results presented in Table 3, for the sake of comparison. The next two columns show that excluding observations from spells that were preceded by occupation changes produces qualitatively similar results, with the coefficient on AFQT × potential experience being 0.038 (0.018) and that on AFQT × spell length being 0.053 (0.025). The same coefficients are slightly smaller [0.033 (0.019) and 0.049 (0.026), respectively] when observations from spells that are preceded by industry changes are excluded, but the coefficients are still significantly greater than zero (at least at a 10% level) and are not significantly different from the main results. Finally, columns G and H show that limiting attention to workers over 30, which leaves just under half of the sample, produces at least as much evidence of employer learning during an employment spell as the results from the full sample, but no evidence of learning over potential experience, which is consistent with the conclusion that the private learning of employers is still important for this subsample even though the importance of market learning has declined. All of these estimates show that the main results of this paper are robust to limiting the sample to cases in which industry and occupation changes are less likely than they are in the full sample.

Another alternative explanation is based on the possibility that human capital deteriorates during nonemployment spells. If this is the case, then what appears to be employer learning during an employment spell might be caused by workers rebuilding capital lost during the preceding nonemployment spell. I consider this possibility by reestimating my results with the sample limited to spells that were preceded by 16 or fewer weeks of nonemployment, which leaves

---

55 These results should not be taken as evidence that occupation- or industry-specific human capital is not important. They merely suggest that it does not explain the results I attribute to asymmetric employer learning. 56 IV results (not shown) were qualitatively similar, although standard errors for them rose more than they did for the OLS estimates.
just over half of the sample. Table 7 shows the results for both the OLS and IV regressions on this restricted sample. The evidence of learning over spell length is no longer statistically significant in the OLS regressions estimated under this restriction; however, the IV estimates are quite similar to those in Table 4, with the coefficient on AFQT × spell length being 0.085 (0.042) and that on AFQT × experience being 0.024 (0.032). Given that this restriction on the sample is based on part of the worker’s actual labor market history, estimates that treat both actual experience and the length of the employment spell as endogenous are probably more appropriate than the OLS estimates.

5. Conclusions and Directions for Future Research

This paper has presented evidence of asymmetric employer learning by using an extension of previous work in which all learning was assumed to be public. If an employer’s private learning is reflected in a worker’s wage and one employer’s private information is transmitted to the next when the worker makes a job-to-job transition, then wage regressions will reflect evidence of employer learning over spells of continuous employment. At the same time, any learning that does take place publicly will appear as learning over experience in the labor market, as in AP and FG. The wage regressions I have estimated show that evidence of employer learning over a spell of continuous employment is at least as strong as evidence of learning over labor market experience. In fact, I have found that much of what is seen as evidence of learning over experience in regressions that assume all learning is public may actually be due to asymmetric employer learning.

I tested alternative explanations for this result and found no evidence that is inconsistent with
the basic story of private learning affecting wages and information being transmitted between employers. I first considered related stories of asymmetric employer learning in which little or no information is transmitted between employers, which imply wage regressions should reveal evidence of learning over tenure with the current employer as well as over time in the current spell of uninterrupted employment; however, these alternative explanations were not supported by the data. Furthermore, I found that adding training variables to the wage regressions does not reduce evidence of asymmetric employer learning. I also found no evidence supporting the possibility that what appears to be evidence of learning during the employment spell is actually related to occupation- or industry-specific human-capital accumulation. Finally, I considered the possibility that the results are driven by human capital deteriorating during long spells of nonemployment and then being rebuilt during employment spells, but found that the results still held when I limited attention to employment spells that were preceded by relatively short nonemployment spells.

The model presented in this paper illustrates how the private learning of employers might influence wages and how that private information might be transmitted from one employer to the next. The model shows that, with a relatively small amount of private information, outside firms can profitably compete with better-informed employers through bidding wars, even when the worker’s productivity is completely general. Furthermore, bidding wars allow a firm that outbids the worker’s employer to observe the employer’s private information by observing the employer’s ultimate bid. I have shown that this kind of competition, combined with the transmission of information between employers, causes workers’ wages to converge to their employers’ private expectations as employment spells increase in length. This convergence, along with the employers’ private learning, then implies that wages become more correlated with
ability as the current employment spell increases in length, which is what the empirical results suggest.

Although this model is not intended to be a general description of the labor markets all workers face, its description of a particular labor market makes at least one important theoretical contribution: Competition between firms in the model results in different wages for workers who have the same publicly observable characteristics. In contrast, the previous literature on asymmetric employer learning could not explain differences in wages that were not due to differences in job level or other observable characteristics, which was one of the main criticisms Gibbons and Waldman (1999) made of this literature.57

This model has other implications for the wages of workers that are consistent with well-known empirical results. Because wages start out below employers’ expectations of workers’ productivity, but are bid up toward those expectations, wages 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) (BGH). Furthermore, the model incorporates real wage cuts, but predicts that they are more rare than wage increases, as documented by BGH. This outcomes follows from wages being lowered only when they would otherwise fall below employers’ expectations of workers’ productivity.

Nonetheless, this model would benefit from further development in future research. While I have shown that the wages of workers who otherwise appear the same differ in my model, I have not characterized the variance of wages between these workers or the volatility of a worker’s wage between periods. At the very least, characterizing these variances would provide another

\[57\text{ A recent paper by Golan (2005) also provides a noteworthy way around this criticism. In her model, more capable workers use bargaining power to earn higher wages than otherwise similar workers of lower ability. It would be worthwhile for future research to explore the empirical implications of her model.} \]
comparison between the implications of this model and the work of FG on wage dynamics under public learning.

More generally, exploring the wage and employment dynamics implied by my model could be interesting. The current paper says nothing about the model’s implications for the wage distribution and how it evolves over time or for the probability of turnover, how it evolves, and what it implies for wage changes. All of these are important issues, and developing them would allow for further comparisons of the model in this paper with other models and possibly with other empirical evidence.\textsuperscript{58}

Finally, the basic framework of this model could be used to revisit issues discussed in the literature on task assignment. The model illustrates 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 signaling 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.

\textsuperscript{58}Eeckhout (2006) develops a model that is similar to mine in which the wages of workers with entirely general human capital are determined through bidding wars under asymmetric information. He then compares the implications of his model with those obtained by models with match-specific productivity.

Two important differences between the model in Eeckhout (2006) and mine are that employed workers in his model do not use their current wage as a reservation price when they encounter a new firm, and he does not use an equilibrium that is analogous to the symmetric equilibrium of Milgrom and Weber (1982). By comparison, he uses an equilibrium in which the current employer bids more aggressively and the outside firm more timidly. (See Section 1.2 for a discussion of such equilibria.)
References


A. Proof of Proposition 2

First note that the wage $w_{xt}$, as described in Equation (4), 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, then

$$\tilde{b} = \max \{ b(\nu_1), \ldots, b(\nu_t) \}.$$  

The wage is then 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, then

$$\tilde{b} = \max \{ b(S_t-\tau+1), \ldots, b(\nu_t) \}.$$  

The same monotonic-convergence argument used earlier applies here. Furthermore, for any $\tau$,

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

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 $\mu$; thus, as job changes occur later in the employment spell, the starting wages for 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 cuts took place in a fixed period in the past and the wage is not lowered again, then the monotonic-convergence argument used earlier again applies to the sequence of wages which 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 $\psi$. (This result follows from the fact that $E(\mu|S_x, S_t)$ is a convergent sequence.)

B. Equation 6 and its Coefficients.

With $Z$ added to the model, the current employer’s expectation of the worker’s productivity is

$$E(\mu|Z, S_x, S_t) = \beta_m m + \beta_Z (Z\delta) + \beta_x S_x + \beta_t S_t,$$
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 because of public learning:

$$\frac{\partial B_{xt}}{\partial x} = (1 - \rho(t)) \left( \frac{\partial \sigma^2_x}{\partial x} (\sigma^2_x \sigma^2_Z)^2 \sigma^2_\mu}{(D')^2} \right) + \rho(t) \frac{\partial \beta_m}{\partial x} < 0.$$ 

The rest of this section will show the derivatives of $B_{mxt}$, $B_{Zxt}$, and $B_{\mu xt}$ in Equation (6) in order to support the conclusions of Section 2.

**B.1. Derivatives with Respect to Experience**

Recalling that $w_{xt} = E(\mu|Z,S_t',\nu_0 = \nu_1,\nu_1)$, we easily simplify this equation to

$$E(w_{xt}) = B_{mxt}m + B_{Zxt}(Z\delta) + B_{\mu xt}\mu + \phi',$$

where

$$\begin{align*}
B_{mxt} &= (1 - \rho(t)) \cdot \frac{\sigma^2_x \sigma^2_z \sigma^2_\mu}{D'} + \rho(t) \cdot \beta_m, \\
B_{Zxt} &= (1 - \rho(t)) \cdot \frac{\sigma^2_x \sigma^2_z \sigma^2_\mu}{D'} + \rho(t) \cdot \beta_Z, \\
B_{\mu xt} &= (1 - \rho(t)) \cdot \frac{\sigma^2_x \sigma^2_z \sigma^2_\mu}{D'} + \rho(t) (\beta_x + \beta_t), \text{ and} \\
\phi' &= (1 - \rho(t)) \cdot \left( \frac{\sigma^2_x \sigma^2_z \sigma^2_\mu}{D'} \eta_x + \frac{2\sigma^2_x \sigma^2_z \sigma^2_\mu}{D'} \eta_1 \right) + \rho(t) (\beta_x \eta_x + \beta_\eta_1).
\end{align*}$$

The first wage the worker was paid in the current employment spell is

$$w_{x't} = E(\mu|Z,S_t',\nu_0 = \nu_1,\nu_1) = \frac{\sigma^2_x \sigma^2_z \sigma^2_\mu}{D'} m + \frac{\sigma^2_x \sigma^2_z \sigma^2_\mu (Z\delta)}{D'} S_t' + \frac{2\sigma^2_x \sigma^2_z \sigma^2_\mu}{D'} \nu_1,$$

where $D' = \sigma^2_x \sigma^2_z \sigma^2_\mu + \sigma^2_x \sigma^2_\mu + \sigma^2_x \sigma^2_\mu + 2\sigma^2_x \sigma^2_x \sigma^2_\mu$.

Inserting these expectations, we can rewrite Equation (5) as

$$\begin{align*}
\beta_m &= \frac{\sigma^2_x \sigma^2_z \sigma^2_\mu}{D'} m, \\
\beta_Z &= \frac{\sigma^2_x \sigma^2_z \sigma^2_\mu}{D'} (Z\delta), \\
\beta_x &= \frac{2\sigma^2_x \sigma^2_z \sigma^2_\mu}{D'} S_t, \\
\beta_t &= \frac{2\sigma^2_x \sigma^2_z \sigma^2_\mu}{D'} \nu_1.
\end{align*}$$

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 because of public learning:

$$\begin{align*}
\beta_m &= \frac{\sigma^2_x \sigma^2_z \sigma^2_\mu}{D'} m, \\
\beta_Z &= \frac{\sigma^2_x \sigma^2_z \sigma^2_\mu}{D'} (Z\delta), \\
\beta_x &= \frac{2\sigma^2_x \sigma^2_z \sigma^2_\mu}{D'} S_t, \\
\beta_t &= \frac{2\sigma^2_x \sigma^2_z \sigma^2_\mu}{D'} \nu_1.
\end{align*}$$
This follows because \( \frac{\partial \sigma^2_x}{\partial x} < 0 \) and \( \frac{\partial \beta}{\partial x} = \frac{\partial \sigma^2_x}{\partial x} (\sigma^2_x \sigma^2_x)^2 \sigma^2_x < 0 \). Also,

\[
\frac{\partial B_{xt}}{\partial x} = (1 - \rho(t)) \left( \frac{\partial \sigma^2_x}{\partial x} (\sigma^2_x \sigma^2_x)^2 \sigma^2_x \right) + \rho(t) \frac{\partial \beta}{\partial x} < 0,
\]

because \( \frac{\partial \sigma^2_x}{\partial x} < 0 \) and \( \frac{\partial \beta}{\partial x} = \frac{\partial \sigma^2_x}{\partial x} (\sigma^2_x \sigma^2_x)^2 \sigma^2_x < 0 \). Further,

\[
\frac{\partial B_{mxt}}{\partial x} = (1 - \rho(t)) \frac{\partial \sigma^2_x}{\partial x}, \left( \frac{\sigma^2_x \sigma^2_x (\sigma^2_x \sigma^2_x + \sigma^2_x \sigma^2_x)}{(D^2)^2} \right) + \rho(t) \left( \frac{\partial \beta}{\partial x} + \frac{\partial \beta}{\partial x} \right),
\]

where

\[
\frac{\partial \beta}{\partial x} = - \frac{\partial \sigma^2_x}{\partial x} (\sigma^2_x \sigma^2_x + \sigma^2_x \sigma^2_x) \sigma^2_x \sigma^2_x \sigma^2_x > 0, \quad \text{and}
\]

\[
\frac{\partial \beta}{\partial x} = \frac{\partial \sigma^2_x}{\partial x} (\sigma^2_x \sigma^2_x)^2 \sigma^2_x < 0.
\]

The first term in \( \frac{\partial B_{mxt}}{\partial x} \) is clearly positive. The second term is also positive:

\[
\frac{\partial \beta}{\partial x} + \frac{\partial \beta}{\partial x} = - \frac{\partial \sigma^2_x}{\partial x} (\sigma^2_x \sigma^2_x + \sigma^2_x \sigma^2_x) \sigma^2_x \sigma^2_x \sigma^2_x > 0.
\]

Therefore, \( \frac{\partial B_{mxt}}{\partial x} > 0 \).

**B.2. Derivatives with Respect to Employment Spell Length**

Under asymmetric learning with some level of competition from outside firms, \( B_{mxt} \) increases with spell length, while \( B_{mxt} \) and \( B_{Zxt} \) decrease:

\[
\frac{\partial B_{mxt}}{\partial t} = \frac{\partial \rho(t)}{\partial t} \left( \beta_m - \frac{\sigma^2_x \sigma^2_x \sigma^2_x}{D^2} \right) + \rho(t) \frac{\partial \beta}{\partial t} < 0,
\]

because \( \beta_m < \frac{\sigma^2_x \sigma^2_x \sigma^2_x}{D^2} \) (i.e., less weight is put on \( m \) in the employer’s expectation) and \( \frac{\partial \beta}{\partial t} < 0 \). Also,

\[
\frac{\partial B_{Zxt}}{\partial t} = \frac{\partial \rho(t)}{\partial t} \left( \beta_Z - \frac{\sigma^2_x \sigma^2_x \sigma^2_x}{D^2} \right) + \rho(t) \frac{\partial \beta}{\partial t} < 0,
\]

45
because $\beta_Z < \frac{\sigma^2_x \sigma^2_t}{\sigma^2}$ and $\frac{\partial \beta_Z}{\partial t} < 0$. Further,

$$\frac{\partial B_{\text{ext}}}{\partial t} = \frac{\partial \rho(t)}{\partial t} \left( \beta_x + \beta_t - \frac{\sigma^2_x \sigma^2_t}{D'v} - 2\frac{\sigma^2_x \sigma^2_t}{D'^2} \right) + \rho(t) \left( \frac{\partial \beta_x}{\partial t} + \frac{\partial \beta_t}{\partial t} \right)$$

$$> 0.$$ 

Intuitively, the first term is positive because the current employer’s expectation is more correlated with the worker’s productivity than was the worker’s initial wage in the current employment spell; that is,

$$(\beta_x + \beta_t) > \left( \frac{\sigma^2_x \sigma^2_t}{D'v} + 2 \cdot \frac{\sigma^2_x \sigma^2_t}{D'^2} \right).$$

The second term is positive because $|\frac{\partial \beta_x}{\partial t}| < |\frac{\partial \beta_t}{\partial t}|$. [The algebra is analogous to the case of $\left( \frac{\partial \beta_x}{\partial x} + \frac{\partial \beta_t}{\partial x} \right).$]
<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Real Hourly Wage</td>
<td>9.881</td>
<td>6.517</td>
<td>2.002</td>
<td>192.054</td>
</tr>
<tr>
<td>Highest Grade Completed</td>
<td>13.253</td>
<td>1.873</td>
<td>12</td>
<td>20</td>
</tr>
<tr>
<td>White</td>
<td>0.699</td>
<td>0.459</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Black</td>
<td>0.258</td>
<td>0.437</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Normalized AFQT</td>
<td>-0.104</td>
<td>0.982</td>
<td>-2.018</td>
<td>2.361</td>
</tr>
<tr>
<td>Employment Spell Length</td>
<td>5.134</td>
<td>4.464</td>
<td>0</td>
<td>23.019</td>
</tr>
<tr>
<td>Tenure</td>
<td>3.608</td>
<td>3.760</td>
<td>0</td>
<td>22.846</td>
</tr>
<tr>
<td>Potential Experience Experience</td>
<td>10.037</td>
<td>5.311</td>
<td>0</td>
<td>25</td>
</tr>
<tr>
<td>Potential Experience in 2000</td>
<td>8.029</td>
<td>4.985</td>
<td>0.019</td>
<td>24.615</td>
</tr>
<tr>
<td>Experience in 2000</td>
<td>18.827</td>
<td>3.167</td>
<td>9</td>
<td>25</td>
</tr>
<tr>
<td>Experience in 2000</td>
<td>15.158</td>
<td>5.055</td>
<td>0.231</td>
<td>24.615</td>
</tr>
<tr>
<td>Urban</td>
<td>0.762</td>
<td>0.426</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

*Notes: Wages are in 1987 dollars. Spell length, tenure, and the experience variables are measured in years. There are 30,374 observations, except for tenure (29,639 nonmissing) and the experience measures in 2000 (2,150).
**Table 2. Coefficient Estimates under Public Learning: OLS and IV Estimates of Wage Regressions**

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

*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, a quartic polynomial in the experience measure, and interactions of grade with the time trend. The experience measure is years of potential experience in OLS regressions, and years of actual experience with potential experience used as an instrument in IV regressions.
### Table 3. Asymmetric Learning: OLS Wage Regressions Using Potential Experience*

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

*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*

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

*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 length, and interactions of grade with the time trend. Potential experience and its interactions are used as instruments for actual experience and its interactions. The residual of spell length regressed on the worker's career-average spell length, actual experience, duration of the current job, and a dummy variable for missing values of duration is used as an instrument for spell length. The residual's interactions are used as instruments for spell length's interactions.
Table 5. Learning over Tenure or Spell Length?*

<table>
<thead>
<tr>
<th></th>
<th>OLS</th>
<th></th>
<th>I</th>
<th>IV</th>
<th></th>
<th>II</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>I</td>
<td>II</td>
<td></td>
<td>I</td>
<td>II</td>
<td></td>
</tr>
<tr>
<td>Grade</td>
<td>0.841</td>
<td>0.985</td>
<td>0.592</td>
<td>0.763</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.133)</td>
<td>(0.134)</td>
<td>(0.415)</td>
<td>(0.354)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade x Experience</td>
<td>-0.036</td>
<td>-0.044</td>
<td>0.043</td>
<td>0.043</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.018)</td>
<td>(0.019)</td>
<td>(0.064)</td>
<td>(0.064)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade x Spell Length</td>
<td>0.066</td>
<td>0.049</td>
<td>0.029</td>
<td>0.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.018)</td>
<td>(0.020)</td>
<td>(0.035)</td>
<td>(0.039)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade x Tenure</td>
<td>-0.004</td>
<td>0.002</td>
<td>-0.099</td>
<td>-0.108</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.021)</td>
<td>(0.022)</td>
<td>(0.155)</td>
<td>(0.135)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>-0.532</td>
<td>-0.028</td>
<td>-0.869</td>
<td>-0.262</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.192)</td>
<td>(0.198)</td>
<td>(1.923)</td>
<td>(1.687)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White x Experience</td>
<td>0.124</td>
<td>0.099</td>
<td>0.177</td>
<td>0.169</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.021)</td>
<td>(0.024)</td>
<td>(0.069)</td>
<td>(0.085)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White x Spell Length</td>
<td>-0.068</td>
<td>-0.136</td>
<td>0.063</td>
<td>-0.045</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.058)</td>
<td>(0.064)</td>
<td>(0.097)</td>
<td>(0.101)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White x Tenure</td>
<td>0.059</td>
<td>0.086</td>
<td>-0.121</td>
<td>-0.122</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.059)</td>
<td>(0.067)</td>
<td>(0.488)</td>
<td>(0.414)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AFQT</td>
<td>0.984</td>
<td>0.399</td>
<td>0.921</td>
<td>0.108</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.092)</td>
<td>(0.124)</td>
<td>(0.117)</td>
<td>(0.329)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AFQT x Experience</td>
<td>.....</td>
<td>0.030</td>
<td>.....</td>
<td>0.006</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.....</td>
<td>(0.015)</td>
<td>.....</td>
<td>(0.053)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AFQT x Spell Length</td>
<td>.....</td>
<td>0.071</td>
<td>.....</td>
<td>0.115</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.....</td>
<td>(0.032)</td>
<td>.....</td>
<td>(0.054)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AFQT x Tenure</td>
<td>.....</td>
<td>-0.024</td>
<td>.....</td>
<td>0.046</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.....</td>
<td>(0.035)</td>
<td>.....</td>
<td>(0.136)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note: Standard errors (in parentheses) are Huber/White, accounting for multiple observations per person. These regressions use 29,312 observations with nonmissing values of tenure and job duration. All regressions include a quartic time trend; a dummy variable for urban residence; quartic polynomials in the experience measure, spell length, and tenure; and interactions of grade with the time trend. OLS regressions use years of potential experience, and IV regressions use actual experience with potential experience as an instrument. The residual of spell duration regressed on the worker's average spell duration, actual experience, and current job duration is used as an instrument for spell duration. The residual's interactions are used as instruments for spell duration's interactions. The residual of tenure regressed on job duration is used as an instrument for tenure, and its interactions are used as instruments for tenure's interactions.
Table 6. OLS Coefficient Estimates under Asymmetric Learning:
Are Previous Results Driven by Occupation- or Industry-Specific Human Capital?*

<table>
<thead>
<tr>
<th></th>
<th>Results from Table 3</th>
<th>No Occ. Change</th>
<th>No Ind. Change</th>
<th>Age&gt;=30</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
</tr>
<tr>
<td>Grade</td>
<td>0.834</td>
<td>0.978</td>
<td>0.832</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>(0.134)</td>
<td>(0.135)</td>
<td>(0.141)</td>
<td>(0.143)</td>
</tr>
<tr>
<td>Grade x Experience</td>
<td>-0.036</td>
<td>-0.044</td>
<td>-0.028</td>
<td>-0.038</td>
</tr>
<tr>
<td></td>
<td>(0.018)</td>
<td>(0.018)</td>
<td>(0.019)</td>
<td>(0.020)</td>
</tr>
<tr>
<td>Grade x Spell Duration</td>
<td>0.062</td>
<td>0.050</td>
<td>0.055</td>
<td>0.042</td>
</tr>
<tr>
<td></td>
<td>(0.014)</td>
<td>(0.015)</td>
<td>(0.015)</td>
<td>(0.016)</td>
</tr>
<tr>
<td>White</td>
<td>-0.513</td>
<td>-0.003</td>
<td>-0.588</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>(0.189)</td>
<td>(0.194)</td>
<td>(0.222)</td>
<td>(0.226)</td>
</tr>
<tr>
<td>White x Experience</td>
<td>0.126</td>
<td>0.100</td>
<td>0.149</td>
<td>0.117</td>
</tr>
<tr>
<td></td>
<td>(0.020)</td>
<td>(0.023)</td>
<td>(0.025)</td>
<td>(0.028)</td>
</tr>
<tr>
<td>White x Spell Duration</td>
<td>-0.034</td>
<td>-0.084</td>
<td>-0.059</td>
<td>-0.109</td>
</tr>
<tr>
<td></td>
<td>(0.035)</td>
<td>(0.039)</td>
<td>(0.042)</td>
<td>(0.046)</td>
</tr>
<tr>
<td>AFQT</td>
<td>0.988</td>
<td>0.397</td>
<td>1.043</td>
<td>0.363</td>
</tr>
<tr>
<td></td>
<td>(0.090)</td>
<td>(0.120)</td>
<td>(0.103)</td>
<td>(0.131)</td>
</tr>
<tr>
<td>AFQT x Experience</td>
<td></td>
<td></td>
<td>0.031</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.015)</td>
<td></td>
</tr>
<tr>
<td>AFQT x Spell Duration</td>
<td></td>
<td></td>
<td>0.054</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.022)</td>
<td></td>
</tr>
</tbody>
</table>

Observations | 30,374  | 30,374  | 23,590  | 23,590  | 22,084  | 22,084  | 14,110  | 14,110  |

*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 and spell duration, and interactions of grade with the time trend.
Table 7. Evidence of Asymmetric Learning
Excluding Observations that Follow Long Unemployment Spells*

<table>
<thead>
<tr>
<th></th>
<th>OLS I</th>
<th>OLS II</th>
<th>IV I</th>
<th>IV II</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade</td>
<td>0.863</td>
<td>1.005</td>
<td>0.857</td>
<td>0.997</td>
</tr>
<tr>
<td></td>
<td>(0.257)</td>
<td>(0.257)</td>
<td>(0.272)</td>
<td>(0.271)</td>
</tr>
<tr>
<td>Grade x Experience</td>
<td>-0.053</td>
<td>-0.063</td>
<td>0.171</td>
<td>-0.016</td>
</tr>
<tr>
<td></td>
<td>(0.035)</td>
<td>(0.035)</td>
<td>(0.047)</td>
<td>(0.044)</td>
</tr>
<tr>
<td>Grade x Spell Length</td>
<td>0.066</td>
<td>0.057</td>
<td>(0.038)</td>
<td>(0.031)</td>
</tr>
<tr>
<td></td>
<td>(0.026)</td>
<td>(0.026)</td>
<td>(0.065)</td>
<td>(0.038)</td>
</tr>
<tr>
<td>White</td>
<td>-0.554</td>
<td>-0.004</td>
<td>-0.703</td>
<td>-0.175</td>
</tr>
<tr>
<td></td>
<td>(0.278)</td>
<td>(0.284)</td>
<td>(0.340)</td>
<td>(0.324)</td>
</tr>
<tr>
<td>White x Experience</td>
<td>0.135</td>
<td>0.099</td>
<td>0.171</td>
<td>0.159</td>
</tr>
<tr>
<td></td>
<td>(0.029)</td>
<td>(0.033)</td>
<td>(0.047)</td>
<td>(0.050)</td>
</tr>
<tr>
<td>White x Spell Length</td>
<td>(0.046)</td>
<td>(0.082)</td>
<td>(0.038)</td>
<td>(0.123)</td>
</tr>
<tr>
<td></td>
<td>(0.056)</td>
<td>(0.058)</td>
<td>(0.065)</td>
<td>(0.070)</td>
</tr>
<tr>
<td>AFQT</td>
<td>0.878</td>
<td>0.240</td>
<td>0.832</td>
<td>0.176</td>
</tr>
<tr>
<td></td>
<td>(0.122)</td>
<td>(0.159)</td>
<td>(0.128)</td>
<td>(0.174)</td>
</tr>
<tr>
<td>AFQT x Experience</td>
<td>.....</td>
<td>0.042</td>
<td>.....</td>
<td>0.024</td>
</tr>
<tr>
<td></td>
<td>.....</td>
<td>(0.021)</td>
<td>.....</td>
<td>(0.032)</td>
</tr>
<tr>
<td>AFQT x Spell Length</td>
<td>.....</td>
<td>0.039</td>
<td>.....</td>
<td>0.085</td>
</tr>
<tr>
<td></td>
<td>.....</td>
<td>(0.030)</td>
<td>.....</td>
<td>(0.042)</td>
</tr>
</tbody>
</table>

*Note: These regressions use 15,302 observations from employment spells that are preceded by nonemployment spells of less than or equal to 16 weeks. 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 and spell length, and interactions of grade with the time trend. OLS regressions use years of potential experience, and IV regressions use actual experience with potential experience as an instrument. 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.