

IZA DP No. 8618

**Does Employer Learning Vary by Schooling Attainment?  
The Answer Depends on How Career Start Dates Are  
Defined**

Audrey Light  
Andrew McGee

November 2014

# **Does Employer Learning Vary by Schooling Attainment? The Answer Depends on How Career Start Dates Are Defined**

**Audrey Light**

*Ohio State University*

**Andrew McGee**

*Simon Fraser University  
and IZA*

Discussion Paper No. 8618  
November 2014

IZA

P.O. Box 7240  
53072 Bonn  
Germany

Phone: +49-228-3894-0

Fax: +49-228-3894-180

E-mail: [iza@iza.org](mailto:iza@iza.org)

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

## **ABSTRACT**

### **Does Employer Learning Vary by Schooling Attainment? The Answer Depends on How Career Start Dates Are Defined**

We demonstrate that empirical evidence of employer learning is sensitive to how one defines the career start date and, in turn, measures cumulative work experience. Arcidiacono, Bayer, and Hizmo (2010) find evidence of employer learning for high school graduates but not for college graduates, and conclude that high levels of schooling reveal true productivity. We show that their choice of start date – based on first-observed school exit and often triggered by school vacations – systematically overstates experience and biases learning estimates towards zero for college-educated workers. Using career start dates tied to a more systematic definition of school exit, we find that employer learning is equally evident for high school and college graduates.

JEL Classification: I21, J24, J31

Keywords: employer learning, schooling, measurement

Corresponding author:

Audrey Light  
Department of Economics  
Ohio State University  
410 Arps Hall  
1945 N. High Street  
Columbus, OH 43210  
USA  
E-mail: [light.20@osu.edu](mailto:light.20@osu.edu)

## I. Introduction

The empirically testable model of public employer learning proposed by Farber and Gibbons (1996) and Altonji and Pierret (2001) relies on a simple view of labor market entry and evolution: Workers enter the labor market at time  $t=0$  with no labor market experience ( $X=0$ ), schooling attainment ( $S$ ) is observed at  $t=0$  and *unchanged* beyond that point, and employers observe a continuous stream of performance signals as labor market experience evolves. In short, the employer learning model follows orthodox human capital models (Becker 1964; Ben-Porath 1967; Mincer 1974) in assuming a well-defined, once-and-for all transition from school to employment.

Longitudinal survey data—*e.g.*, the 1979 National Longitudinal Survey of Youth (NLSY79), which is widely used for empirical implementation of the employer learning model<sup>1</sup>—reveal that school-to-work transitions are often less clear-cut than theoretical models assume. Young people interrupt their school enrollment, work while in school, and work both discontinuously and part-time while nonenrolled.<sup>2</sup> Independent of these behaviors, NLSY79 interviews often take place in the summer (especially in early survey rounds), so analysts who allow nonenrollment status at the interview date to trigger the start of the career can inadvertently “start the clock” long before the respondent leaves school or enters the labor market. Given the nature of the data, it is unclear how the career start date should be defined—yet this is potentially a critical decision insofar as it determines which wage observations are included in the analysis, how cumulative labor market experience is measured, and whether schooling attainment continues to increment once the career is deemed to be underway.

In this paper, we ask whether these measurement issues affect the seminal test for employer learning proposed by Altonji and Pierret (2001), hereafter AP, and the extension to that test used by Arcidiacono, Bayer and Hizmo (2010), hereafter ABH. The AP test calls for a log-wage model with  $S$ ,  $Z$ ,  $S \cdot X$ , and  $Z \cdot X$  among the regressors, where  $X$  is labor market experience,  $S$  (schooling) is a signal of pre-market productivity observed by employers at  $t=0$ , and  $Z$  (typically a test score) represents a component of pre-market productivity that employers cannot observe *ex ante*. A negative estimated coefficient for  $S \cdot X$  and a positive estimated coefficient for  $Z \cdot X$  are evidence that wage determination depends less on initial productivity signals and more on true productivity as employers learn. In an extension to this test, ABH ask whether employer

---

<sup>1</sup>NLSY79-based tests of the employer learning model include Farber and Gibbons 1996; Altonji and Pierret 2001; Pinkston 2006; Lange 2007; Schönberg 2007; Arcidiacono, Bayer and Hizmo 2010; Mansour 2012; and Light and McGee 2015.

<sup>2</sup>Studies documenting these phenomena include Light (1995a, 1995b) and Chuang (1997) on interrupted schooling; Ruhm (1995), Light (2001), Hotz *et al.* (2002), and Parent (2006) on in-school employment; and Keane and Wolpin (1997), Booth (2002) and Neumark (2002) on early-career employment stability. Michael and Tuma (1984) and Light (1998) focus more generally on school-to-work transitions.

learning differs by workers' schooling attainment. They estimate separate log-wage models for workers with 12 and 16 years of schooling, and find that the estimated  $Z \cdot X$  coefficient is positive for the former and zero for the latter. That is, they find evidence of employer learning for high school ( $S=12$ ) workers only, which they interpret as evidence that college-educated ( $S=16$ ) workers are able to signal their true pre-market productivity at the outset of their careers.

To learn whether evidence of employer learning is affected by measurement, we conduct both the AP and ABH tests using alternative NLSY79 datasets. First, we work with the exact data used by ABH, whose sample selection rules and definition of career start date are drawn from AP and Lange (2007). Second, we use a preferred version of the data in which the career start date and, in turn, measures of  $S$  and  $X$  conform closely to the employer learning model. Whereas ABH's career start date is the "year last enrolled" reported by respondents the first time they are not enrolled at the interview date—which may seem like an innocuous definition of  $t=0$ , but in fact corresponds to summer or winter enrollment breaks for those respondents who happen to be interviewed between school terms—we follow Farber and Gibbons (1996), Pinkston (2006) and Schönberg (2007) in pinning our career start date to a well-defined school-to-work transition that is unaffected by the NLSY79 interview schedule. Specifically, we define  $t=0$  as the start of the first nonenrollment spell lasting at least 12 months.<sup>3</sup> ABH define schooling attainment as the *time-varying* value prevailing each time a wage is reported; we define  $S$  as the *time-constant* level observed at the start of the career. Because these alternative datasets differ with respect to the measure of  $X$ , the measure of  $S$ , and the wage observations included in the sample, we also use a series of intermediate datasets that hold constant select factors; *e.g.*, we work with a dataset that uses our preferred  $X$  and  $S$  measures, but contains only those observations that appear in the ABH sample.

By systematically manipulating the data, we demonstrate the following: First, ABH's findings are sensitive to the definition of career start date. Using their data, we reproduce their evidence of employer learning for  $S=12$  workers but not for  $S=16$  workers. Using our preferred data, we find that employer learning is equally evident for both schooling groups. Second, ABH's "zero effect" for  $S=16$  workers can be attributed to their use of career start dates that often precede school exit, and the attendant overstatement of potential experience. This mismeasurement arises for  $S=16$  workers to a much greater extent than for  $S=12$  workers because the longer NLSY79 respondents stay in school, the more likely they are to be interviewed during a short enrollment break. Third, ABH's (selectively) early career start date does *not* have the advantage of capturing in-school employment experience that is missed by our later start date. Mean levels of *actual* experience are identical when based on the ABH start date

---

<sup>3</sup>Our findings are invariant to whether we define the cutoff as 6, 12, or 15 months.

and our start date, and it is therefore unsurprising that we find equal evidence of employer learning for  $S=12$  and  $S=16$  workers when we use either start date but substitute a measure of actual experience for potential experience. Fourth, ABH's findings are invariant to whether  $S$  is allowed to increment as experience accrues. Use of a time-varying schooling measure is inconsistent with the employer learning model, but the variation in  $S$  is not substantial enough to affect the estimates. Fifth, AP's findings—which are based on a pooled sample of workers with schooling levels ranging from 8 to 20—are robust to the definition of career start date. Because mismeasurement of  $X$  is concentrated among workers with high schooling levels, it does not drive the estimates when those workers contribute a minor portion of the variation used for identification.

## II. Employer Learning

We begin this section with a condensed presentation of the AP employer learning model (see also Farber and Gibbons 1996) to highlight the fact that the career start date ( $t=0$ ), schooling ( $S$ ) and cumulative labor market experience ( $X$ ) are well-defined within the context of the model. After describing in section II.B AP's empirical test for employer learning and the extension used by ABH, in II.C we consider how empirical tests of employer learning are likely to be affected by mismeasurement of the career start date,  $S$ , and  $X$ .

### A. The Altonji and Pierret Employer Learning Model

The process of public employer learning begins at  $t=0$ , when the worker enters the labor market with a given level of productivity that cannot be directly observed by employers. AP express initial log-productivity as:

$$y_{i0} = rS_i + \alpha_1 q_i + \lambda Z_i + \eta_i$$

where  $S_i$  is observed by (all) employers and by the econometrician,  $q_i$  is observed by (all) employers but not by the econometrician,  $Z_i$  is observed by the econometrician only, and  $\eta_i$  is observed by neither party. It is important to recognize that the econometrician's empirical analogs to  $S_i$  and  $Z_i$  (typically highest grade completed and a cognitive test score) must be measured at  $t=0$  and unchanged beyond that point to be consistent with this theoretical framework.

Employers form an expectation of factors they *cannot* observe ( $Z_i$  and  $\eta_i$ ) at  $t=0$  on the basis of factors they *can* observe ( $S_i$  and  $q_i$ ):

$$E(Z_i | S_i, q_i) + v_i = \gamma_1 q_i + \gamma_2 S_i + v_i$$

$$E(\eta_i | S_i, q_i) + e_i = \alpha_2 S_i + e_i.$$

Employers then use these expressions to form an expectation of initial log-productivity:

$$E(y_{i0} | S_i, q_i) = (r + \lambda\gamma_2 + \alpha_2)S_i + (\lambda\gamma_1 + \alpha_1)q_i,$$

where  $\lambda v_i + e_i$  represents the initial error in employers' assessments of initial productivity. Over

time, firms update their expectations about  $y_{i0}$  using observations of the worker's performance history ( $D_{it}$ )—information that, *by definition*, is unavailable at or before  $t=0$ . The model therefore imposes two (related) requirements on the analyst's definition of  $t=0$ : the career start date must (i) correspond to the point in the lifecycle when employer learning plausibly begins; and (ii) precede all relevant observations of the worker's performance history.

The AP model assumes that labor markets are competitive and that workers' log-wages equal their expected log-productivity. The log-wage paid by an employer at  $t$  is given by:

$$\begin{aligned} w_{it} &= E(y_{it}|S_i, q_i, D_{it}) \\ &= (r + \lambda\gamma_2 + \alpha_2)S_i + (\lambda\gamma_1 + \alpha_1)q_i + H^*(t_{it}) + E(\lambda v_i + e_i|D_{it}) + \zeta_{it} \end{aligned} \quad (1)$$

where  $H^*(t_{it})$  and  $\zeta_{it}$  represent additions to log-productivity that occur after  $t=0$  and factors outside the model, all of which are assumed to be orthogonal to  $S_i$ ,  $q_i$ ,  $Z_i$ , and  $\eta_i$ .

## B. Empirical Tests of the Employer Learning Model

AP's empirical test exploits the fact that econometricians use components of initial productivity that they observe ( $S_i$  and  $Z_i$ ) to estimate a misspecified version of equation 1:

$$w_{it} = b_{st}S_i + b_{zt}Z_i + H^*(t_{it}) + \varepsilon_{it}. \quad (2)$$

Using AP's derivation, we have

$$E(\hat{b}_{st}) = (r + \lambda\gamma_2 + \alpha_2) + \phi_{qs} + \phi_{st} = b_{s0} + \phi_{st} \quad (3a)$$

$$E(\hat{b}_{zt}) = \phi_{qz} + \phi_{zt} = b_{z0} + \phi_{zt} \quad (3b)$$

where  $\phi_{qs}$  and  $\phi_{qz}$  are coefficients from hypothetical regressions of  $(\lambda\gamma_1 + \alpha_1)q_i$  on  $S_i$  and  $Z_i$ , and  $\phi_{st}$  and  $\phi_{zt}$  are coefficients from hypothetical regressions of  $E(\lambda v_i + e_i|D_{it})$  on  $S_i$  and  $Z_i$ . These are the bias terms in the econometrician's OLS estimators for  $b_{st}$  and  $b_{zt}$  in equation 2.

AP express the time-varying component of each bias as

$$\phi_{st} = \theta_t \phi_s \quad \text{and} \quad \phi_{zt} = \theta_t \phi_z$$

where  $\phi_s$  and  $\phi_z$  are coefficients from hypothetical regressions of the initial error  $(\lambda v_i + e_i)$  on  $S_i$  and  $Z_i$ , and

$$\theta_t = \frac{\text{cov}(E(\lambda v_i + e_i|D_{it}), Z_i)}{\text{cov}(\lambda v_i + e_i, Z_i)} = \frac{\text{cov}(E(\lambda v_i + e_i|D_{it}), v_i)}{\text{cov}(\lambda v_i + e_i, v_i)}.$$

Based on the preceding derivations, AP are able to make the following arguments. First, upon making the innocuous assumptions that  $Z_i$  is a scalar,  $\text{cov}(\lambda v_i + e_i, v_i) > 0$  and  $\text{cov}(S_i, Z_i) > 0$ , we can sign the time-constant terms:  $\phi_s < 0$  and  $\phi_z > 0$ . Second, the time-varying term  $\theta_t$  increases monotonically from zero at  $t=0$ , when the performance history  $D_{it}$  has yet to be observed, to a maximum value of one at  $T$ , when the performance history has fully revealed true pre-market productivity. Together, these arguments imply that  $\hat{b}_{st}$  (the econometrician's estimated coefficient for  $S_i$ ) is expected to decrease in  $t$ , while  $\hat{b}_{zt}$  (the estimated coefficient for  $Z$ ) is expected to increase in  $t$ .

AP operationalize this test by estimating

$$w_{it} = \beta_0 + \beta_1 S_i + \beta_3 Z_i + \beta_4 S_i \cdot X_{it} + \beta_5 Z_i \cdot X_{it} + \beta_6 X_{it} + \epsilon_{it}, \quad (4)$$

where  $X_{it}$  (cumulative labor market experience) is the empirical analog to elapsed time since the beginning of the career  $t$  (subscript  $t$  in  $w_{it}$ ,  $X_{it}$ , and  $\epsilon_{it}$  refers to years in our longitudinal data),  $S_i$  is “highest grade completed” at  $t=0$ , and  $Z_i$  is a pre-market measure (typically a test score) that is correlated with productivity at  $t=0$ . (In practice, the specification used by AP and others includes higher-order experience terms and a number of additional regressors.) A finding that  $\hat{\beta}_4 < 0$  and  $\hat{\beta}_5 > 0$  is consistent with the prediction that  $\hat{b}_{st}$  decreases and  $\hat{b}_{zt}$  increases in  $t$ —*i.e.*, it is evidence that, over time, log-wages are tied less to initially-observable signals ( $S_i$ ) and more to information revealed by ongoing performance signals.<sup>4</sup>

ABH modify the AP test to determine how evidence of employer learning compares for two “types” of workers defined by their schooling levels. They estimate the log-wage model

$$w_{it} = \beta_0 + \beta_3 Z_i + \beta_5 Z_i \cdot X_{it} + \beta_6 X_{it} + \epsilon_{it}, \quad (5)$$

for separate subsamples of workers with  $S_i = 12$  and  $S_i = 16$ , and test for equality of the two estimates of  $\beta_5$ . While the magnitudes of  $\hat{\beta}_5$  are not directly comparable across samples because they depend on sample covariances between  $Z_i$  and all other regressors, ABH’s conclusions ultimately depend only on “sign” tests:  $\hat{\beta}_5 > 0$  for the  $S_i = 12$  sample and  $\hat{\beta}_5 = 0$  for the  $S_i = 16$  sample.<sup>5</sup>

### C. Measurement Issues

We begin by considering how empirical tests for employer learning are affected when cumulative labor market experience ( $X_{it}$ ) overstates the interval ( $t$ ) in which employers receive performance signals. As shown in II.B, the bias term  $\theta_t \phi_z$  underlies the empirical test for employer learning. Assuming the hypothetical relationship between employers’ initial “productivity assessment error” and the test score representing  $Z_i$  (a relationship that defines  $\phi_z$ ) is unaffected by mismeasurement of  $X_{it}$ , we need only ask how  $\theta_t$  is affected when  $X_{it}$  overstates the true interval  $t$ . Because  $\theta_t$  represents the extent to which employer learning—*viz.*, revisions to the initial “productivity assessment error” made upon observing performance history  $D_{it}$ —covaries with  $Z_i$ , the answer is obvious: if we use a proxy for  $t$  that overstates the period over which performance signals are received, the covariance represented by  $\theta_t$  is understated and  $E(\hat{b}_{zt})$  increases in  $t$  more slowly than it otherwise would (or, in the extreme, does not increase

<sup>4</sup>The model focuses on the process by which employers learn *initial* productivity ( $y_{i0}$ ), and assumes that any augmentation of initial productivity via on-the-job training is captured by  $H^*(t)$  and is *independent* of  $S_i$  and  $Z_i$ . Farber and Gibbons (1996) and AP acknowledge that complementarities between on-the-job training and  $S_i$  can lead to the finding  $\hat{\beta}_4 \geq 0$ .

<sup>5</sup>Light and McGee (2015) extend the AP test (following Farber and Gibbons 1996) by replacing  $Z_i$  with a standardized, residual value that is orthogonal to all other regressors. The use of this alternative  $Z_i$  allows the magnitude of  $\hat{\beta}_5$  to be compared across samples and specifications. For comparability with AP and ABH, we use non-residual measures of  $Z_i$  throughout the current empirical analysis.

in  $t$ ). Turning to log-wage models 4-5, the estimated coefficient for  $Z_i \cdot X_{it}$  (which is driven by the omitted variable bias  $\theta_t \phi_z$ ) will tend toward zero. Overstating  $t$  is equivalent to looking for evidence of employer learning over an interval when, in fact, employer learning does *not* take place.

In principle, analysts can avoid overstating  $t$  in two ways: by measuring  $X_{it}$  as elapsed time from a career start date that plausibly represents the start of employer learning, or by measuring  $X_{it}$  as *actual* work experience. The first approach is well-represented in the literature, and career start dates based on a carefully-chosen notion of both school exit (Pinkston 2006) and labor market entry (Farber and Gibbons 1996; Schönberg 2007) have been proposed.<sup>6</sup> Nonetheless, no analyst can guarantee that his or her preferred date corresponds precisely with the start of employer learning—and a “too early” career start date that captures time spent nonemployed or employed in uninformative jobs will necessarily cause the estimated employer learning parameter to tend towards zero.

A measure of *actual* work experience potentially eliminates the mismeasurement of  $X_{it}$  due to an injudiciously chosen career start date. In fact, actual experience can be accumulated from a “too early” career start date with impunity if the goal is to avoid counting nonemployment in the experience measure and/or to include potentially informative “early” work experience. However, the use of actual experience is problematic for three reasons. First, in the absence of a carefully-chosen career start date, actual experience might include in-school employment that is irrelevant to the employer learning process. Second, while the employer learning model assumes employers use performance signals observed from  $t=0$  onward to assess *pre-market* log-productivity ( $y_{i0}$ ), actual experience is informative about labor force attachment, job mobility, on-the-job training, and other factors that reflect *changes* to worker productivity after  $t=0$ . AP and Pinkston (2006, 2009) attempt to skirt this problem by using potential experience as an instrument for actual experience, but employers are likely to assess productivity changes by observing actual experience *conditional* on elapsed time. Third, to the extent that actual experience *is* informative about  $y_{i0}$ —*e.g.*, because workers with high pre-market productivity tend to work continuously—including a polynomial in actual experience in the log-wage model is inconsistent with the assumption that  $H^*(t_{it})$  is unrelated to employer learning.

Despite its limitations in the context of the employer learning model, an actual experience measure serves a valuable purpose for our analysis. While we argue that our career start date is sensible and consistent with the employer learning model, it is vulnerable to the criticism that it misses informative work experience acquired prior to school exit. Similarly, we argue that ABH’s start date is defined inconsistently across individuals and is often triggered by summer

---

<sup>6</sup>See table A3 (appendix A) for a summary of how the career start date,  $X$ , and  $S$  are measured in a range of employer learning studies based on the NLSY79.

vacations, yet it can be defended on grounds that it (selectively) captures in-school work experience. By constructing both potential and actual experience measures for each start date and using each measure (separately) to conduct the AP and ABH tests of employer learning, we can isolate the effect of nonwork time included in the potential experience measures, and of early work time captured uniquely by the ABH actual experience measure.

Estimates of log-wage models 4-5 can also be affected when the empirical proxy for  $S_i$  does not accurately represent the signal observed by employers at  $t=0$ . This measurement problem arises when individuals increase their schooling attainment after the career start date—a situation that arises when the start date is triggered by nonenrollment corresponding to summer breaks from school, but also when individuals return to school after lengthy nonenrollment spells. AP, Lange (2007), Schönberg (2007) and ABH use time-varying “highest grade completed,” we use time-constant  $S_i$  in the current study and in Light and McGee (2015), and other analysts are not explicit about whether their schooling measure is time-varying or time-constant (see table A3).

If analysts use the “wrong” value of  $S_i$  when estimating equation 4, estimates of  $\beta_1$  and  $\beta_4$  do not correspond to the derivation for  $E(\hat{b}_{st})$  shown in 3a; all three components of this expectation ( $r + \lambda\gamma_2 + \alpha_2$ ,  $\phi_{qs}$ , and  $\phi_s$ ) are valid only if the econometrician uses the same level of schooling that employers use at  $t=0$  to form initial productivity assessments. Nonetheless, as long as increments to  $S_i$  are seen for relatively few sample members and/or deviate relatively little from values observed at  $t=0$ , we do not expect estimates of equation 4 to be significantly affected. Estimates for equation 5 are more likely to be affected because incremental schooling will systematically place individuals into the  $S_i = 16$  sample (with an attendant overstatement of  $X_{it}$ ) on the basis of schooling attainment reported after the career start date. To avoid inconsistency with the learning model, our preferred strategy is to tie the career start date to a defensible definition of school exit, and then terminate the observation window at the onset of any reenrollment spell that causes  $S_i$  to increment.

Table A3 (Appendix A) compares measures of career start dates, schooling, and experience used for a broad sample of employer learning studies based on the NLSY79. Our goal is *not* to assess the robustness of employer learning estimates to the many career start dates and experience measures that have been used in the literature. While some measures of career start date,  $X_{it}$ , and  $S_i$  conform more closely to the employer learning model than others, there is no “right” way to identify career starting dates or labor market experience. More importantly, experimentation described in the literature (*e.g.*, Farber and Gibbons 1996; Pinkston 2006) and conducted as part of our analysis indicate that estimates of equation 4 tend to be robust to alternative measurement strategies. The reason for this is clear: each measurement issue we consider is more likely to affect highly-schooled workers than their less-schooled counterparts, and variation contributed by highly-schooled workers does not drive the identification of

equation 4 parameters. In contrast, parameters of equation 5 are identified solely for  $S_i = 12$  and  $S_i = 16$  subsamples. Our analysis centers on a comparison of ABH’s measurement strategy and our preferred approach because measurement proves to matter primarily when testing the employer learning model for a sample of college-educated workers.

### **III. Data**

Our data are from the 1979 National Longitudinal Survey of Youth (NLSY79), which is the data source used by ABH and AP. The original NLSY79 sample (interviewed in 1979) consisted of 12,686 male and female respondents born in 1957-64. Respondents were interviewed annually from 1979-94 and biennially thereafter, with data currently available through 2012. We confine our attention to data from 1979-2004 to conform to ABH; AP use 1979-92 data. The NLSY79 is the survey of choice in the empirical employer learning literature because it (i) allows log-wage models to be estimated from labor market entry onward; and (ii) provides cognitive test scores and family background variables that plausibly represent pre-market productivity factors ( $Z$ ) that employers do not observe *ex ante*.

#### **A. Construction of Alternative Samples**

Our analysis relies on comparisons among several alternative samples. We begin by reproducing the sample that ABH used to compute their baseline estimates (columns 1 and 3 of their table 2). This sample—which conforms closely to the sample selection rules and variable definitions used by AP and Lange (2007)—is restricted to workers with either 12 or 16 years of schooling, and uses potential experience (elapsed time since the start of the career) to proxy  $t$ . We compare this sample to our preferred version, in which we measure the career start date, schooling attainment, and potential experience differently than ABH but duplicate their selection rules and variable definitions in all other respects. We refer to these two samples as the ABH and LM (for Light and McGee) “ $S=12/S=16$  potential experience” samples. We construct analogous ABH and LM samples in which we replace potential experience with detailed measures of actual experience. When working with both potential and actual experience measures, we also use “intermediate” samples in which, for example, we discard observations that are included in the ABH sample but excluded from the LM version. To conduct the AP test of employer learning, we relax the selection criteria requiring workers to have either  $S=12$  or  $S=16$  to produce pooled ABH and LM “all  $S$  levels” samples (where  $S$  ranges from 8 to 20), using both potential and actual experience.

In the remainder of this subsection, we describe the selection criteria and variable definitions used to construct each sample, highlighting key differences between ABH’s approach and ours. Because the integrity of our analysis hinges on transparency about the comparison of our data and the data used by ABH, we provide a more detailed discussion of these issues in Appendix A. The appendix also contains summary lists of sample selection rules (table A1) and variable

definitions (table A2).

Starting with criteria that determine which of the 12,686 NLSY79 respondents are included in the samples (rows *a-e* in table A1), ABH confine their attention to 5,066 male, non-Hispanic respondents with valid scores for the Armed Forces Qualifications Test (AFQT). These three criteria are commonly used by analysts for empirical tests of employer learning models. We adhere to these selection criteria in constructing each alternative sample for this analysis.<sup>7</sup>

Another respondent-specific selection rule (rows *d-d'* in table A1) leads to a key difference between the ABH and LM samples: individuals must have a valid *career start date* to remain in the sample. This criterion drops respondents who attrit from the survey prior to crossing the chosen threshold. More importantly, the manner in which the start date is defined is critical to the subsequent measurement of schooling and experience. ABH identify the first interview date at which each respondent reports himself *not* enrolled in school, and use the contemporaneously-reported “year last enrolled” (ignoring the reported month) as their career start date. We refer to this “start year” variable as  $SY^{abh}$ .

In our judgment,  $SY^{abh}$  has three shortcomings. First, it starts the clock on labor market experience for respondents who are experiencing a short-term nonenrollment spell, including summer vacation or other inter-term breaks. Second, it fails to define the career start date uniformly for all respondents, given that the vagaries of NLSY79 interview schedules determine which respondents are interviewed during the summer or during other short enrollment breaks.<sup>8</sup> Third, ABH do not utilize information on the month of school exit; in combination with a similar aggregation of interview dates to calendar years, this introduces “rounding” error in potential experience.

To underscore the seriousness of the first two shortcomings noted above, consider two respondents who (*i*) are age 16 at the time of the 1979 interview; (*ii*) attend school continuously for 16 years until graduating from college in May 1985; (*iii*) have a three-month enrollment break in May, June, and July between all 16 school years; and (*iv*) are nonemployed until June 1985. For the purpose of our illustration, the only difference between these individuals is that Respondent A is interviewed by the NLSY79 in July 1979 and in October of every subsequent year, while Respondent B is always interviewed in October. ABH would define A’s career start year as 1979 despite the fact that, at age 16, he remains six years away from leaving school and entering the labor market; ABH would correctly define B’s start year as 1985. Clearly, the ABH career start date is based on each respondent’s first *observed* nonenrollment spell rather than a

---

<sup>7</sup>We retain only 5,065 non-Hispanic men with valid AFQT scores upon correcting what we judge to be a minor error in ABH’s identification of respondent sex; see the discussion of criterion *a* in Appendix A.

<sup>8</sup>Between 1979 and 1998 (the maximum value of  $SY^{abh}$ ), 41% of NLSY79 interviews conducted with non-Hispanic, male respondents took place in May through August.

definition of “true” school exit or labor market entry, and as such is subject to the arbitrary nature of the NLSY79 interview schedule.<sup>9</sup>

To remedy these shortcomings, we define the career start date ( $SD^{lm}$ ) as the month and year that starts the first nonenrollment spell lasting at least 12 months. We identify this date using two alternative information sources: “month and year last enrolled” reported by respondents who are nonenrolled at the interview date, and monthly enrollment timelines available from 1980 onward; details appear in Appendix A. While we believe  $SD^{lm}$  is sensibly-defined and measured both accurately and uniformly for all respondents, we acknowledge that our choice of a 12-month nonenrollment spell to trigger the start of the career is inherently arbitrary. Through a series of experiments, we established that our key findings are invariant to whether we shorten the required nonenrollment spell to six months or lengthen it to 15 months.<sup>10</sup>

Additional selection criteria (rows *f-o* in table A1) determine which wage observations are used for men selected into the sample. ABH confine their attention to the “current or last” wage reported in interview years 1979-2004. They require the average hourly wage to be between \$1 and \$100, the respondent’s residence to be urban or rural (*i.e.*, nonmissing), the class of worker to be private sector or government, and the respondent’s employment status at the interview date to be working or with a job. We follow these selection rules in constructing all samples.

Selection criteria dictating which portion of the 1979-2004 observation window is “in range” (rows *g-h* and *l-l'* in table A1) represent another deviation between our strategy and ABH’s. In principle, the first rule should simply be that the interview date at which a wage is reported must be past the career start date.<sup>11</sup> ABH instead require that the interview year ( $IntY$ ) exceed the career start year ( $SY^{abh}$ ) by at least  $\Delta S$ , where  $\Delta S$  is the difference between “highest grade completed” in the interview year and the minimum “highest grade completed” from  $SY^{abh}$  onward. As illustrated in Appendix A, this rule selectively deletes early-career wage observations for sample members whose schooling increments after  $SY^{abh}$ . ABH do not justify this rule which, in our judgment, is not the preferred way to account for schooling activity after the career begins.

---

<sup>9</sup>Note that the sooner respondents A and B leave school, the less severe is the mismeasurement of A’s start year. For example, if both respondents are age 18 in 1979 and leave school permanently that year with  $S=12$ , ABH correctly identify their start year as 1979.

<sup>10</sup>After smoothing over one-month enrollment gaps (which can represent reporting error) and nonenrollment spells that appear to correspond to school vacations, only 19% of sample members have a nonunique school exit date. Among these individuals, the majority interrupt their enrollment for at least 12 months; among individuals with shorter breaks, the gap between first and last school exit is usually less than two years. For these reasons, cutoffs of 6, 12 and 15 months deliver identical or similar data for the vast majority of sample members.

<sup>11</sup>The ABH and LM samples use wages for the “current or last” job and select workers who are employed at the interview date, so the interview date is effectively the date at which the wage is earned.

In contrast to ABH’s rule, we simply require that the interview date ( $IntD$ ), defined as the interview month and year, exceed the career start date ( $SD^{lm}$ )—*i.e.*, the wage must be earned after the career begins. To contend with the fact that reenrollment and increments to “highest grade completed” *are* observed after the career has begun (although far less often when  $SD^{lm}$  is used as the start date than when  $SY^{abh}$  is used), we require the interview date to precede any reenrollment spell that triggers a subsequent increment to “highest grade completed.” In other words, we terminate the observation window when a respondent returns to school, given that any change in the pre-market schooling signal is inconsistent with the employer learning model.

To complete the selection of “in range” observations, ABH require potential experience (elapsed time since the start of the career) to be less than 13 years in light of evidence that employer learning is confined to the early part of the career (Lange 2007). They define potential experience as  $X^{abh} = IntY - SY^{abh}$ , so in their final sample  $X^{abh}$  takes on integer values from one to 12. We conform to this selection rule, but define potential experience as  $X^{lm} = (IntD - SD^{lm})/12$  so our potential experience measure ranges from 0.083 to 12.917 years. We believe our use of month-level measures is preferred to ABH’s more aggregated measures—but far more critical is the fact that ABH’s tendency to define “too early” a career start date translates into their assigning “too high” a level of potential experience, as demonstrated in III.C.

The remaining selection rule used for the “ $S=12/S=16$ , potential experience” samples (row  $m$  in table A1) produces the final discrepancy between ABH’s data and ours: ABH select observations for which  $S^{abh}$ —defined as “highest grade completed” *at the date of the interview*—is either 12 or 16. By using time-varying schooling, they introduce temporal separation between an individual’s career start year ( $SY^{abh}$ ) and his first appearance in the data with a given signal ( $S^{abh}$ ).<sup>12</sup> We require  $S^{lm}$  to equal either 12 or 16, but in contrast to ABH we define schooling attainment as time-constant “highest grade completed” at the career start date. To construct “all  $S$  levels” counterparts to these samples, time-varying  $S^{abh}$  and time-constant  $S^{lm}$  are restricted to values 8-20.

ABH do not report estimates based on actual work experience, but we explore this extension to determine whether discrepancies between their findings and ours are mitigated when we measure *actual* work experience rather than elapsed time since school exit. To construct comparable “actual experience” versions of the ABH and LM samples, we (i) convert both  $SY^{abh}$  and  $SD^{lm}$  to their corresponding start weeks; and (ii) count the number of weeks the respondent works at least 20 hours for a nonmilitary employer from the start week to the relevant interview week (and divide this cumulative measure by 52). Because weekly employment variables are only available from January 1978 onward, we drop all respondents with start weeks prior to this

---

<sup>12</sup>This allows 23 men to appear in ABH’s  $S=12$  subsample at low levels of  $X^{abh}$  and in their  $S=16$  subsample at higher levels of  $X^{abh}$ .

date. (See Appendix A for details.) The actual experience variables are  $ActX^{abh}$  and  $ActX^{lm}$ .

Table 1 reports sample sizes for each ABH and LM sample. Using the “S=12, potential experience” subsample for illustration, the ABH sample contains 11,795 observations for 1,926 men with  $S^{abh}=12$ , while the LM sample consists of 13,489 observations for 2,074 men with  $S^{lm}=12$ . ABH’s sample is smaller than ours primarily because they drop over 100 respondents with pre-1979 start years. Table 1 indicates that 10,909 of these wage observations (representing 92.5% of the entire ABH sample) appear in both the ABH and LM samples. For these 10,909 common observations, the ABH and LM samples necessarily have the same values for schooling and for variables that we define identically to ABH (logwage, AFQT scores, race, calendar year, urban status), but they often have different values for experience. We explore these discrepancies in section III.C. One goal of our analysis is to learn whether differences between ABH’s findings and ours are due to (i) different experience and/or schooling values for “common” observations; or (ii) observations that appear in one sample but not the other.

## B. Variables Used in Log-Wage Models

The preceding discussion defined several key variables used in our analysis: career start dates ( $SY^{abh}$  and  $SD^{lm}$ ), schooling ( $S^{abh}$  and  $S^{lm}$ ), potential experience ( $X^{abh}$  and  $X^{lm}$ ), and actual experience ( $ActX^{abh}$  and  $ActX^{lm}$ ). Definitions of all variables are summarized in table A2, and summary statistics for the alternative samples appear in table 2.

The log-wage models used throughout our analysis conform to specifications used by ABH. The dependent variable is the log of the average hourly wage (in cents), deflated by the CPI-U. All specifications include raw AFQT scores (standardized by age in 1980) as the empirical analog to  $Z$ , a cubic polynomial in either potential experience ( $X^{abh}$  or  $X^{lm}$ ) or actual experience ( $ActX^{abh}$  or  $ActX^{lm}$ ) as the empirical analog to  $t$ , and an interaction between AFQT scores and experience. All specifications also include a dummy variable indicating whether the respondent is black and a black-experience interaction, a dummy variable indicating whether the respondent lives in an urban area, and year fixed effects. Specifications that use “all  $S$  levels” samples include either  $S^{abh}$  or  $S^{lm}$  and their interactions with experience.

## C. Comparing Start Dates, Experience, and Schooling Across Samples

ABH trigger the start of the career ( $SY^{abh}$ ) the first time a respondent reports himself to be nonenrolled at the interview date. Our start date ( $SD^{lm}$ ) requires respondents to be at the start of a nonenrollment spell that lasts at least 12 months. While it is clear that ABH’s start year will often precede ours, in table 3 we demonstrate precisely how the two measures compare after aggregating our start date to the corresponding start year ( $SY^{lm}$ ) for comparability with  $SY^{abh}$ . We conduct the comparison for both the  $S=12$  and  $S=16$  potential experience subsamples, and we segment each subsample into observations that are common to the LM and ABH samples versus

those that appear only in the ABH sample.<sup>13</sup>

We focus first on the 10,909 observations for 1,672 men that appear in both the ABH and LM  $S=12$  subsamples. The upper right hand quadrant of table 3 reveals that we assign the same start year as ABH for 90.3% of men and 91.5% of observations in this subsample. The ABH start year precedes our start year for almost all remaining observations, most often by only one year. Our finding that the ABH and LM career start dates are in close agreement for this subsample is unsurprising: most NLSY79 respondents who terminate their schooling after completing the 12<sup>th</sup> grade do so before their 1979 interview (when they are ages 14-22) or within a year or two of that date, so there are few opportunities to be interviewed during an enrollment break prior to  $SY^{lm}$ . Discrepancies arise when there *are* such opportunities; *i.e.*, when younger respondents are interviewed during summer vacations or other short enrollment breaks prior to the completion of the 12<sup>th</sup> grade.

We see far less agreement between  $SY^{abh}$  and  $SY^{lm}$  among the 290 men and 886 observations in the “ABH only” portion of the  $S=12$  subsample. As summarized in table A4a, most of these observations are excluded from the LM sample because (i) the wage is reported prior to  $SD^{lm}$ ; (ii) the wage is reported after a return to school that terminates our observation window; and/or (iii) we assign a “highest grade completed” other than 12. While none of these reasons precludes  $SY^{abh}$  and  $SY^{lm}$  from being in agreement, this subsample is dominated by individuals with enrollment interruptions. As a result, we assign the same start year as ABH for only 52% of men and 65% of observations, and the mean difference in start years (conditional on being positive) is almost four years.

The most striking pattern seen in table 3 is that  $SY^{abh}$  is *far* more likely to precede  $SY^{lm}$  among individuals with  $S=16$  than among their less-schooled counterparts. In the “both” subsamples,  $SY^{abh} = SY^{lm}$  for 90% of men with  $S=12$  but only 65% of men with  $S=16$ , while  $SY^{abh}$  precedes  $SY^{lm}$  for 9.3% of men with  $S=12$  and almost 35% of men with  $S=16$ . The differences across schooling levels are less dramatic in the “ABH only” subsamples, where  $SY^{abh}$  precedes  $SY^{lm}$  for 45% of men with  $S=12$  and 56% of men with  $S=16$ . ABH systematically start the career earlier than we do for  $S=16$  workers because the longer an individual stays in school, the more likely he is to experience a short-term enrollment interruption (including summer vacation) *and* to be interviewed during a nonenrollment spell.

The finding that  $SY^{abh}$  precedes  $SY^{lm}$  for 35% of men appearing in both  $S=16$  samples raises the question: Are the discrepancies in start years dominated by individuals who complete 16

---

<sup>13</sup>Observations in the “ABH only” subsamples are dropped from the LM subsamples because they fail to meet the selection criteria described in II.A, but we can assign values for  $SD^{lm}$  and  $S^{lm}$  prior to dropping the observations. In contrast, we cannot always assign  $SY^{abh}$  and  $S^{abh}$  to observations in the “LM only” subsamples because item nonresponse is often what leads to their deletion from the ABH sample.

years of school well beyond the “expected” age of college completion? To assess this issue, we compute age at  $SY^{abh}$  and age at  $SD^{lm}$  for the 183 men in the “both/ $S=16$ ” sample for whom the ABH and LM start years disagree (table 3). For 112 of these individuals (61% of the sample), both ages are between 21 and 24, which we view as “expected” ages of college completion; another 12 individuals (7% of the sample) are older than 24 at both  $SY^{abh}$  and  $SD^{lm}$ . Another 23 individuals (13% of the sample) are younger than 21 at  $SY^{abh}$  but in the “expected” age range at  $SD^{lm}$ , while the remaining 36 individuals (19% of the sample) are in the “expected” age range at  $SY^{abh}$  but older than 24 at  $SD^{lm}$ . Individuals are older by construction at the LM start date than at the ABH start year, but fewer than one in five discrepancies correspond to a later-than-expected age at the LM start date *only* (and, as demonstrated later in this subsection, schooling attainment can be less than 16 at  $SY^{abh}$ ). We pursue this issue in section IV by determining whether our findings are sensitive to the age at school exit.

In table 4, we demonstrate that differences in the ABH and LM start dates translate into “mirror image” differences in potential experience.<sup>14</sup> Using the  $S=12$  “both” subsample for illustration, (i)  $SY^{abh} > SY^{lm}$  for 0.4% of observations (table 3) while  $X^{abh} < X^{lm}$  for 1.8% of observations (table 4); (ii)  $SY^{abh} = SY^{lm}$  for 91.5% of observations (table 3) and  $X^{abh} = X^{lm}$  for 88.4% of observations (table 4); and (iii)  $SY^{abh} < SY^{lm}$  for 8.2% of observations (table 3) while  $X^{abh} > X^{lm}$  for 9.7% of observations (table 4).<sup>15</sup> Among observations for which  $SY^{abh} < SY^{lm}$ , the mean difference between  $X^{abh}$  and  $X^{lm}$  is 1.18 years, which is close in magnitude to the mean difference in start years of 1.33 seen in table 3. The mean difference between  $X^{abh}$  and  $X^{lm}$  falls to 0.43 among all observations for which  $X^{abh} > X^{lm}$ ; these observations include many cases where  $SY^{abh} = SY^{lm}$ , but  $X^{lm}$  and  $X^{abh}$  differ slightly because of rounding. In fact, discrepancies between tables 3 and 4 arise solely because we measure elapsed time from start date to interview date in months (divided by 12) while ABH aggregate to elapsed time in years.

Table 5 shows the same summary statistics as table 4 for the  $S=12$  and  $S=16$  *actual* experience samples. If ABH tend to start the career too early for  $S=16$  workers, thereby overstating experience when measuring it as elapsed time since the start date, then switching from potential to actual experience will reduce their mismeasurement. However, if ABH’s early start date enables them to measure *actual* employment experience that precedes  $SD^{lm}$  (and that is potentially informative to employers), then discrepancies between  $X^{abh}$  and  $X^{lm}$  will remain when actual experience is substituted.

Table 5 reveals considerable agreement between  $ActX^{abh}$  and  $ActX^{lm}$ . The (rounded) difference between  $ActX^{abh}$  and  $ActX^{lm}$  equals zero for 96% of observations in the  $S=12$

---

<sup>14</sup>Table 4 is identical to table 3 in structure, but because  $X^{abh}$  and  $X^{lm}$  are time-varying we dispense with distributions based on samples containing one observation per person.

<sup>15</sup>The distributions in table 4 are for the difference  $X^{abh} - X^{lm}$  rounded to the nearest integer.

subsamples and over 80% of observations in the  $S=16$  subsamples. Moreover, the mean (unrounded) difference between the two experience measures conditional on being positive is only 0.15-0.39 for the four subsamples. (As shown in table 2, unconditional mean levels of actual experience are identical in the ABH and LM samples for both schooling groups.) It appears that ABH overstate *potential* experience (especially for  $S=16$  workers) because they measure elapsed time from “too early” a career start date, and include a considerable amount of nonemployment. Once we accumulate only those weeks in which individuals work, we can start the clock at either the ABH or LM start date and obtain roughly identical measures of experience.

ABH’s time-varying schooling measure is highest grade completed at the date of the interview, while our time-constant variable is highest grade completed at the career start date. In table 6, we assess discrepancies between the two variables, and consider the extent to which ABH’s measure differs from what employers would observe at the start of the career. Clearly, an observation appears in both the ABH and LM  $S=12$  or  $S=16$  subsamples only when  $S^{abh}$  and  $S^{lm}$  are identically equal to 12 or 16, so there are no discrepancies between the two measures in the “both” subsamples. Among observations that are used only by ABH, however,  $S^{abh}$  and  $S^{lm}$  are in agreement for only 8.4% (10%) of observations in the  $S=12$  ( $S=16$ ) subsample, and the mean difference in schooling conditional on  $S^{abh}$  being greater than  $S^{lm}$  is as high as 2.23 years in the  $S=16$  subsample. These discrepancies are unsurprising, given that observations are dropped from the LM sample because they precede the start date or follow a reenrollment spell, or because  $S^{lm}$  does not equal 12 or 16.

The more noteworthy finding in table 6 is that ABH’s time-varying  $S^{abh}$  exceeds the schooling value reported at ABH’s career start year for only 1% of observations in the “both”  $S=12$  subsample, but for 8.1% of observations in the “both”  $S=16$  subsample and 82.4% (62.6%) of observations in the “ABH only”  $S=12$  ( $S=16$ ) subsamples. Combining the “ABH only” and “both” subsamples into one,  $S^{abh}$  exceeds its value at the ABH career start date for 7% of observations in the  $S=12$  subsample and 14% of observations in the  $S=16$  subsample. By allowing this variable to be time-varying, ABH use a “pre-market signal” that differs from what employers would observe at  $X=0$  for a nontrivial number of observations.

#### **IV. Evidence of Employer Learning**

We begin by demonstrating that ABH’s key finding—that employer learning occurs for  $S=12$  workers but *not* for  $S=16$  workers—is driven by their choice of career start date. In column 1 of table 7, we report estimates for equation 5 using the ABH “ $S=12/S=16$  potential experience” samples; these estimates are identical to those reported in columns 1 and 3 of ABH’s table 2. The estimated coefficient for  $AFQT \cdot X/10$  is 0.126 for  $S=12$  workers and an imprecisely estimated 0.012 for  $S=16$  workers. Based on this evidence, ABH conclude that employers

observe pre-market productivity “nearly perfectly” at the outset of the career for  $S=16$  workers, but must learn it over time for  $S=12$  workers. In column 8 of table 7, we report analogous estimates based on the LM “ $S=12/S=16$  potential experience” sample. The estimated coefficients for  $AFQT \cdot X/10$  are about 0.10 for both schooling groups, which leads us to conclude that employer learning exists “equally” at both schooling levels.<sup>16</sup>

To ascertain what accounts for the difference between ABH’s findings and ours, in columns 2-7 of table 7 we report estimates based on restricted versions of the ABH and LM samples. ABH’s start date leads them to include wage observations that we deem to be earned prior to the start of the employer learning process, so in column 2 we discard from the ABH samples all observations that precede  $SD^{lm}$ . In column 3, we instead discard observations that succeed the date (if applicable) when we terminate the observation window due to school reenrollment. Neither sample restriction reverses ABH’s findings. In column 4, we drop from the ABH samples those observations for which ABH’s time-varying schooling measure ( $S^{abh}$ ) differs from our time-constant schooling measure ( $S^{lm}$ ). As demonstrated by table 6, the column 4 restriction causes most observations in the “ABH only” subsamples to be dropped—so in column 5 we go one step further and drop *all* observations from the “ABH only” subsamples. In columns 4-5 the estimated coefficient for  $AFQT \cdot X/10$  for  $S=16$  workers increases to an imprecisely estimated 0.05. Given that a disproportionate share of “ABH only” observations have discrepant ABH and LM starting dates (table 6), the estimates in columns 4-5 suggest that the measurement of career start dates (and the associated measurement of potential experience) is what drives the difference between ABH’s findings and ours rather than ABH’s use of time-varying schooling. This conjecture is supported by column 6, where we eliminate all observations for which  $SY^{abh} < SY^{lm}$  and obtain an estimated coefficient for  $AFQT \cdot X/10$  that is almost identical to the 0.097 in column 8. Column 6 demonstrates that once we eliminate observations with overstated potential experience from the ABH data, we obtain “our” results—*viz.*, equal evidence of employer learning for  $S=12$  and  $S=16$  workers.

We conduct additional experiments to establish that the column 8 findings are robust to changes in sample composition and adjustments to the definition of career start date. First, in column 7 of table 7 we use the LM data and variables, but drop observations that are not common to the ABH sample. The estimated coefficient for  $AFQT \cdot X/10$  is 0.108, which is

---

<sup>16</sup>Using conventional significance levels, we fail to reject the null hypothesis that the two parameter estimates are equal. We reach the same conclusion using alternative estimates (not reported) in which  $AFQT$  is replaced with standardized residual scores designed to be orthogonal to all other measures. In principle, only the use of residual scores allows a meaningful comparison of the magnitude of employer learning across samples (Light and McGee 2015). Because correlations between  $AFQT$  and other regressors are similar for the two schooling subsamples, however, we reach the same conclusion with respect to cross-sample tests of equality regardless of whether we use  $AFQT$  or residual test scores.

virtually identical to the estimate in column 8. This demonstrates that the column 8 findings are *not* driven by observations that are included in the LM sample but not in the ABH sample. Second, we revisit a concern raised in section III that our findings might be influenced by men who complete 16 years of school at later-than-expected ages. To pursue this issue, we restrict the  $S=16$  subsample used in column 8 of table 7 to 3,344 observations for 456 individuals whose age at  $SD^{lm}$  is 21-24. Using this sample of men who complete college at an “expected” age, we obtain an estimated coefficient for  $AFQT \cdot X/10$  (not tabulated) equal to 0.084. Clearly, our evidence of employer learning for  $S=16$  workers is not driven by late college completers. Third, to further investigate the possibility that we start the career too late, we redefine  $SD^{lm}$  as the start of the first period of nonenrollment lasting six months, rather than 12 months. These estimates are not tabulated, but the estimated coefficients for  $AFQT \cdot X/10$  are virtually identical to what is seen in column 8 (0.103 and 0.092 for the  $S=12$  and  $S=16$  samples, respectively). Fourth, for respondents with jobs in progress (and less than one month from ending) at  $SD^{lm}$ , we increment  $X^{lm}$  by the “pre-career” job duration to avoid discounting potentially informative work experience. This adjustment is made for relatively few respondents, and has a negligible effect on both mean levels of  $X^{lm}$  and on the estimates shown in column 8.

These experiments assuage concerns that  $SD^{lm}$  is defined “too late.” Clearly, the column 8 estimates do *not* depend on observations that post-date the ABH observation window, on men who finish college at older ages, on school exits lasting 12 months instead of six, or on the exclusion of select work experience that pre-dates the career start date. To explore further the latter issue, we turn to estimates based on actual work experience rather than elapsed time since the career start date. In doing so, we can determine whether  $SY^{abh}$  (despite its documented flaws) has the advantage of measuring early work experience that contributes to employer learning but is excluded from experience measures that start the clock at  $SD^{lm}$ .

In table 8 we reproduce each set of estimates appearing in table 7 after switching to the “ $S=12/S=16$  actual experience” samples. As discussed in II.C, actual experience is a problematic proxy for  $t$  (time over which public employer learning takes place) because it is likely to inform employers not only about pre-market productivity but also about labor force attachment, job mobility, and other factors that change over time. Nonetheless, we switch from potential experience to actual experience to see if ABH’s findings are robust to this change. We already established that (i) the LM and ABH start dates produce *identical* mean levels of actual experience for both schooling groups (table 2) and (ii) discrepancies between  $ActX^{abh}$  and  $ActX^{lm}$  in table 5 are *much* smaller than similar discrepancies in potential experience. Therefore, we expect the effect of ABH’s “too early” career start year to be mitigated when we replace potential experience with a measure of actual time spent working.

The table 8 estimates differ from those in table 7 in a number of key ways. First, although

for  $S=12$  workers the estimated AFQT-experience coefficient is largely invariant to whether we use ABH or LM variables, the switch from potential experience to actual experience causes this estimate to shrink from roughly 0.12 to roughly 0.07. Second, for  $S=16$  workers the switch to actual experience causes the estimated AFQT-experience coefficient to increase dramatically in columns 1-2. The data used for columns 1-2 are dominated by observations in which ABH “start the clock” on  $S=16$  workers who are still in school, so it is unsurprising that a measure of actual experience weakens ABH’s conclusion that employer learning does not occur for the college educated. Third, as a result of the patterns just described, in every column of table 8 we fail to reject the null hypothesis that the estimated coefficients for  $AFQT \cdot ActX$  are equal across schooling groups. The use of actual experience mitigates the effect of “too early” a career start date to the point that any use of ABH data (columns 1-6) reveals evidence of employer learning for  $S=12$  and  $S=16$  workers alike.<sup>17</sup>

As the final step of our investigation, we ask whether estimates based on “all  $S$  levels” samples are as sensitive as estimates for  $S=16$  samples to differences between ABH and LM career start dates. In table 9, we report estimates for equation 4 using data for workers with  $S$  ranging from 8 to 20; similar specifications and pooled schooling samples are used by Farber and Gibbons (1996), AP, Lange (2007) and virtually all other employer learning studies. Table 9 reveals that LM data produce slightly smaller estimated  $AFQT \cdot X$  coefficients than do the ABH data, especially for the potential experience samples (0.077 versus 0.094). However, neither these differences nor other differences between LM and corresponding ABH estimates are statistically significant. Just as tables 7-8 reveal that estimates for  $S=12$  samples are robust to whether ABH or LM data are used, table 9 demonstrates that estimates for “all  $S$  levels” samples are similarly robust. The likelihood of observing short-term enrollment interruptions (including summer breaks) in the NLSY79 increases with schooling attainment, so ABH’s understatement (overstatement) of career start dates (potential experience) is concentrated among college-educated workers. This mismeasurement necessarily drives the findings in subsamples of  $S=16$  workers, but it does not dominate the data in “all  $S$  levels” samples.

## V. Concluding Comments

Defining the career start date as the onset of the first nonenrollment spell lasting at least 12 months, we find that employer learning does *not* vary by schooling attainment. Specifically, we find “equal” evidence of employer learning for workers who have completed 12 and 16 years of

---

<sup>17</sup>As noted in II.A, in switching from  $X$  to  $ActX$  we discard observations for which the start date precedes January 1978. To isolate the effects of that sample selection rule, we reproduce the estimates in columns 1 and 8 of table 7 using actual experience samples, but potential experience variables. The estimated coefficients (standard errors) for  $AFQT \cdot X/10$  are 0.135 (.019) for the ABH  $S=12$  sample and 0.011 (.048) for the ABH  $S=16$  sample, and 0.114 (.018) and 0.097 (.044) for the corresponding LM samples. Clearly, these estimates are robust to the deletion of pre-1978 school leavers.

school. Our findings are robust to (i) whether we measure experience as elapsed time since  $X=0$  or actual weeks worked; (ii) shortening the 12-month nonenrollment cutoff to six months; (iii) excluding individuals who complete college after age 24; and (iv) including select work experience gained prior to the career start date. Our results contradict ABH, who find evidence of employer learning for  $S=12$  workers but not for  $S=16$  workers. However, our results are consistent with Lang and Siniver (2011) and Bordon and Braga (2013) who, to our knowledge, are the only other analysts to test for employer learning using samples of college-educated workers. These studies report clear-cut evidence of employer learning for college graduates in Israel and Chile, respectively.<sup>18</sup>

We provide evidence that ABH's finding is an artifact of their measurement strategy: they "start the clock" on an individual's career the first time he is nonenrolled *at the date of an NLSY79 interview*, regardless of whether that point-in-time nonenrollment spell represents a permanent school exit, a temporary school exit, or simply a summer or winter school break. College-goers are far more likely than terminal high school graduates to be interviewed during a school vacation prior to a "true" school exit, so ABH's definition leads them to overstate potential labor market experience (elapsed time since  $X=0$ ) for  $S=16$  workers relative to  $S=12$  workers. This, in turn, causes the estimated coefficient that identifies employer learning to be systematically biased toward zero. When we replace the ABH potential experience variable with actual experience—which can be accumulated from as early a starting point as one desires, given that it only accumulates time spent on the job—we reject the notion that employer learning occurs for  $S=12$  workers only.

Our analysis informs the employer learning literature by calling into question ABH's finding that employer learning (and workers' ability to signal pre-market productivity, which defines the "need" for employer learning) varies with schooling attainment. ABH's analysis is a key contribution to the literature that asks whether employer learning occurs homogeneously throughout the labor market. While other analysts allow employer learning to differ by blue collar versus white collar occupation (Bauer and Haisken-DeNew 2001), initial occupation (Mansour 2012), and the "importance" of a particular skill to the current occupation (Light and McGee 2015), ABH was the first study to ask whether learning differs by *worker* type rather than job type. Unfortunately, their provocative and widely-cited finding appears to be driven by systematic mismeasurement of potential experience for highly-schooled workers.

Our analysis also informs a broader issue that is often overlooked in the empirical literature. We argue that (i) school-to-work transitions tend to be less clear-cut than is assumed by

---

<sup>18</sup>Lang and Siniver (2011) find that employers need seven years to learn fully the productivity of "regular" college graduates, although they find no evidence of employer learning for graduates of elite institutions. The NLSY79 sample is dominated by college-goers at non-elite institutions.

theoretical models; *(ii)* analysts are therefore left to define a career start date as they see fit; and *(iii)* the definition of career start date can affect one's findings. We can point to a small number of studies dedicated to exploring the ambiguity of career start dates (Michael and Tuma 1984; Light 1998) and an equally small number of studies within the employer learning literature in which the authors take seriously the need to define an appropriate start date (Farber and Gibbons 1996; Pinkston 2006, 2009). Based on our findings, it appears that analysts exploring a range of "early career" topics would be well-advised to give careful consideration to the choice of career start date and associated measurement issues.

## References

- Altonji, Joseph G. and Charles R. Pierret. "Employer Learning and Statistical Discrimination." *Quarterly Journal of Economics* 116 (February 2001): 313-50.
- Arcidiacono, Peter, Patrick Bayer and Aurel Hizmo. "Beyond Signaling and Human Capital: Education and the Revelation of Ability." *American Economic Journal: Applied Economics* 2 (October 2010): 76-104.
- Bauer, Thomas K. and John P. Haisken-Denew. "Employer Learning and the Returns to Schooling." *Labour Economics* 8 (May 2001): 161-80.
- Becker, Gary. *Human Capital*, 3<sup>rd</sup> edition. Chicago: University of Chicago Press, 1993.
- Ben-Porath, Yoram. "The Production of Human Capital and the Life Cycle of Earnings." *Journal of Political Economy* 75 (August 1967): 352-65.
- Booth, Alison L., Marco Francesconi and Jeff Frank. "Temporary Jobs: Stepping Stones or Dead Ends?" *The Economic Journal* 112 (June 2002): F189-213.
- Bordón, Paola and Breno Braga. "Employer Learning, Statistical Discrimination and University Prestige." Unpublished paper, June 2013.
- Chuang, Hwei-Lin. "High School Youths' Dropout and Re-enrollment Behavior." *Economics of Education Review* 16 (April 1997): 171-86.
- Farber, Henry S. and Robert Gibbons. "Learning and Wage Dynamics." *Quarterly Journal of Economics* 111 (November 1996): 1007-47.
- Hotz, V. Joseph, Lixin Colin Xu, Marta Tienda, and Avner Ahituv. "Are There Returns to the Wages of Young Men from Working While in School?" *Review of Economics and Statistics* 84 (May 2002): 221-36.
- Keane, Michael P. and Kenneth I. Wolpin. "The Career Decisions of Young Men." *Journal of Political Economy* 105 (June 1997): 473-522.
- Lang, Kevin and Erez Siniver. "Why Is an Elite Undergraduate Education Valuable? Evidence from Israel." *Labour Economics* 18 (December 2011): 767-77.
- Lange, Fabian. "The Speed of Employer Learning." *Journal of Labor Economics* 25 (January 2007): 1-35.
- Light, Audrey. "The Effects of Interrupted Schooling on Wages." *Journal of Human Resources* 30 (Summer 1995a): 472-502.
- Light, Audrey. "Hazard Model Estimates of the Decision to Reenroll in School." *Labour Economics* 2 (December 1995b): 381-406.
- Light, Audrey. "Estimating Returns to Schooling: When Does the Career Begin?" *Economics of Education Review* 17 (February 1998): 31-45.
- Light, Audrey. "In-School Work Experience and the Returns to Schooling." *Journal of Labor Economics* 19 (January 2001): 65-93.
- Light, Audrey and Andrew McGee. "Employer Learning and the 'Importance' of Skill." *Journal of Human Resources* 50 (Winter 2015), forthcoming.

- Mansour, Hani. "Does Employer Learning Vary by Occupation?" *Journal of Labor Economics* 30 (April 2012): 415-44.
- Michael, Robert T. and Nancy Brandon Tuma. "Youth Employment: Does Life Begin at 16?" *Journal of Labor Economics* 2 (October 1984): 464-76.
- Mincer, Jacob. *Schooling, Experience, and Earnings*. New York: Columbia University Press (for NBER), 1974.
- Neumark, David. "Youth Labor Markets in the United States: Shopping Around Vs. Staying Put." *Review of Economics and Statistics* 84 (August 2002): 462-82.
- Parent, Daniel. "Work While in High School in Canada: Its Labour Market and Educational Effects." *Canadian Journal of Economics* 39 (November 2006): 1125-50.
- Pinkston, Joshua C. "A Test of Screening Discrimination with Employer Learning." *Industrial and Labor Relations Review* 59 (January 2006): 267-84.
- Pinkston, Joshua C. "A Model of Asymmetric Employer Learning with Testable Implications." *Review of Economic Studies* 76 (January 2009): 367-94.
- Ruhm, Christopher. "The Extent and Consequences of High School Employment." *Journal of Labor Research* 16 (Summer 1995): 293-303.
- Schönberg, Uta. "Testing for Asymmetric Learning." *Journal of Labor Economics* 25 (October 2007): 651-91.

### Sample Selection Criteria and Variable Definitions (ABH and LM samples)

We begin by detailing the sample selection criteria (summarized in **table A1**) that are used to construct the ABH and LM samples.<sup>19</sup> As part of this discussion, we explain how several key variables are defined; **table A2** provides a comprehensive list of variable definitions used by ABH and LM, while **table A3** shows how key variable definitions and selection rules used by ABH and LM compare to other NLSY79-based studies. We conclude this appendix by summarizing the selection criteria that cause observations to be excluded from the ABH sample or the LM sample (**tables A4a-b**).

#### Sample selection criteria (table A1)

- a. The NLSY79 sample consists of 12,686 male and female respondents born in 1957-64. All ABH and LM samples drop female respondents. LM drop 6,283 females using the official sex identifier based on the screener (reference number R02148.00; question name SAMPLE\_SEX). ABH drop only 6,282 female respondents because they use 1982 “interviewer remarks” on respondent sex (reference number R08102.00; question name Q15-2) to override SAMPLE\_SEX in one instance.
- b. All ABH and LM samples drop 1,000 men who are Hispanic. Remaining respondents are black or nonblack, non-Hispanic. ABH are left with 5,404 non-Hispanic male respondents rather than 5,403 because of their method for identifying sex (item a).
- c. In 1980, 11,914 NLSY79 respondents (94% of the original sample) were administered the Armed Services Vocational Aptitude Battery (ASVAB). Scores from a subset of those tests are used to construct a score for the Armed Forces Qualifications Test (AFQT). All ABH and LM samples drop the 338 non-Hispanic men for whom AFQT scores are unavailable.
- d. To identify a career start date, ABH identify the earliest year in which the respondent reports himself to be nonenrolled at the time of the interview. Nonenrolled respondents are asked the month and year of their last enrollment as a follow-up question. (Subsequent to the 1979 interview, this information is elicited only from respondents identified as having attended school since the last interview.) ABH use the reported “year last enrolled” as the career start date, discarding the corresponding “month last enrolled.” We call this variable  $SY^{abh}$  for (career) start year; it is variable *gy* (for “graduation year”) in ABH’s Stata code.

If  $SY^{abh}$  is prior to 1979, ABH drop the respondent unless he provides valid responses to additional questions on the number of weeks worked in 1975-77. This pre-1978 employment information is presumably intended to back-fill an actual experience measure from the career start year to January 1978, when detailed work histories begin for most NLSY79 respondents. This restriction is unnecessary for an analysis based on potential experience; ABH include an actual experience variable in their summary statistics (their table 1) and use it to produce a set of structural estimates reported in their online appendix, but the estimates in their published article rely exclusively on a measure of potential experience.

$SY^{abh}$  is defined for 4,727 of the 5,066 individuals who meet criteria a-c; 3,997 men remain after the pre-1979 deletion rule is applied.

- d.’ LM samples use  $SD^{lm}$  as the career start date, which is the month and year that starts the first

<sup>19</sup>The Stata file (named *createdata.do*) used by ABH to create their samples is available at:

<http://www.aeaweb.org/articles.php?doi=10.1257/app.2.4.76>

We describe ABH’s sample selection rules in a different order than what appears in their Stata program, and we omit several rules that prove to be nonbinding.

## Appendix A

nonenrollment spell lasting at least 12 months. (We define all month/year variables as (year·12)+month). We identify this date using two alternative information sources: First, we use “month and year last enrolled” reported by respondents who are nonenrolled at the interview date. Starting with 1979 and proceeding forward, if (i) the interview date is at least 12 months past this “last enrolled” date or (ii) the next interview date is at least 12 months past this date and there is no increment to schooling attainment during the interim, we define our first potential start date ( $SD1$ ). Second, we use the monthly enrollment timeline constructed for all respondents from January 1980 onward to identify the date of the first nonenrollment spell lasting at least 12 months; this alternative measure is  $SD2$ .

We choose among the alternative start dates using these decision rules, in order: (i)  $SD^{lm} = SD2$  for 14% of respondents for whom  $SD1$  is undefined (typically due to missed interviews and/or missing “month and year last enrolled”); (ii)  $SD^{lm} = SD1$  for 43% of respondents for whom  $SD1$  and  $SD2$  are within three months of each other; (iii)  $SD^{lm} = SD1$  for 35% of respondents for whom  $SD1$  precedes January 1980; (iv)  $SD^{lm} = SD2$  for 4% of respondents because “month and year last enrolled” reported in an interview other than the one used to define  $SD1$  is the same date as  $SD2$ , which we use as evidence that  $SD2$  is an accurate date; (v)  $SD^{lm} = SD1$  for the remaining 4% of respondents.

$SD^{lm}$  is defined for 5,039 of the 5,065 men who meet criteria *a-c*.

- e. Respondents who meet criteria *a-d* or *a-d'* remain in the sample if at least one wage observation meets subsequent criteria.
- f. NLSY79 respondents were interviewed annually from 1979 to 1994 and biennially from 1996 to 2010. All ABH and LM samples use data for 1979-2004.
- g. ABH require an otherwise valid wage to be reported in an interview year that exceeds  $SY^{abh}$  by at least  $\Delta S$ , where  $\Delta S$  is the difference between  $S^{abh}$  (highest grade completed that year) and the minimum  $S^{abh}$  from  $SY^{abh}$  onward. (A precise definition of  $S^{abh}$  is given in our discussion of row *m* and in table A2.) Consider an example where  $SY^{abh}=79$  and the sequence of  $S^{abh}$  values in interview years 1979-1983 is 12,14,15,16,16, so the  $\Delta S$  sequence is 0,2,3,4,4. ABH’s selection rule leads them to drop the 1980-82 observations and use only the 1979 and 1983 observations. If we change the  $S^{abh}$  sequence to 12,13,14,15,16, ABH do not drop any observations. ABH do not discuss their rationale for the  $\Delta S$  aspect of selection rule *g*.
- g.’ LM samples use a straightforward alternative to rule *g*: the interview date at which an otherwise valid wage is reported must be greater than  $SD^{lm}$ .
- h. LM samples impose an additional selection criterion not used by ABH: the interview date at which an otherwise valid wage is reported must precede any reenrollment in school that triggers an increment to highest grade completed. Far fewer increments to schooling occur after  $SD^{lm}$  than after  $SY^{abh}$ , but when individuals *do* reenroll in school after a nonenrollment spell lasting at least 12 months we terminate the observation window.
- i. All ABH and LM samples use wages for the “current or last job” reported at each interview, and require that the nominal average hourly be between \$1 and \$100. The wage variable is a NLSY79 constructed variable based on information reported directly by respondents.
- j. All ABH and LM samples require the residential location at the time of the interview to be classified as either urban or rural; this selection rule simply excludes observations where the urban/rural variable takes on a missing value. We use this criterion (along with most others)

## Appendix A

for comparability with ABH.

- k. All ABH and LM samples require the “class of worker” variable associated with the “current or last job” to be coded as either “private company” or “government.” This excludes workers who are self-employed or working without pay or, for 1994-2004, working for a family business or nonprofit organization. All ABH and LM samples also require that employment status at the interview date be either “working” or “with a job.”
- l. ABH require  $0 < X^{abh} < 13$  where  $X^{abh}$  (*ptexp* in ABH’s Stata program) is potential experience, defined as  $IntY - SY^{abh}$  (interview year minus career start year). By aggregating both career start dates and interview dates to the year-level, ABH lose variation in potential experience. For example, an individual who leaves school in January 1980 and is interviewed in December 1981 and an individual who leaves school in December 1980 and is interviewed in January 1981 are both assigned  $X^{abh} = 1$  despite having been out of school for 23 months and 1 month, respectively. Valid observations for  $X^{abh}$  take on integer values from one to 12.
- l.’ LM require  $0 < X^{lm} < 13$ , where  $X^{lm}$  (potential experience) is defined as  $(IntD - SD^{lm})/12$  (interview data minus career start date, divided by 12); valid observations for  $X^{lm}$  range from 0.083 to 12.917.

Note that criteria *a-l* or *a-l'* are imposed for *all* ABH and LM samples used in this analysis.

- m. To select a sample of workers with  $S=12$  or  $S=16$ , ABH require the variable  $S^{abh}$  (*educ* in their Stata program) to equal 12 or 16.  $S^{abh}$  is a *time-varying* measure of highest grade completed at the given interview date. Specifically,  $S^{abh}$  is the (revised) created variable identifying the respondent’s highest grade completed on May 1 of each interview year (although ABH inadvertently used the nonrevised variables for 2002 and 2004).
- m.’ LM require the variable  $S^{lm}$  to equal 12 or 16, where  $S^{lm}$  is the (revised) created variable identifying the respondent’s highest grade completed on May 1 of the interview year corresponding to  $SD^{lm}$ .  $S^{lm}$  is a *time-constant* variable. Because we terminate the observation window at the start of a reenrollment spell that triggers an increment to highest grade completed (criterion *h*), there is little scope for employers to observe a level of schooling attainment other than  $S^{lm}$ .

Rules *a-m* produce the ABH “ $S=12/S=16$ , potential experience” sample used by ABH for the estimates in columns 1 and 3 of their table 2. Rules *a-m'* produce the LM “ $S=12/S=16$ , potential experience” sample. See table 1 for sample sizes.

- n. To produce a sample that is not restricted to  $S^{abh}=12$  or 16, ABH replace rule *m* with a requirement that time-varying  $S^{abh}$  takes on a value between 8 and 20.
- n.’ To produce a comparable LM sample, we replace rule *m'* with a requirement that time-constant  $S^{lm}$  takes on a value between 8 and 20.

Rules *a-l+n* produce the ABH “all *S* levels, potential experience” sample. Rules *a-l'+n'* produce the LM “all *S* levels, potential experience” sample. See table 1 for sample sizes.

- o. To produce alternative versions of the ABH samples in which potential experience is replaced with actual experience, we necessarily depart from their data construction strategy. Our goal is to construct a detailed measure of actual experience based on the weekly “hours” array and “employment status” array that are part of the NLSY79 work history. First, we match  $SY^{abh}$  with the corresponding “month last enrolled” ( $SM^{abh}$ ) that ABH discard. Second, we drop men who start their career prior to January 1978, when the work history

## Appendix A

arrays begin. Third, we use the week corresponding to the midpoint of the month as the career start week, and the week corresponding to each interview date as the stop week. Fourth, we use the arrays to count the number of weeks between these start and stop weeks in which the individual works at least 20 hours/week for a nonmilitary employer. This cumulative measure (divided by 52) is  $ActX^{abh}$ , which represents *our* measure (not ABH's measure) of actual experience for the ABH sample.

*o.*' We produce "actual experience" versions of the LM samples using the strategy described under criterion *o*. When applied to LM samples, the career start week is the week corresponding to the midpoint of the month defined by  $SD^{lm}$ .

Rules  $a-l+m+o$  ( $a-l+n+o$ ) produce the ABH " $S=12/S=16$ , actual experience" (all  $S$  levels, actual experience") sample. Rules  $a-l'+m'+o'$  ( $a-l'+n'+o'$ ) produce the corresponding LM samples. See table 1 for sample sizes.

### Observations excluded from the ABH or LM samples (tables A4a-b)

Table 1 reveals that 886 (468) wage observations are included in the ABH  $S=12$  ( $S=16$ ) potential experience subsamples but are excluded from the LM samples, while 2,580 (297) wage observations are included in the LM  $S=12$  ( $S=16$ ) potential experience subsamples but excluded from the ABH sample. To conclude our comparison of ABH and LM sample construction, we summarize the selection criteria that account for these "non-common" observations.

Table A4a reveals that 16% of "ABH only" observations in the  $S=12$  subsample and 19% of observations in the  $S=16$  subsample do not appear in the LM sample because they are reported prior to the LM career start date ( $SD^{lm}$ ). This is unsurprising given that ABH often start the career earlier than LM by using the start date of summer vacations and other short-term enrollment breaks. Another 26% of  $S=12$  observations and 62% of  $S=16$  observations are excluded from the LM sample because they are reported after we terminate the observation window due to school reenrollment. This indicates that ABH identify a nontrivial number of  $S=16$  workers who attain that schooling level after leaving school for at least 12 months and then reenrolling. Another 55% of  $S=12$  observations and 18% of  $S=16$  observations are excluded from the LM sample because our "pre-career" measure of schooling attainment ( $S^{lm}$ ) is neither 12 nor 16 for those workers. This occurs if, for example, an individual starts his career with  $S=11$  and subsequently increments "highest grade completed" to 12. Given that ABH use an aggregated (or "rounded") measure of potential experience, there are a small number of additional observations where our more accurate measure ( $X^{lm}$ ) is out of range.

Turning to the "LM only" observations summarized in table A4b, 25% (28%) of observations in the  $S=12$  ( $S=16$ ) subsample are excluded from the ABH sample because the career start date ( $SY^{abh}$ ) is undefined—*i.e.*, "year last enrolled" takes on a missing value the first time the respondent is nonenrolled at the interview date. We are often able to substitute alternative career start dates based on monthly enrollment timelines when "month and year last enrolled" are missing. Another 63% of  $S=12$  observations are excluded from the ABH sample because  $SY^{abh}$  precedes 1979 (and, presumably, information on annual weeks worked in 1975-77 is missing). This factor only affects the  $S=12$  subsample because no sample member (all of whom were born in 1957-64) completed 16 years of school prior to 1978. Another 11% of  $S=12$  observations and 65% of  $S=16$  observations are dropped by ABH because potential experience is out of range. Because the ABH start date often precedes the LM start date, we include many observations for which  $X^{abh} > 13$ . The small number of remaining observations are excluded from the ABH sample because their (time-varying) schooling variable is out of range.

Appendix A

**Table A1:** Selection Criteria for Alternative ABH and LM Samples

Selection criterion	ABH	LM
<b>Keep individual if</b>		
a. Male	✓	✓
b. Non-Hispanic	✓	✓
c. Valid AFQT score	✓	✓
d. Career start year ( $SY^{abh}$ ) is defined; <i>i.e.</i> , “year last enrolled” reported at the first “currently nonenrolled” interview is 1979 or later (or earlier than 1979 only if weeks worked before 1979 are reported)	✓	
d'. Career start date ( $SD^{lm}$ ) is defined; <i>i.e.</i> , month/year that starts the first 12-month-long nonenrollment spell precedes the last interview date		✓
e. $\geq 1$ wage observation using criteria below	✓	✓
<b>Keep annual wage observation for each interview year if</b>		
f. Interview year between 1979 and 2004	✓	✓
g. Interview year exceeds $SY^{abh}$ by at least $\Delta S$ years, where $\Delta S$ is the amount by which highest grade completed at interview ( $S^{abh}$ ) exceeds the minimum, post- $SY^{abh}$ highest grade	✓	
g'. Interview date (month/year) exceeds $SD^{lm}$		✓
h. Interview date (month/year) precedes first reenrollment date		✓
i. Nominal average hourly wage is \$1-\$100	✓	✓
j. Residence at interview date is urban or rural	✓	✓
k. Class of worker is private or government, and employment status is working or with a job	✓	✓
<b>For “potential experience” and all other samples</b>		
l. $0 < X^{abh} < 13$ ( $X^{abh}$ is elapsed years from $SY^{abh}$ to interview year)	✓	
l'. $0 < X^{lm} < 13$ ( $X^{lm}$ is elapsed months from $SD^{lm}$ to interview date $\div 12$ )		✓
<b>For “S=12/S=16” samples</b>		
m. $S^{abh}$ (highest grade completed at interview date) = 12 or 16	✓	
m'. $S^{lm}$ (highest grade completed at $SD^{lm}$ ) = 12 or 16		✓
<b>For “all S levels” samples</b>		
n. Replace $m$ with $8 \leq S^{abh} \leq 20$	✓	
n'. Replace $m'$ with $8 \leq S^{lm} \leq 20$		✓
<b>For “actual experience” samples</b>		
o. $SY^{abh}$ and corresponding start month no earlier than January 1978; combine with $l$ , and either $m$ or $n$ . <sup>a</sup>	✓	
o'. $SD^{lm}$ no earlier than January 1978; combine with $l'$ , and either $m'$ or $n'$		✓

<sup>a</sup>This selection rule is not used by ABH; we apply it to their data to define actual experience measures comparable to those defined for the LM sample.

**Table A2: Variable Definitions**

Variable	Definition
<b>Variables common to ABH and LM samples</b>	
$\log\text{wage}^*$	Log of average hourly wage (in cents), deflated by the CPI-U for all urban consumers (1990=100)
1 if black <sup>*</sup>	1 if non-Hispanic black, 0 if non-Hispanic nonblack
1 if urban <sup>*</sup>	1 if residence at interview date is urban, 0 if rural
AFQT score <sup>*</sup>	Raw AFQT score standardized by age [empirical analog to Z]
Year <sup>*</sup>	Calendar year dummies (1979-2004) indicating when wage is reported
$\text{Int}Y$	Year in which interview takes place; $\text{Int}M$ is corresponding month
$\text{Int}D^{\S}$	Date at which interview takes place ( $\text{Int}Y \cdot 12 + \text{Int}M$ )
<b>Variables specific to ABH or LM samples</b>	
$\text{SY}^{abh}$	ABH career start year (named $gy$ by ABH), defined as “year last enrolled” reported at first interview when nonenrolled [empirical analog to $t=0$ ]
$\text{SM}^{abh\ \S}$	“Month last enrolled” corresponding to $\text{SY}^{abh}$
$\text{SD}^{lm}$	LM career start date, defined as year/month ( $\text{SY}^{lm} \cdot 12 + \text{SM}^{lm}$ ) that begins first nonenrollment spell lasting at least 12 months [empirical analog to $t=0$ ]
$\text{SY}^{lm}$	Year corresponding to $\text{SD}^{lm}$ ; used for comparison with $\text{SY}^{abh}$
$\text{RD}^{lm}$	LM reenrollment date, defined as year/month (year $\cdot$ 12 + month) that begins first reenrollment spell that leads to an increment in highest grade completed
$\text{S}^{abh\ *}$	ABH (time-varying) schooling attainment (named $educ$ by ABH), defined as (revised, created) highest grade completed on May 1 of year in which wage is reported <sup>a</sup> [empirical analog to S]
$\text{S}^{lm\ *}$	LM (time-constant) schooling attainment, defined as (revised, created) highest grade completed on May of year corresponding to $\text{SD}^{lm}$ [empirical analog to S]
$\text{X}^{abh\ *}$	ABH potential experience (named $ptexp$ by ABH), defined as $\text{Int}Y - \text{SY}^{abh}$ ; takes on integer value from 1 to 12 [empirical analog to $t$ ]
$\text{X}^{lm\ *}$	LM potential experience, defined as $(\text{Int}D - \text{SD}^{lm})/12$ ; takes on values from 0.0833 to 12.9167 [empirical analog to $t$ ]
$\text{Act}X^{abh\ *\ \S}$	ABH actual experience, defined as (number of weeks worked at least 20 hours/week for a nonmilitary employer)/12, from the week corresponding to the midpoint of $\text{SM}^{abh}$ and $\text{SY}^{abh}$ to the week corresponding to $\text{Int}D$ ; defined only if $\text{SM}^{abh}$ and $\text{SY}^{abh}$ are no earlier than January 1978 [empirical analog to $t$ ]
$\text{Act}X^{lm\ *}$	LM actual experience, defined as for ABH, using $\text{SD}^{lm}$ as starting point [empirical analog to $t$ ]

\*Variables used in regressions.

<sup>\S</sup> Variables are not used by ABH, but are used to here to construct ABH actual experience ( $\text{Act}X^{abh}$ ) for comparison with  $\text{Act}X^{lm}$ .

<sup>a</sup>ABH use nonrevised “May 1 highest grade completed” in 2002-04.

Appendix A

**Table A3:** Measurement Strategies Used by Select NLSY79-Based Employer Learning Studies

Study	Career start date ( $t=0$ )	Schooling ( $S$ )	Experience ( $X$ )
Farber & Gibbons 1996	Transition from a “primarily nonworking” period lasting at least one year to a “primarily working” period lasting at least three years. <sup>a</sup>	Unclear whether $S$ defined at $t=0$ or $t$	<i>PotX</i> : Elapsed time since $t=0$ <i>ActX</i> : none
Altonji & Pierret 2001 (AP)	Month/year of last enrollment reported at 1 <sup>st</sup> non-enrolled interview	$S$ measured at $t$ (time-varying; must be non-decreasing); keep observations after reenrollment	<i>PotX</i> : Age- $S-6$ <i>ActX</i> : Weeks worked >30 hours since $t=0$ , divided by 50 IV for <i>ActX</i> : <i>PotX</i>
Pinkston 2006	Month/year of final school exit	Unclear whether $S$ defined at $t=0$ or $t$	<i>PotX</i> : Ages- $S-6$ <i>ActX</i> : Weeks worked since $t=0$ , divided by 52
Lange 2007	Same as ABH	$S$ measured at $t$ (time-varying)	<i>PotX</i> : Elapsed years since $t=0$
Schönberg 2007	Transition from “non-working” period lasting at least one year to “working” period lasting at least two years <sup>b</sup>	$S$ measured at $t$ (time-varying)	<i>PotX</i> : none <i>ActX</i> : Cumulative job durations since $t=0$ and other measures <sup>c</sup>
Arcidiacono <i>et al.</i> 2010 (ABH)	Year of last enrollment reported at 1 <sup>st</sup> nonenrolled interview	$S$ measured at $t$ (time-varying); keep observations after reenrollment	<i>PotX</i> : Elapsed years since $t=0$
Mansour 2012	Month/year of 1 <sup>st</sup> school exit (unclear whether 1 <sup>st</sup> school exit is identified as in AP)	Unclear whether $S$ defined at $t=0$ or $t$ ; keep observations after reenrollment	<i>PotX</i> : None <i>ActX</i> : Same as AP IV for <i>ActX</i> : none
Light & McGee 2015	Same as AP	$S$ measured at $t=0$ (time-constant); terminate observation window upon reenrollment	<i>PotX</i> : Elapsed months since $t=0$ , divided by 12
Light & McGee 2013 (this study)	Month/year that starts 1 <sup>st</sup> nonenrollment spell lasting at least 12 months	$S$ measured at $t=0$ (time-constant); terminate observation window upon reenrollment	<i>PotX</i> : Elapsed months since $t=0$ , divided by 12 <i>ActX</i> : Weeks worked >20 hours on nonmilitary jobs since $t=0$ , divided by 52 IV for <i>ActX</i> : none

<sup>a</sup>Respondents are considered “primarily working” during intervals between interviews (approximately one year in duration) if they work at least half the weeks and average at least 30 hours/week.

<sup>b</sup>Respondents are considered to be “working” if they work at least 26 weeks during a calendar year.

<sup>c</sup>Alternative experience measures used to test for asymmetric employer learning.

Appendix A

**Table A4a:** Diagnostics for “ABH Only” Observations (included in ABH but not in LM) in “S=12/S=16, Potential Experience” Sample

Excluded from LM sample due to:	$S^{abh}=12$	
	# obsns	% obsns
Prior to career start date ( $SD^{lm}$ )	144	16.3
After reenrollment	233	26.3
$S^{lm} \neq 12$ or 16	491	55.4
$X^{lm} \leq 0$ or $X^{lm} \geq 13$	18	2.0
Total	886	100.0
	$S^{abh}=16$	
Prior to career start date ( $SD^{lm}$ )	90	19.2
After reenrollment	289	61.8
$S^{lm} \neq 12$ or 16	85	18.2
$X^{lm} \leq 0$ or $X^{lm} \geq 13$	4	0.9
Total	468	100.0

Note: Reasons for being excluded from the LM sample are not mutually exclusive; observation counts are for observations not already counted as a preceding reason.

**Table A4b:** Diagnostics for “LM only” Observations (included in LM but not in ABH) in “S=12/S=16, Potential Experience” Sample

Excluded from ABH sample due to:	$S^{lm} = 12$	
	# obsns	% obsns
$SY^{abh} = .$ (career start year undefined)	643	24.9
$SY^{abh} < 1978$	1,626	63.0
$X^{abh} \leq 0$ or $X^{abh} \geq 13$	287	11.1
$S^{abh} \neq 12$ or 16	24	0.9
Total	2,580	100.0
	$S^{lm} = 16$	
$SY^{abh} = .$ (career start year undefined)	83	27.9
$X^{abh} \leq 0$ or $X^{abh} \geq 13$	194	65.3
$S^{abh} \neq 12$ or 16	20	6.7
Total	297	100.0

Note: Reasons for being excluded from the ABH sample are not mutually exclusive; observation counts are for observations not already counted as a preceding reason.

**Table 1:** Sample Sizes for Alternative ABH and LM Samples

Sample	ABH	LM	Sample Overlap			
			ABH only	LM only	Both	%ABH in both
S=12, potential experience <sup>a</sup>						
Number of observations	11,795	13,489	886	2,580	10,909	92.5
Number of men	1,926	2,074	290	618	1,672	86.8
S=16, potential experience <sup>a</sup>						
Number of observations	4,112	3,941	468	297	3,644	88.6
Number of men	650	544	155	141	526	80.9
All S levels, potential experience <sup>b</sup>						
Number of observations	25,692	28,300	3,558	6,166	22,134	86.2
Number of men	3,674	4,399	927	1,594	3,488	94.9
S=12, actual experience <sup>c</sup>						
Number of observations	9,866	10,050	804	988	9,062	91.9
Number of men	1,525	1,417	259	317	1,301	85.2
S=16, actual experience <sup>c</sup>						
Number of observations	4,098	3,941	454	297	3,644	88.9
Number of men	645	544	150	141	526	81.6
All S levels, actual experience <sup>d</sup>						
Number of observations	22,975	22,731	3,347	3,103	19,628	85.4
Number of men	3,154	3,351	845	999	2,994	94.9

Note: individuals can appear in both the “ABH only” (or “LM only”) and “both” subsamples, so summing the number of men across these columns does not produce the total number of men for the total ABH (or LM) sample.

<sup>a</sup>The ABH (LM) sample is based on criteria  $a-m$  ( $a-m'$ ) in table A1. Because  $S^{abh}$  is time-varying, 23 men appear in both the S=12 and S=16 ABH samples (*i.e.*, 2,553 unique individuals appear in the combined S=12 and S=16 samples).

<sup>b</sup>The ABH (LM) sample is based on criteria  $a-l$  and  $n$  (or  $a-l'$  and  $n'$ ) in table A1.

<sup>c</sup>The ABH (LM) sample is based on criteria  $a-m$  and  $o$  ( $a-m'$  and  $o'$ ) in table A1. Because  $S^{abh}$  is time-varying, 20 men appear in both the S=12 and S=16 ABH samples.

<sup>d</sup>The ABH (LM) sample is based on criteria  $a-l$  and  $n-o$  ( $a-l'$  and  $n'-o'$ ) in table A1.

**Table 2:** Means and Standard Deviations for Alternative ABH and LM Samples

Variable	Potential experience			Actual experience		
	S=12	S=16	All S	S=12	S=16	All S
<b>ABH samples</b>						
logwage (avg. hourly wage in cents) *	6.65 (.43)	7.10 (.46)	6.77 (.50)	6.63 (.43)	7.10 (.46)	6.78 (.51)
$S^{abh}$ (highest grade completed) *	—	—	13.20 (2.28)	—	—	13.35 (2.34)
AFQT score (raw, age-standardized) *	-.08 (.93)	.97 (.60)	.20 (1.00)	-.08 (.94)	.97 (.60)	.24 (1.00)
1 if black *	.31	.18	.28	.33	.18	.28
1 if urban *	.74	.87	.78	.74	.87	.79
$X^{abh}$ (potential experience in years) *	6.61 (3.30)	5.92 (3.30)	6.47 (3.31)	6.48 (3.38)	5.91 (3.29)	6.37 (3.35)
$ActX^{abh}$ (actual experience in weeks divided by 52) *	—	—	—	4.91 (3.22)	5.14 (3.21)	4.96 (3.20)
$SY^{abh}$ (career starting year)	79.48 (2.31)	82.82 (2.63)	80.52 (2.94)	80.16 (1.88)	82.84 (2.60)	81.05 (2.62)
Number of observations	11,795	4,112	25,692	9,866	4,098	22,975
Number of men	1,926	650	3,674	1,525	645	3,154
<b>LM samples</b>						
logwage (avg. hourly wage in cents) *	6.68 (.43)	7.13 (.46)	6.76 (.49)	6.65 (.43)	7.13 (.46)	6.78 (.51)
$S^{lm}$ (highest grade completed) *	—	—	12.65 (2.28)	—	—	13.00 (2.33)
AFQT score (raw, age-standardized) *	-.05 (.92)	.97 (.61)	.06 (1.02)	-.08 (.94)	.97 (.61)	.14 (1.01)
1 if black *	.31	.19	.30	.33	.19	.30
1 of urban *	.74	.87	.77	.74	.87	.78
$X^{lm}$ (potential experience in months divided by 12) *	6.49 (3.47)	5.55 (3.49)	6.24 (3.52)	6.33 (3.61)	5.55 (3.49)	6.03 (3.60)
$ActX^{lm}$ (actual experience in weeks divided by 52) *	—	—	—	4.95 (3.28)	5.15 (3.36)	4.90 (3.29)
$SD^{lm}$ (career starting date, months since 1900)	958.08 (29.98)	1008.68 (31.58)	971.32 (40.62)	970.88 (22.95)	1008.8 (31.58)	984.04 (34.57)
$SY^{lm}$ (year corresponding to $SD^{lm}$ )	79.32 (2.52)	83.61 (2.67)	80.46 (3.42)	80.40 (1.92)	83.61 (2.67)	81.53 (2.90)
Number of observations	13,489	3,941	28,300	10,050	3,941	22,731
Number of men	2,074	544	4,399	1,417	544	3,351

Note: See table A2 (appendix A) for variable definitions.

\*Variables used in regressions, along with experience squared and cubed, an AFQT-experience interaction, a black-experience interaction, a S-experience interaction (when using “all S” samples), and year fixed effects.

**Table 3:** Comparison of ABH and LM Starting Years  
for “S=12/S=16, Potential Experience” Samples

	“ABH only” subsample				“Both” subsample			
	$S^{abh}=12$				$S^{abh} = S^{lm}=12$			
$SY^{lm} - SY^{abh}$ is:	# men	% men	# obs	% obs	# men	% men	# obs	% obs
-2 or smaller	7	2.4	24	2.7	6	.4	18	.2
-1	2	.7	2	.2	3	.2	20	.2
0	151	52.1	572	64.6	1,509	90.3	9,982	91.5
1	45	15.5	145	16.4	118	7.1	706	6.5
2 or more	85	29.3	143	16.1	36	2.2	183	1.7
All	290	100.0	886	100.0	1,672	100.0	10,909	100.0
Mean $(SY^{lm} - SY^{abh}) > 0^a$ (std. dev.)	3.92 (3.96)		3.80 (4.59)		1.44 (1.03)		1.33 (.82)	
	$S^{abh}=16$				$S^{abh} = S^{lm}=16$			
$SY^{lm} - SY^{abh}$ is:	# men	% men	# obs	% obs	# men	% men	# obs	% obs
-2 or smaller	4	2.6	15	3.2				
-1	2	1.3	11	2.4				
0	62	40.0	259	55.3	343	65.2	2,465	67.7
1	26	16.8	71	15.2	112	21.3	841	23.1
2 or more	61	39.4	112	23.9	71	13.5	338	9.3
All	155	100.0	468	100.0	526	100.0	3,644	100.0
Mean $(SY^{lm} - SY^{abh}) > 0^a$ (std.dev.)	3.39 (2.47)		3.22 (2.65)		2.25 (2.03)		1.82 (1.58)	

Note:  $SY^{lm}$  is the calendar year containing  $SD^{lm}$ ;  $SY^{abh}$  is the career starting year defined by ABH. Observations in the “ABH only” sample do not appear in the LM sample, but we are able to define  $SD^{lm}$  and  $SY^{lm}$  for all men/observations used by ABH.

<sup>a</sup>Mean gap conditional on being positive.

**Table 4:** Comparison of ABH and LM Potential Experience for “S=12/S=16, Potential Experience” Samples

	“ABH only”		“Both”	
	$S^{abh}=12$		$S^{abh} = S^{lm}=12$	
$X^{abh} - X^{lm}$ (rounded) is:	# obs	% obs	# obs	% obs
-2 or smaller	24	2.7	20	.2
-1	45	5.1	82	1.6
0	497	56.1	9,230	88.4
1	163	18.4	1,373	8.0
2 or more	157	17.7	204	1.7
All	886	100.0	10,909	100.0
Mean $(X^{abh} - X^{lm})   >0^a$ (std. dev.)	2.08 (3.76)		0.43 (.46)	
Mean $(X^{abh} - X^{lm})   SY^{lm} > SY^{abh}$ (std. dev.)	3.79 (4.69)		1.18 (.95)	
	$S^{abh}=16$		$S^{abh} = S^{lm}=16$	
$X^{abh} - X^{lm}$ (rounded) is:	# obs	% obs	# obs	% obs
-2 or smaller	15	3.2		
-1	21	4.5	31	.9
0	263	56.2	2,534	69.5
1	55	11.8	738	20.3
2 or more	114	24.4	341	9.4
All	468	100.0	3,644	100.0
Mean $(X^{abh} - X^{lm})   >0^a$ (std. dev.)	2.48 (2.71)		0.90 (1.36)	
Mean $(X^{abh} - X^{lm})   SY^{lm} > SY^{abh}$ (std. dev.)	3.13 (2.76)		1.53 (1.69)	

Note:  $X^{abh}$  and  $X^{lm}$  are potential experience defined for the ABH and LM samples, respectively. Observations in the “ABH only” sample do not appear in the LM sample, but we are able to define  $X^{abh}$  for all observations used by ABH.

<sup>a</sup>Mean difference between  $X^{abh}$  and  $X^{lm}$  (not rounded) conditional on being positive.

**Table 5:** Comparison of ABH and LM Actual Experience  
for “S=12/S=16, Actual Experience” Samples

	“ABH only”		“Both”	
	$S^{abh}=12$		$S^{abh} = S^{lm}=12$	
$ActX^{abh} - ActX^{lm}$ (rounded) is:	# obs	% obs	# obs	% obs
-2 or smaller	9	1.4	6	.1
-1	7	1.1	10	.1
0	603	95.8	8,691	95.9
1	19	2.9	276	3.1
2 or more	8	1.2	79	.9
All	646	100.0	9,062	100.0
Mean ( $ActX^{abh} - ActX^{lm}$ )   $ActX^{abh} > ActX^{lm}$ (std. dev.)	0.20 (.31)		0.15 (.36)	
Mean ( $ActX^{abh} - ActX^{lm}$ )   $SY^{lm} > SY^{abh}$ (std. dev.)	0.29 (.42)		0.56 (.82)	
	$S^{abh}=16$		$S^{abh} = S^{lm}=16$	
$ActX^{abh} - ActX^{lm}$ (rounded) is:	# obs	% obs	# obs	% obs
-2 or smaller	14	3.9		
-1	12	3.3		
0	302	83.0	3,244	89.0
1	24	6.6	225	6.2
2 or more	12	3.3	175	4.8
All	364	100.0	3,644	100.0
Mean ( $ActX^{abh} - ActX^{lm}$ )   $ActX^{abh} > ActX^{lm}$ (std. dev.)	0.33 (.79)		0.39 (.94)	
Mean ( $ActX^{abh} - ActX^{lm}$ )   $SY^{lm} > SY^{abh}$ (std. dev.)	0.64 (1.22)		0.66 (1.24)	

Note:  $ActX^{abh}$  and  $ActX^{lm}$  are actual experience defined for the ABH and LM samples, respectively. Sample sizes for the “ABH only” sample reflect the fact that we cannot define  $ActX^{abh}$  for all observations used by ABH.

<sup>a</sup>Mean difference between  $ActX^{abh}$  and  $ActX^{lm}$  (not rounded) conditional on being positive.

**Table 6:** Comparison of ABH and LM Schooling  
for “S=12/S=16, Potential Experience” Samples

	“ABH only”		“Both”	
	$S^{abh}=12$		$S^{abh} = S^{lm}=12$	
$S^{abh} < S^{lm}$	# obs	% obs	# obs	% obs
$S^{abh} = S^{lm}$	85	9.6	10,909	100.0
$S^{abh} > S^{lm}$	74	8.4		
	727	82.1		
All	886	100.0	10,909	100.0
Mean $(S^{abh} - S^{lm}) >0^a$	1.80			
(std. dev.)	(.94)			
$S^{abh} = S^{abh}$ at $SY^{abh}$	156	17.6	10,801	99.0
$S^{abh} > S^{abh}$ at $SY^{abh}$	730	82.4	108	1.0
All	886	100.0	10,909	100.0
Mean $(S^{abh} - S^{abh} \text{ at } SY^{abh})   >0^a$	1.84		1.48	
(std. dev.)	(.94)		(.68)	
	$S^{abh}=16$		$S^{abh} = S^{lm}=16$	
$S^{abh} < S^{lm}$	# obs	% obs	# obs	% obs
$S^{abh} = S^{lm}$	45	9.6	3,644	100.0
$S^{abh} > S^{lm}$	47	10.0		
	376	80.3		
All	468	100.0	3,644	100.0
Mean $(S^{abh} - S^{lm}) >0^a$	2.23			
(std. dev.)	(1.27)			
$S^{abh} = S^{abh}$ at $SY^{abh}$	175	37.4	3,349	91.9
$S^{abh} > S^{abh}$ at $SY^{abh}$	293	62.6	294	8.1
All	468	100.0	3,644	100.0
Mean $(S^{abh} - S^{abh} \text{ at } SY^{abh})   >0^a$	2.59		2.02	
(std. dev.)	(1.16)		(1.08)	

Note:  $S^{abh}$  and  $S^{lm}$  are highest grade completed defined for the ABH and LM samples, respectively;  $S^{abh}$  at  $SY^{abh}$  is the value of time-varying  $S^{abh}$  at the ABH career start year. Observations in the “ABH only” sample do not appear in the LM sample, but we are able to define  $S^{lm}$  for all observations used by ABH.

<sup>a</sup>Mean difference between  $S^{abh}$  and  $S^{lm}$  (or its value at  $SY^{abh}$ ) conditional on being positive.

**Table 7: Regression Estimates Using ABH and LM**  
 “ $S=12/S=16$ , Potential Experience” Samples

Variable	1	2	3	4	5	6	7	8
$S^{abh}=12$ or $S^{lm}=12$								
AFQT	.006 (.013)	.011 (.013)	.005 (.013)	.006 (.013)	.005 (.013)	.007 (.014)	.013 (.013)	.024 (.012)
AFQT·X/10 (a)	.126 (.018)	.121 (.018)	.127 (.018)	.121 (.018)	.123 (.018)	.123 (.019)	.114 (.018)	.104 (.015)
Black	-.063 (.027)	-.055 (.027)	-.068 (.027)	-.066 (.027)	-.063 (.027)	-.067 (.029)	-.070 (.026)	-.058 (.024)
Black·X/10	-.036 (.035)	-.044 (.035)	-.027 (.035)	-.032 (.035)	-.034 (.036)	-.026 (.037)	-.026 (.034)	-.045 (.031)
X	.092 (.013)	.090 (.013)	.094 (.013)	.091 (.013)	.092 (.013)	.096 (.014)	.089 (.011)	.090 (.010)
Root MSE	.394	.395	.391	.389	.390	.394	.390	.391
# observations	11,795	11,651	11,562	10,983	10,909	10,618	10,909	13,489
# individuals	1,926	1,862	1,856	1,675	1,672	1,673	1,672	2,074
$S^{abh}=16$ or $S^{lm}=16$								
AFQT	.148 (.035)	.136 (.036)	.153 (.036)	.144 (.037)	.140 (.038)	.120 (.041)	.120 (.035)	.109 (.033)
AFQT·X/10 (b)	.012 (.048)	.021 (.049)	.038 (.049)	.051 (.049)	.055 (.050)	.102 (.055)	.108 (.050)	.097 (.044)
Black	.110 (.056)	.075 (.057)	.117 (.056)	.104 (.058)	.087 (.057)	.105 (.067)	.090 (.053)	.082 (.052)
Black·X/10	-.130 (.069)	-.098 (.071)	-.102 (.065)	-.092 (.066)	-.075 (.067)	-.105 (.083)	-.064 (.067)	-.057 (.063)
X	.142 (.023)	.129 (.024)	.141 (.024)	.132 (.024)	.129 (.025)	.148 (.027)	.118 (.021)	.111 (.019)
Root MSE	.425	.424	.419	.420	.420	.407	.414	.415
# observations	4,112	4,022	3,823	3,691	3,644	2,750	3,644	3,941
# individuals	650	620	572	528	526	411	526	544
P-value for $H_0$ : (a)=(b)	.026	.053	.087	.185	.199	.728	.898	.879

Note: Columns 1-6 use ABH variables ( $X^{abh}$ ,  $S^{abh}$ ); columns 7-8 use LM variables ( $X^{lm}$ ,  $S^{lm}$ ). All regressions include  $X^2$ ,  $X^3$ , an urban indicator, and year fixed effects. Standard errors are corrected for nonindependence of observations reported over time by the same individual.

Column 1: ABH sample.

Column 2: ABH sample, but drop all observations that precede LM career start date ( $SD^{lm}$ ).

Column 3: ABH sample, but drop all observations that follow a reenrollment in school.

Column 4: ABH sample, but drop all observations where  $S^{abh} \neq S^{lm}$ .

Column 5: Observations common to ABH and LM samples; use ABH variables.

Column 6: ABH sample, but drop all observations where  $SY^{abh} < SY^{lm}$ .

Column 7: Observations common to ABH and LM samples; use LM variables.

Column 8: LM sample.

**Table 8:** Regression Estimates Using ABH and LM  
“S=12/S=16, Actual Experience” Samples

Variable	1	2	3	4	5	6	7	8
$S^{abh}=12$ or $S^{lm}=12$								
AFQT	.049 (.012)	.054 (.012)	.050 (.012)	.047 (.012)	.047 (.012)	.051 (.013)	.048 (.012)	.051 (.012)
AFQT·ActX/10 (a)	.073 (.021)	.067 (.021)	.070 (.021)	.074 (.021)	.074 (.021)	.068 (.022)	.073 (.021)	.074 (.020)
Black	-.008 (.024)	.000 (.025)	-.009 (.025)	-.006 (.025)	-.003 (.025)	-.003 (.026)	-.010 (.025)	-.004 (.024)
Black·ActX/10	-.067 (.044)	-.077 (.044)	-.068 (.044)	-.064 (.044)	-.067 (.044)	-.067 (.046)	-.057 (.045)	-.074 (.042)
ActX	.152 (.012)	.153 (.013)	.155 (.012)	.156 (.012)	.156 (.012)	.156 (.013)	.154 (.012)	.150 (.012)
Root MSE	.382	.383	.380	.377	.377	.383	.378	.379
# observations	9,866	9,722	9,664	9,147	9,062	8,746	9,062	10,050
# individuals	1,525	1,461	1,472	1,305	1,301	1,289	1,301	1,417
$S^{abh}=16$ or $S^{lm}=16$								
AFQT	.150 (.033)	.134 (.034)	.167 (.034)	.153 (.035)	.151 (.036)	.134 (.040)	.141 (.035)	.129 (.033)
AFQT·ActX/10 (b)	.042 (.047)	.060 (.048)	.040 (.049)	.063 (.050)	.064 (.051)	.105 (.055)	.082 (.052)	.070 (.046)
Black	.092 (.054)	.055 (.056)	.133 (.053)	.117 (.056)	.105 (.057)	.113 (.067)	.117 (.055)	.111 (.054)
Black·ActX/10	-.063 (.065)	-.020 (.069)	-.104 (.066)	-.086 (.067)	-.074 (.070)	-.094 (.078)	-.093 (.070)	-.089 (.066)
ActX	.147 (.021)	.130 (.022)	.151 (.021)	.139 (.022)	.134 (.022)	.148 (.025)	.130 (.021)	.131 (.019)
Root MSE	.414	.414	.409	.410	.411	.400	.409	.410
# observations	4,098	4,008	3,823	3,691	3,644	2,736	3,644	3,941
# individuals	645	615	572	528	526	406	526	544
<sup>a</sup> P-value for H <sub>0</sub> : (a)=(b)	.544	.888	.573	.845	.849	.536	.876	.937

Note: Columns 1-6 use ABH variables ( $ActX^{abh}$ ,  $S^{abh}$ ); columns 7-8 use LM variables ( $ActX^{lm}$ ,  $S^{lm}$ ).

All regressions include  $ActX^2$ ,  $ActX^3$ , an urban indicator, and year fixed effects. Standard errors are corrected for nonindependence of observations reported over time by the same individual.

Column 1: ABH sample.

Column 2: ABH sample, but drop all observations that precede LM career start date ( $SD^{lm}$ ).

Column 3: ABH sample, but drop all observations that follow a reenrollment in school.

Column 4: ABH sample, but drop all observations where  $S^{abh} \neq S^{lm}$ .

Column 5: Observations common to ABH and LM samples; use ABH variables.

Column 6: ABH sample, but drop all observations where  $SY^{abh} < SY^{lm}$ .

Column 7: Observations common to ABH and LM samples; use LM variables.

Column 8: LM sample.

Table 9: Regression Estimates Using ABH and LM “All  $S$  Levels” Samples

Variable	Potential $X$		Actual $X$	
	ABH	LM	ABH	LM
$S$	.083 (.005)	.080 (.005)	.069 (.004)	.073 (.004)
$S \cdot X$	-.026 (.007)	-.005 (.006)	-.001 (.007)	.001 (.007)
AFQT	.033 (.011)	.048 (.009)	.055 (.010)	.058 (.009)
AFQT $\cdot X/10$	.094 (.014)	.077 (.012)	.076 (.016)	.071 (.015)
Black	-.018 (.021)	-.027 (.018)	-.015 (.018)	-.013 (.017)
Black $\cdot X/10$	-.074 (.025)	-.064 (.022)	-.041 (.030)	-.059 (.028)
$X$	.144 (.013)	.114 (.011)	.148 (.013)	.134 (.012)
$X^2/10$	-.091 (.015)	-.087 (.012)	-.135 (.016)	-.123 (.014)
$X^3/100$	.029 (.007)	.030 (.006)	.057 (.009)	.050 (.008)
Root MSE	.420	.405	.408	.399
# observations	25,692	28,300	22,975	22,731
# individuals	3,674	4,399	3,154	3,351

Note: The ABH (LM) potential experience specifications use variables  $X^{abh}$  and  $S^{abh}$  ( $X^{lm}$  and  $S^{lm}$ ). The ABH (LM) actual experience specifications use variables  $ActX^{abh}$  and  $S^{abh}$  ( $ActX^{lm}$  and  $S^{lm}$ ). All specifications include an urban indicator and year fixed effects. Standard errors are corrected for non-independence of observations reported over time by the same individual.