

# The Relationship Between Wage Growth and Wage Levels

Tricia Gladden<sup>1</sup>  
Bureau of Labor Statistics

Christopher Taber  
(corresponding author)  
Department of Economics  
University of Wisconsin-Madison  
1180 Observatory Drive  
Madison, WI 53706, USA  
ctaber@ssc.wisc.edu  
(608) 263-7791  
fax: (608)262-2033

May 19, 2008

<sup>1</sup>We would like to thank Joe Altonji, Harry Holzer, Elizabeth Lower-Basch, Tom MaCurdy, Bruce Meyer, Robert Moffitt, Derek Neal, John Rust, Geoffrey Wallace, two anonymous referees, and seminar participants at Syracuse University, the Missouri Economic Conference and the Midwest Economic Association for helpful comments. We gratefully acknowledge the financial support of the small grants program sponsored by the Joint Center for Poverty Research. All remaining errors are our own.

## **Abstract**

We estimate the covariance between the permanent component of wages and a random coefficient on experience in models both with potential experience and with actual experience. Actual experience is allowed to be arbitrarily correlated with both the permanent component of wages and the random component on experience. We find no evidence that workers of higher ability experience faster wage growth. Our point estimates suggest that a worker with a one standard deviation higher level of permanent ability would have a return to annual potential experience that is 0.61 of a percentage point lower. The analogous point estimate for actual experience is 0.87 of a point lower. Contrary to the popular perception, wage growth among low skill workers appears to be at least as high as that for a medium skilled worker. Key words: Panel Data, Wage Growth, Low Skilled Labor Markets, Random Coefficients

# 1 Introduction

Increasing labor force participation among low skill workers has been a major goal of policy makers over the last two decades. Examples of policies that share this goal include expansions in the Earned Income Tax Credit and the Personal Responsibility and Work Opportunity Reconciliation Act (welfare reform). An important consequence of these policies has been additional work experience for low skilled workers. One of the most robust findings in labor economics is that wages increase with work experience. However, very little of this work has focused on wage growth among low wage workers. As a result, we have little information on the impact of additional labor force participation on future wages for this group. One possible explanation for this gap in the literature is that there are serious econometric issues behind the wage growth process involving parameter heterogeneity, endogeneity, and selection issues. We attempt to address these issues in this paper. A second explanation for this lack of research is that policy makers may believe that low wage workers are locked into “dead end jobs” in which wages stagnate. We find no evidence to support this claim. Our results suggest that wage growth among low skill workers is statistically indistinguishable from wage growth among medium skilled workers. In fact, our point estimates indicate that wages may grow slightly faster for low skill workers than medium skilled workers.

Since the goal of this work is to study the early labor market experience

of low skill workers, we limit our sample in two important ways. First, we study only workers who have completed less than a year of college. Second, we examine only the first ten years after a worker enters the labor force.

We use a correlated random effects model to estimate the relationship between wage levels and wage growth. That is we let wages take the form:

$$(1) \quad w_{it} = \theta_i + \beta_i E_{it} + u_{it},$$

where  $w_{it}$  represents log wages,  $\theta_i$  and  $\beta_i$  are individual specific random coefficients,  $E_{it}$  is a measure of experience, and  $u_{it}$  is an error term. The random effect,  $\theta_i$ , represents the unobserved skill of a worker. Our goal is to estimate the relationship between the permanent component,  $\theta_i$ , and the random coefficient on experience,  $\beta_i$ . In an ideal world we would nonparametrically estimate the full joint distribution of  $(\theta_i, \beta_i)$ . We could then simulate the expected level of wage growth for different levels of unobserved skill (i.e.  $E[\beta_i | \theta_i]$ ). Unfortunately the data is not rich enough to estimate this conditional expectation precisely. Instead we focus on the simplest summary statistic of the relationship between  $\theta_i$  and  $\beta_i$ , its covariance. We believe that with richer data our methodology could be extended to uncover this full joint distribution.

We estimate models using both potential experience, defined as years since leaving school, and actual experience, defined as the total number of weeks worked since leaving school. In our primary specification, we allow actual experience,  $E_{it}$ , to be arbitrarily correlated with both  $\theta_i$  and  $\beta_i$  (but

uncorrelated with  $u_{it}$ ). Our point estimates of this relationship are negative and statistically insignificant. They suggest that a one standard deviation increase in permanent ability reduces the return to annual potential experience by 0.61 of a percentage point. The analogous point estimate for actual experience is a reduction of 0.87 of a percentage point. Contrary to the popular perception, wage growth among low skill workers appears to be at least as high as that for a medium skilled worker.<sup>1</sup>

We know of at least three reasons why these empirical results are interesting and important. First, as mentioned above, they are informative for policy makers interested in the effects of programs that promote work experience on wages. Our work represents basic research that is not intended to answer a specific policy question. However, our results suggest that using estimates from the large literature on returns to experience may give reasonable estimates for the returns for low wage workers. Second, these results are informative for labor economists interested in understanding the wage process. The parameter  $\theta_i$  is often called “ability” while  $\beta_i$  is called “ability to learn.” Under this interpretation one might strongly suspect that these two abilities are positively correlated. Our results indicate that they are not, a result that is surprising and of relevance to labor economists interested in

---

<sup>1</sup>It is important to point out the difference between this result and the Mincer “overtaking” result. Mincer’s story is one about age: workers with higher levels of schooling leave school later and thus experience their wage growth at a later age than do less educated workers. Our story is different because we are measuring wage growth across experience rather than across age where one would not expect this result. In fact the standard Mincer model predicts experience profiles that are parallel in experience.

understanding the determinants of wages. Third, models of the wage process are often important inputs into simulated models. For example, public economists often want to simulate distribution effects of the social security system (e.g. Gokhale and Kotlikoff, 2002) or federal income taxes (e.g. Fullerton and Rogers, 1993). Another example is given by financial economists who often want to model the amount of wage risk that workers face (e.g. Heaton and Lucas, 1996). These models are often based on wage models such as (1). While  $cov(\theta_i, \beta_i)$  is only one piece of this process it may be an important piece. Our results present a range of reasonable values that can be assumed for this component.

Our methodological contribution is to show how to use panel data to estimate  $cov(\theta_i, \beta_i)$  in model (1) where  $E_{it}$  is endogenous and potentially correlated with  $(\theta_i, \beta_i)$ . This should be useful both for future estimation of wage growth models and for estimation of other random coefficient panel data models.

The empirical approach taken in this paper is somewhat different from much of the literature on estimation of the wage process. Our goal is not to uncover a single set of parameters for the “best” model of wages. Instead, our goal is to improve understanding of the relationship between unobserved skill and wage growth—a topic on which there has been almost no previous work. Rather than trying to discover a single best set of estimates, we estimate many different models that use different assumptions in different ways. We show that the results are robust across the different specifications. The paper

proceeds as follows: section 2 reviews the related literature and section 3 describes our econometric approach. In section 4 we discuss the National Longitudinal Survey of Youth (NLSY79) data that we use. Section 5 presents a regression approach and Section 6 presents a GMM approach. We provide conclusions in Section 7.

## 2 Previous Work

This paper builds on our previous work (Gladden and Taber, 2000) in which we measure the variation in wage growth across members of different observable skill classes. As an example of the type of analysis in that paper, suppose one wants to compare wage growth between two groups of individuals. Let  $g_i$  be a dummy variable that distinguishes between the groups. Our primary comparison is between individuals who obtained a high school diploma versus those that did not.<sup>2</sup> We report results from the regression:

$$w_{it} = \beta_0 + \beta_1 g_i + \beta_2 AE_{it} + \beta_3 AE_{it} g_i + u_{it},$$

where  $AE_{it}$  represents actual work experience. The key parameter of interest in this work is  $\beta_3$ : the difference in returns to experience between high school graduates and workers who do not complete high school. The problem in estimating this parameter is that, if higher wage workers participate in the labor force at a higher rate, actual experience may be positively correlated with  $u_{it}$ . We deal with this possibility in a number of different ways,

---

<sup>2</sup>Our sample included no individuals who had completed a year of post-secondary education.

including instrumental variables and fixed effects. We consistently estimate  $\beta_3$  to be small in magnitude and insignificant, indicating that the return to work experience for dropouts is not significantly different than the return to experience for high school graduates. We also look for differences in earnings growth between gender, racial, and socioeconomic groups. We find no evidence that family income or other measures of family background affect wage growth. We do find that race and gender relate to wage growth in the manner found by previous research. White men tend to have larger returns to experience than white women or black men. Interestingly, in some specifications we find that black women actually receive higher returns to experience than white women or black men.

Several authors have taken a direct approach by studying the relationship between wage growth and welfare receipt. Using data from the Panel Study of Income Dynamics, Moffitt and Rangarajan(1989) find that mothers who are typical welfare recipients have steeper wage growth than typical non-recipients, but warn of selection bias. Burtless (1995) looks at the return to potential experience and finds that wages grow more slowly for welfare mothers than others. Corcoran and Loeb (1999) examine actual experience and find that welfare recipients in the NLSY have slightly lower wage growth than other workers. Connolly and Gottschalk (2000) worry about job matching and find evidence that wages grow faster for more educated workers. French, Mazumder, and Taber (2006) look at changes in the return to experience for low skilled across time.

The main weakness of previous work (including our own) is that it only examines the relationship between wage progression and observable measures of skill. Observables explain only a small amount of wage dispersion. While the observable measures of skill are statistically significant in explaining wage growth, the  $R^2$  in our regressions are approximately 0.15. This leaves a tremendous amount of wage variation that cannot be explained by observable measures of skill. Our goal in this work is to assess the relationship between unobservable skill and wage growth, or more specifically the relationship between wage growth and wage levels. We build on the substantial literature on the covariance structure of wages in labor economics including Abowd and Card (1989) and more recent work by Baker (1997), Baker and Solon (2003), and Lillard and Reville (1999). This literature attempts to understand the evolution of wages over the life-cycle.

Although it is not the focus of their studies, Baker (1997) and Lillard and Reville (1999) both find a negative relationship between wage levels and wage growth. This work differs from ours along many dimensions. First, both authors use the PSID and do not focus on earnings growth early in the life-cycle. We show that this early part of the life-cycle is crucial for identification in our framework. Second, they do not focus on low skill workers. Third, they follow the larger literature on earnings dynamics that has largely ignored the information that can be provided by actual experience as opposed to potential experience. Our previous work indicates that this may be a major limitation as the results differ substantially when actual experience is used.

On the methodological side, there exists a huge literature that allows for random coefficients. In many cases the random coefficient is allowed to be correlated with the coefficient on which it multiplies. For example the “treatment effect” literature (e.g. Heckman and Robb, 1987, Imbens and Angrist, 1994, or Heckman, 1997) models the treatment effect as a random coefficient on the endogenous treatment. Wooldridge (1997) and Heckman and Vytlacil (1998) consider a more general version of this model. We differ from this literature by focusing on a summary statistic about the joint distribution (the covariance) of two of these parameters as opposed to a summary statistic of the treatment effect (like the “average treatment effect”). We also take advantage of panel data. Other papers such as Baker (1997) or Barry, Levinsohn, and Pakes (1995) allow for correlation of random coefficients as one part of a larger model. However, they do not focus explicitly on the estimation of this relationship making the method of estimation and the source of identification not transparent.

### 3 Econometric Approach

In the context of the model, our goal is to discover the relationship between  $\theta_i$  and  $\beta_i$ . We will consider two measures of experience: *potential experience* defined as time since entering the labor force and *actual experience* defined as the actual amount of time that the individual has worked since leaving school. We first consider the simpler model in which experience is defined as

potential experience ( $PE_{it}$ ) so that log wages are defined as

$$(2) \quad w_{it} = \theta_i + \beta_i PE_{it} + u_{it},$$

where potential experience increases by one each year.<sup>3</sup> We put no restriction on the relationship between  $\theta_i$  and  $\beta_i$  but assume that  $u_{it}$  is uncorrelated with both of them. A simple method of examining this relationship is to look at the covariance between wage differences and initial wage levels

$$\begin{aligned} cov(w_{iT} - w_{iT-\tau}, w_{i0}) &= cov(\beta_i \tau + u_{iT} - u_{iT-\tau}, \theta_i + u_{i0}) \\ &= \tau cov(\theta_i, \beta_i) + cov(u_{iT} - u_{iT-\tau}, u_{i0}). \end{aligned}$$

Under standard time series properties, if  $T-\tau$  is large enough,  $cov(u_{iT} - u_{iT-\tau}, u_{i0})$  should be close to zero. Let  $\bar{w}_0$  be the sample mean of  $w_{i0}$ . As long as  $u_{iT}$  is asymptotically uncorrelated for each  $i$ ,

$$\frac{1}{N} \sum_{i=1}^N \frac{(w_{iT} - w_{iT-\tau})}{\tau} (w_{i0} - \bar{w}_0) \xrightarrow{p} cov(\theta_i, \beta_i)$$

as  $N \rightarrow \infty$  and  $T \rightarrow \infty$ . Our strategy for estimating  $cov(\theta_i, \beta_i)$  is based on this idea.

An alternative interpretation of our approach is an m-dependence type assumption in which one assumes that  $u_{iT-\tau}$  is uncorrelated with  $u_{i0}$ . That is we could assume that

$$(3) \quad cov(u_{iT}, u_{iT-h}) = 0 \text{ when } h > h^*.$$

---

<sup>3</sup>Previous works suggests that the wage profile is approximately linear during the first ten years. To simplify the analysis we do not include a quadratic term in the wage equation at this point. We test this assumption below.

Formally this is weaker than m-dependence as we only need to assume mean independence rather than full independence. However, for expositional purposes we will continue to refer to it as m-dependence. For example an MA(4) satisfies this restriction with  $h^* = 4$ . As long as  $T - \tau > h^*$ ,

$$\frac{1}{N} \sum_{i=1}^N \frac{(w_{iT} - w_{iT-\tau})}{\tau} (w_{i0} - \bar{w}_0) \xrightarrow{p} cov(\theta_i, \beta_i)$$

as  $N \rightarrow \infty$ .

The key issue is whether  $T$  is large enough in practice for either of these approximations to be valid. In our specifications,  $T$  is usually a number between 6 and 9. As a rough exercise to test this approximation, consider the estimates presented in Baker (1997) Table 4, row 5. Baker assumes a random coefficient model of wages similar to ours, and assumes that the errors follow an ARMA(1,2) process. In the ARMA(1,2) model with  $T = 8$

$$cov(u_{i0}, u_{i8} - u_{i7}) = (\rho - 1)(\rho^2 + \rho\mu_1 + \mu_2)\rho^6\sigma_\varepsilon^2$$

where  $\rho$  is the autocorrelation parameter,  $\mu_1$  and  $\mu_2$  are the moving average parameters, and  $\sigma_\varepsilon^2$  is the variance of the white noise component. Notice that if  $\rho$  is close to one or close to zero this covariance will be small. With Baker's estimates:  $\rho = 0.519$ ,  $\mu_1 = -0.112$ ,  $\mu_2 = -0.040$  and  $\sigma_\varepsilon^2 = 0.092$ , the numerical value of  $cov(u_{i0}, u_{i8} - u_{i7})$  would be -0.00015 which is tiny relative to Baker's estimate of  $cov(\theta_i, \beta_i)$  which is -0.013. Thus it appears that this approximation should work well in practice.

It is not essential for our approach to assume that wages are linear in

experience. Generalize the model to

$$(4) \quad w_{it} = \theta_i + b_i(PE_{it}) + u_{it},$$

where  $b_i$  is a nonlinear function and  $b_i(0) = 0$ . Assume further  $u_{it}$  is  $m$ -dependent so that  $cov(u_{iT}, u_{iT-h}) = 0$  when  $h > h^*$ , then for all  $T - 1 > h^*$

$$cov(w_{i0}, w_{iT} - w_{iT-1}) = cov(\theta_i, b_i(T) - b_i(T - 1))$$

which is the relationship between unobserved ability and wage growth during this period.

Without making stronger assumptions on  $b_i$ , using data on initial wages is necessary. The key aspect of initial wages is that experience is 0 so the return to experience at that point is irrelevant. To see this suppose that the first period of data we observe was  $T = 1$ . Without parametric assumptions on  $b_i$  we could never separate the model (4) from an alternative model

$$w_{it} = \theta_i^* + b_i^*(PE_{it}) + u_{it},$$

with

$$\begin{aligned} \theta_i^* &= \theta_i + b_i(1) \\ b_i^*(PE_{it}) &= b_i(PE_{it}) - b_i(1). \end{aligned}$$

Clearly  $cov(\theta_i, b_i(T) - b_i(T - \tau)) \neq cov(\theta_i^*, b_i^*(T) - b_i^*(T - 1))$  in general. Thus in the absence of parameteric restrictions on  $b_i$ , having initial wages would be essential for identification.

In addition, without making strong assumptions about  $u_{it}$ , one can not identify  $cov(w_{i0}, b_i(1))$ . In principal one could use later periods in life to estimate the covariance structure of  $u_{it}$ . However, the early labor force experience of low wage workers is notoriously unstable with job changes common. It does not seem attractive to assume that the covariance structure for a 40 year old high school graduate is informative about the covariance structure for an 18 year old high school graduate.

Our second specification incorporates actual experience into the model as

$$w_{it} = \theta_i + \beta_i AE_{it} + u_{it},$$

where  $AE_{it}$  is measured as total weeks of experience (divided by 52) at the beginning of the current calendar year. A potential problem with using actual experience is that we would expect it to be correlated with wage levels and wage growth. We will not put any restriction on this relationship. In order to obtain estimates we assume that  $u_{it}$  is independent of  $(\theta_i, \beta_i, AE_{it})$ . In this case, the natural extension of the previous model does not work. Even if  $cov(u_{iT} - u_{iT-\tau}, u_{i0}) = 0$ ,

$$\begin{aligned} cov(w_{i0}, w_{iT} - w_{iT-\tau}) &= cov(\theta_i, \beta_i (AE_{iT} - AE_{iT-\tau})) \\ &\neq cov(\theta_i, \beta_i) (AE_{iT} - AE_{iT-\tau}) \end{aligned}$$

in general. However, using a simple transformation of the data we can extend

the model to the case of actual experience. Note that:

$$\begin{aligned}
 (5) \quad cov \left( w_{i0}, \frac{w_{iT} - w_{iT-\tau}}{AE_{iT} - AE_{iT-\tau}} \right) &= cov \left( \beta_i + \frac{u_{iT} - u_{iT-\tau}}{AE_{iT} - AE_{iT-\tau}}, \theta_i + u_{i0} \right) \\
 &= cov(\theta_i, \beta_i) + cov \left( u_{i0}, \frac{u_{iT} - u_{iT-\tau}}{AE_{iT} - AE_{iT-\tau}} \right).
 \end{aligned}$$

Once again if  $u_{iT}$  is asymptotically uncorrelated we can obtain a consistent estimate of  $cov(\theta_i, \beta_i)$  as  $T$  gets large or by using m-dependence.<sup>4</sup>

One nice aspect of this approach is that we are able to avoid making any assumption at all about the relationship between  $(\theta_i, \beta_i)$  and  $AE_{it}$ . In general it would be very complicated to model the relationship between the two. Second, our method makes the relationship between the data and the estimates clear. A small estimate of  $cov(\theta_i, \beta_i)$  simply means that there is little relationship between the workers' initial wages and their wage growth later in the lifecycle.

However, an important sample selection problem arises when we implement a GMM estimator based on (5). It requires that we observe wage data for three specific periods:  $0, T - \tau$ , and  $T$ . However, if some individuals in the data do not work in one of those periods then there may be selection bias in our estimate of  $Cov(\theta_i, \beta_i)$ . A major advantage of our approach is that it allows us to address this problem. We do not have data on earnings in

---

<sup>4</sup>One potential problem is that  $\theta_i$  may be correlated with  $(u_{it} - u_{it-\tau}) / (AE_{it} - AE_{it-\tau})$  due to sample selection bias. While it is impossible to know for sure, we have tried a number of robustness checks to see if this can be driving the results (see Gladden and Taber, 2004, for another specification which should be robust to this problem). We also argue below that this result is likely to lead to positive biases which make our results even more surprising. We have found no evidence that this problem is empirically important.

every year for every sample member, but since the parameters  $\beta_i$  and  $\theta_i$  do not vary over time we can estimate them using the data we have. As long as almost all individuals work at least three years at some point in the panel,  $cov(\theta_i, \beta_i)$  is identified.

To see this consider the model in which  $u_{it}$  is serially uncorrelated. In this case, calculating a consistent estimate of

$$cov\left(w_{i0}, \frac{w_{i2} - w_{i1}}{AE_{i2} - AE_{i1}}\right)$$

yields a consistent estimate of  $cov(\theta_i, \beta_i)$ . However, the sample selection bias may be severe in this case. We can only construct this covariance for those individuals in the sample who work in years zero, one, and two. In a sample of low wage women, there are likely to be a considerable number of sample members who do not work in at least one of these three years. However, this problem can be easily solved. To estimate  $cov(\theta_i, \beta_i)$  in this case, we can use the first three years that the woman actually works. Let  $t_{1i}$  be the first year the woman works,  $t_{2i}$  the second, and  $t_{3i}$  the third. By definition actual experience is zero at  $t_{1i}$ . In this case

$$cov(\theta_i, \beta_i) \approx cov\left(w_{it_{1i}}, \frac{w_{it_{3i}} - w_{it_{2i}}}{AE_{it_{3i}} - AE_{it_{2i}}}\right).$$

We can estimate  $cov(\theta_i, \beta_i)$  using the sample analog of the right hand side of this equation.<sup>5</sup> If every member of the sample worked for at least three

---

<sup>5</sup>It is important to note that this solves the sample selection problem relating to labor supply. There may be an additional problem of missing data if some individuals are not surveyed in all years. The solution for that problem is not as clean as we need to impute actual experience for them.

different years during the panel (with  $t_{2i}$  sufficiently greater than  $t_{1i}$ ), the sample selection problem would disappear.

While using different years allows us to address the sample selection problem, it leads to an additional problem. Few of the observations in our sample provide all 10 wages. This occurs for 4 reasons: some individuals started working before 1979, the year of the first NLSY interview. After 1994, data on wages is collected only every other year. Many individuals, especially women, have some years in which they do not work. Finally, some individuals drop out of the sample. As a result  $(t_{1i}, t_{2i}, t_{3i})$  varies across different members of the sample.

In the spirit of the discussion above, if we construct moments that are weighted sums of several covariances, we are able to minimize this selection problem. We condition on individuals for whom we observe  $w_{i0}$ , the wage during the first year after they left school. Define a dummy variable that takes the value of 1 if we observe wages for individual  $i$  in periods  $T$  and  $T - \tau$ , and zero otherwise. That is let

$$d_{iT\tau} = \begin{cases} 1 & \text{if we observe } w_{iT}, w_{iT-\tau}, \text{ and } (AE_{iT} - AE_{iT-\tau}) \\ 0 & \text{otherwise.} \end{cases}$$

We consider estimating  $cov\left(w_{i0}, \frac{w_{iT} - w_{iT-\tau}}{AE_{iT} - AE_{iT-\tau}}\right)$  for various values of  $T$  and  $\tau$ . Let  $R$  be the set of  $(T, \tau)$  that we consider. Define a variable  $m_i$  as

$$(6) \quad m_i \equiv \frac{\sum_{(T,\tau) \in R} d_{iT\tau} (w_{i0} - \bar{w}_0) \left( \frac{w_{iT} - w_{iT-\tau}}{AE_{iT} - AE_{iT-\tau}} \right)}{\sum_{(T,\tau) \in R} d_{iT\tau}}.$$

Since  $E(m_i) = cov(\theta_i, \beta_i)$ , we can then estimate the covariance as

$$(7) \quad cov(\widehat{\theta}_i, \widehat{\beta}_i) = \frac{1}{N} \sum_{i=1}^N m_i.$$

In this calculation, each individual receives equal weight regardless of the number of observed wages that we have for them. Since we expect the number of observed wages to depend on  $(\theta_i, \beta_i)$  this yields a consistent estimate of  $cov(\theta_i, \beta_i)$  while an estimate obtained by putting equal weights on each observed wage typically would not.<sup>6</sup>

Unfortunately, the method above does not completely solve the problem of sample selection. In our NLSY sample, there are 6491 individuals with 12 or fewer years of school. About 4350 of these individuals leave school after 1977 so that we can observe one of their first two wages, and continue in the sample 6 or more years after leaving school.<sup>7</sup> If we assume that there is no cohort effect, and that attrition from the sample does not cause sample selection, we could obtain unbiased estimates using these 4350 individuals. However, only about 3250 of these individuals have non-missing wage observations for one of the first two years out of school, and non-missing observations for two or more wages in periods 6-9. Thus, we still have a selected sample. The basic problem is that we are estimating the effect of work experience on wages, and this is impossible to measure for people who do not work. Thus we estimate the covariance between  $\theta_i$  and  $\beta_i$  conditional

---

<sup>6</sup>Clearly one can not simultaneously deal with the sample selection problem and non-parametric  $b_i$ . We must ignore one problem in dealing with the other.

<sup>7</sup>About 150 observations are lost due to attrition from the sample. The remainder is lost because the older members of the sample leave school before 1977.

on individuals who have non-missing wage observations for one of the first two years out of school and non-missing wage observations for two or more wages in periods 6-9.

There is an additional potential source of bias. The fact that we only observe wages for individuals who work likely causes an upward bias in the  $u_{it}$  for our observed sample. Focusing on potential experience, we can obtain the equation

$$\begin{aligned} & cov(w_{i0}, (w_{iT} - w_{iT-\tau}) \mid d_{iT\tau} = 1) \\ & = cov(\theta_i, \beta_i \mid d_{iT\tau} = 1) + cov(\theta_i, u_{iT} - u_{iT-\tau} \mid d_{iT\tau} = 1) \\ & \quad + cov(u_{i0}, \beta_i \mid d_{iT\tau} = 1) + cov(u_{i0}, u_{iT} - u_{iT-\tau} \mid d_{iT\tau} = 1). \end{aligned}$$

The bulk of this paper focuses on the first and fourth terms on the right hand side of this equation. However, even if the variables  $\theta_i$  and  $\beta_i$  are uncorrelated with  $u_{iT}$  unconditionally, they may be correlated with  $u_{iT}$  conditional on working. We generally expect that workers who receive large positive wage shocks will be more likely to work, and thus will be more likely to have observed wages. This means that the expected value of  $u_{iT}$  conditional on working will tend to be positive, and that the second and third terms in the equation above may not be zero, causing our estimates to be biased.

First consider the second term in this expression. In a classic Heckman style selection model, if labor supply were the same in periods  $T$  and  $T - \tau$  this selection bias would bias both  $u_{iT}$  and  $u_{iT-\tau}$  upward by the same amount and not affect the covariance. However, both labor supply and wages increase

with age so  $u_{iT-\tau}$  is likely to be biased *upward* by more than  $u_{iT}$ . As a result, conditioning on working both periods tends to bias  $u_{iT} - u_{iT-\tau}$  *downward*. Since low wage workers tend to work less this bias will be relatively larger for low wage workers than high wage workers so it will tend to *positively* bias our estimates of  $cov(\theta_i, \beta_i)$ . This makes our results more surprising.

The third term could go the other direction. We have very little evidence on the relationship between working in the first period and  $\beta_i$  but one would generally expect it to be positive.<sup>8</sup> Thus the selection bias on  $u_{i0}$  will tend to be larger for individuals with smaller values of  $\beta_i$  this could lead to a negative bias. However, we will show (in Table 3) that if anything the estimates of  $cov(\theta_i, \beta_i)$  are more negative for men than for women, making it hard to believe that this could be driving our results.

## 4 The Data

We use data from the National Longitudinal Study of Youth 1979 (NLSY79). The NLSY79 is a panel data set begun in 1979 with youth aged 14 to 22. We use the cross-sectional sample as well as the oversamples of blacks and Hispanics. The survey is conducted annually until 1994 and bi-annually since then, and respondents are questioned on a large range of topics, including schooling, wages, and work experience.

Our goal is to focus on low to moderate skilled workers, so we use the

---

<sup>8</sup>Economic theory has no prediction as the income and substitution effects go in different directions.

subsample of individuals with 12 or fewer completed years of schooling. In order to focus on the early part of the career so we only use wage information collected during the first 10 years after an individual has left school. One advantage of the NLSY is that individuals report the number of weeks worked for each year in the sample. This information is also obtained retrospectively for the years preceding the survey. This allows us to construct a measure of actual experience that is the key variable in our analysis. We calculate labor market experience in the following manner. An individual is assumed to enter the labor force at the time of the interview immediately following the last year that he was enrolled in school. At the time of this interview, we assume he has no experience, and experience accumulates each year by the number of weeks worked. We impute experience for missing years by averaging the number of weeks worked in the year immediately preceding and in the year immediately following the missing year.

One potentially difficult issue is precisely defining the time of entry into the labor force. Ideally, entry would begin on the date that an individual leaves school and enters the labor force. Due to data limitations, we assume that an individual begins his working life with zero experience at the time of the first interview after they leave school. The issue of individuals returning to school does not appear to be problematic in our sample. We do not include anyone who completes a year of post-secondary education. While a substantial number of high school graduates return to college after working in the labor force for some time, these people are not included in our data.

Second, individuals who drop out of school and later receive a General Equivalency Degree (GED) are treated as dropouts. This assumption is justified by Cameron and Heckman (1993) who show that individuals with GEDs have earnings that are closer to dropouts than to high school graduates. The few students who drop out, complete a GED, and then attend college are not included in the sample. Thus, the only group of students who will be problematic are those who drop out of high school and return to conventional high school to complete a grade or get a standard high school diploma, but do not move on to college. Very few individuals report this pattern of schooling: only about 7% of high school non-completers and 1% of eventual high school graduates leave school for over a year and then return.

## 5 Regression Results

We begin with a regression exercise as a preliminary approach. The key relationship studied in this paper is  $cov(w_{i0}, w_{iT} - w_{iT-1})$ . Under a linear random effects model, this can be interpreted as the covariance between  $\theta_i$  and  $\beta_i$ . In the potential experience model if we regress  $(w_{iT} - w_{iT-1})$  on  $w_{i0}$ , the slope coefficient will converge to

$$\frac{cov(\theta_i, \beta_i)}{var(w_{i0})}.$$

Thus OLS estimates a normalized version of the covariance that is helpful because the regression framework provides a nice scale for judging the magnitude of the covariance.

Table 1 presents the results from a regression of wage growth per year of potential experience ( $w_{i,T} - w_{i,T-1}$ ) on initial wages. To avoid correlation between the error terms, we construct wage differences using wages from  $T=7, 8, 9,$  and  $10$ . These results are presented in the first four columns of Table 1. Columns (1) and (3) present unweighted results; columns (2) and (4) present results that are weighted so that each individual receives equal importance regardless of the number of wages observed. The point estimate of the relationship between wage levels and wage growth is actually negative, but insignificant at conventional levels. The magnitude of the effect is small and fairly precisely estimated. Workers who earn 0.5 log wage points less than the median experience approximately 0.005 (one half of a percentage point) more wage growth per year than the median worker. Since a 0.5 log wage change seems large to us, the change of 0.005 relative to the level of 0.05 seems small, particularly relative to the differences across racial groups.

Since there may be reasons to not trust the first wage after entering the labor market, as a robustness check we run the same regression for the wage in the second year as well. These results are presented columns (5)-(8) of Table 1. The results are similar, although the magnitudes of the point estimates are slightly bigger and the standard errors are somewhat smaller in this case leading to statistically significant results. It is straight forward to modify the framework to formally justify the use of the second wage simply by allowing the wage process described above to initialize at that point. The implicit assumption would be that in the first year after leaving school,

workers aren't serious about their job prospects and don't buckle down until the second year so that experience gained during the first year is useless. The main lesson from this exercise is that the basic results are robust.

We next account for actual experience by using the analogue of the moment from equation (5). Specifically, we regress  $\frac{w_{i,T}-w_{i,T-1}}{AE_{i,T}-AE_{i,T-1}}$  on initial wages and wages one year out. These results are presented in Table 2. Again, we present both weighted and unweighted results. The results are similar to those in Table 1: fairly small negative effects that are borderline significant. The point estimates suggest that workers who earn 0.5 log wage points less than the median worker experience wage growth that is about one percentage point higher. This is nontrivial, but we interpret it as a fairly small effect.

There is some concern that the  $cov(\theta_i, \beta_i)$  may be different for men and women. To address this, Table 3 presents results of a regression of  $\frac{w_{i,T}-w_{i,T-1}}{AE_{i,T}-AE_{i,T-1}}$  on initial wages and wages one year out stratified by gender. In this case the point estimates are closer to zero for women than for men. However, the differences are not statistically significant. Given that much precision is lost when we separate women from men, we continue to use the pooled sample. We should note that although we do not report all of the results, we tried many specifications comparing women to men. Across specifications  $cov(\theta_i, \beta_i)$  was consistently a larger negative number for men than for women, but the difference was not typically significant.

## 6 Generalized Method of Moments Approach

We now move to the main approach that focuses on estimation using equation (7). A major advantage of this approach relative to the regression approach in the previous section is that we can address the sample selection problem more formally by using data from anyone with at least 3 wages. In particular, we do not have to assume that all early wages are uncorrelated with later wage gains in order to obtain a consistent estimate.

Recall that in section 3 we defined  $R$  as the set of  $(T, \tau)$  that we consider. For reasons we will discuss below we focus on the following set:

$$R = \{T, \tau : T \leq 9, \tau \in \{1, 2, 3\}, T - \tau \geq 5\}.$$

We refer to the initial wage as 0, so that  $t = 9$  is actually the tenth wage potentially observed after the individual entered the labor force. Since we use no wage before period 5, when  $\tau = 1$  we use  $w_6 - w_5, w_7 - w_6, w_8 - w_7$ , and  $w_9 - w_8$ ; when  $\tau = 2$  we use  $w_7 - w_5, w_8 - w_6$ , and  $w_9 - w_7$  and when  $\tau = 3$  we use  $w_8 - w_5$  and  $w_9 - w_6$ . Formally, we can justify this with the  $m$ -dependence assumption defined in equation (3) with  $h^* = 4$ .

In Tables 4 and 5 we present the covariance of initial wage and wage growth for potential and actual experience respectively. We build up to the main result by first presenting results using more restrictive sets. In the first column we present results using initial wages and one year wage differences (i.e.  $R = \{T, \tau : T \leq 9, \tau = 1, T - \tau \geq 5\}$ ); in the second column we use two

year wage differences; in the third column we use three year wage differences; and in the final column we aggregate one, two and three year wage differences. The first row presents the covariance of initial wage with wage differences. As robustness checks, the second row presents the covariance of the first year wage with wage differences, and the third row presents combines the covariances of initial- and first period wages with wage differences.

Our main result for all cases is a small negative point estimate of the covariance that is usually statistically insignificant. In some cases, the point estimate is significantly different from zero, particularly when using three period wage differences. However, when we use the larger sample of individuals by combining the data in the fourth column, the point estimate approaches zero and the effect is no longer significant.

Comparing the fourth column with the first three reveals the relevance of sample selection bias. The combined result in the fourth column is not a simple average of the first three columns because to deal with the sample selection bias we are weighting by individual not by wage observations. In general, the estimated effect controlling for sample selection is smaller than the estimated effect using fewer individuals as in columns (1)-(3).

One issue that arises is in judging the magnitude of the result. Our preferred estimate of the covariance of initial wage and wage growth, presented in Table 5 column 4 row 1, is -0.0021, which seems like a small number, but judging the magnitude of a covariance is difficult. As an informal method to get a sense of the magnitude of the covariance, we posit the following linear

approximation

$$E(\beta_i | \theta_i) = \gamma_0 + \gamma_1 \theta_i.$$

This implies that,

$$\gamma_1 = \frac{\text{cov}(\theta_i, \beta_i)}{\sigma_\theta^2}$$

where  $\sigma_\theta^2$  is the variance of  $\theta_i$ . To judge the magnitude imagine ask the following question: How much more wage growth would we expect from an individual whose ability is one standard deviation above the mean? This effect can be written as

$$\begin{aligned} E(\beta_i | \theta_i = \mu_\theta + \sigma_\theta) - E(\beta_i | \theta_i = \mu_\theta) &= \gamma_1 \sigma_\theta \\ &= \frac{\text{cov}(\theta_i, \beta_i)}{\sigma_\theta}. \end{aligned}$$

We can calculate this effect as long as we jointly estimate  $\sigma_\theta^2$  along with  $\text{cov}(\theta_i, \beta_i)$ .

It turns out that estimation of  $\sigma_\theta^2$  can be handled in a manner that is analogous to estimation of  $\text{cov}(\theta_i, \beta_i)$ . First consider the potential experience case. In general, we can identify  $\sigma_\theta^2$  using the fact that:

$$\sigma_\theta^2 = \text{cov}(w_{i0}, w_{iT}) - PE_{iT} \times \text{cov}(\theta_i, \beta_i).$$

In practice, we construct moments using data from all periods that are available for each individual. Moments for actual experience are constructed analogously, although the algebra is a bit more complicated.<sup>9</sup>

---

<sup>9</sup>The moments actually used to estimate  $\sigma_\theta^2$  are described below. Let  $I_{iT}$  be a dummy variable indicating whether  $w_{iT}$  is observed for individual  $i$  at time  $T$ . Assume that

In practice we estimate the three parameters  $\mu_\theta$ ,  $cov(\theta_i, \beta_i)$ , and  $\sigma_\theta^2$  with three moments: the moment we have been using for estimation of  $cov(\theta_i, \beta_i)$  defined in equation (5), the moments just presented for estimation (presented explicitly in footnote 9), and the sample mean of  $w_{i0}$  to estimate  $\mu_\theta$  (i.e.  $E(w_{i0} - \mu_\theta) = 0$ ). The results from this estimation procedure are presented in Table 6. Note that the first row of the table corresponds exactly to the analogous numbers in Table 5.

In the fourth row of Table 6 we combine the results from the other rows to get an estimate of  $\frac{cov(\theta_i, \beta_i)}{\sigma_\theta}$ . For potential experience the point estimate suggests that a one standard deviation increase in unobserved ability is associated with average wage growth of 0.61 percent less per year. For actual experience a one standard deviation increase in unobserved ability is associated with a decrease in average wage growth of 0.87 of a percentage point per full year of work. Neither of these results is statistically significant.

Another way of judging the magnitude of this coefficient is by measuring its contribution to the wage variance. It is well known that the variance  $\overline{cov(u_{i0}, u_{iT})} = 0$  for  $T \geq t^\ell$ , and that the return to experience is linear between  $t^\ell$  and  $t^u$ . We can then use the moment condition

$$E \left( \frac{\sum_{T=t^\ell}^{t^u} I_{iT} [(w_{i0} - \mu_\theta) w_{iT} - PE_{iT} cov(\theta_i, \beta_i)]}{\sum_{T=t^\ell}^{t^u} I_{iT}} \right) = \sigma_\theta^2.$$

For the actual experience model, the expression is more complicated,

$$(8) \quad E \left( \frac{(w_{i0} - \mu_\theta)}{\sum_{T=t^\ell}^{t^u} I_{iT}} \sum_{T=t^\ell}^{t^u} I_{iT} \left[ w_{iT} - \frac{\sum_{(s,\tau) \in R} d_{is\tau} (w_{i0} - \bar{w}_0) \left( \frac{w_{is} - w_{is-\tau}}{AE_{is} - AE_{is-\tau}} \right) AE_{iT}}{\sum_{(s,\tau) \in R} d_{is\tau}} AE_{iT} \right] \right) = \sigma_\theta^2.$$

of log wages is increasing both over the lifecycle and over time (during the period of this sample). This makes the negative coefficient on the covariance surprising. To gauge the importance of this covariance in log wages we focus on potential experience and note that following specification (2),

$$\text{var}(w_{it} | PE_{it}) = \text{var}(\theta_i) + \text{var}(\beta_i)PE_{it}^2 + \text{cov}(\theta_i, \beta_i)PE_{it} + \text{var}(u_{it} | PE_{it}).$$

To gauge the importance of the term  $\text{cov}(\theta_i, \beta_i)PE_{it}$ , in Figure 1 we plot both  $\text{var}(w_{it} | PE_{it})$  and  $[\text{var}(w_{it} | PE_{it}) - \text{cov}(\theta_i, \beta_i)PE_{it}]$  versus  $PE_{it}$ . Since our panel is aging the increase in the variance is due to changes both over time and over age. One can see that both the gross and net variance of wages do increase over the lifecycle. While the negative effect is not trivial, it offsets only a fraction of the overall lifecycle increase in variance.

One potential criticism is the specification of the model. There are two major assumptions that we have made:

- log wages are linear in experience over the first ten years following labor force entry
- $\text{cov}(u_{i0}, u_{it}) \simeq 0$  for  $t > T - \tau$ .

We view the first problem as a question of whether  $T$  is small enough. It is well known that lifecycle earnings are concave and decline later in life. However they are approximately linear in the first part of the lifecycle, so that a linear specification should fit the data for sufficiently small  $T$ . We view the second problem as a question of whether  $T - \tau$  is far enough away

from 0 that the errors are approximately uncorrelated. A natural way of testing our specification relative to both of these potential problems is to test the overidentification condition as to whether  $E \left[ (w_0 - \mu_0) \left( \frac{w_{iT} - w_{iT-\tau}}{AE_{iT} - AE_{iT-\tau}} \right) \right]$  varies across  $T$ .

With the sample selection problem, the model is not overidentified in the classic sense that we can not obtain consistent estimates of all of the moments conditional on  $t$  and  $\tau$ . That is, without this problem we could treat the model as a GMM problem with a separate moment condition

$$E \left[ (w_0 - \mu_0) \left( \frac{w_{iT} - w_{iT-\tau}}{AE_{iT} - AE_{iT-\tau}} \right) \right] = cov(\beta_i, \theta_i)$$

for each  $T$  and  $\tau$ . We could estimate this model and use the standard GMM test of overidentification restrictions.

We address the sample selection problem in two different ways. First, we ignore it by only using the observations we have to estimate each of these moments. The p-value from this test is presented in the fifth row of Table 6. One can see that we do not reject either the potential or actual experience model. Our second approach is to use the average value from the observations we do have on  $i$  when we don't have data on  $\left( \frac{w_{iT} - w_{iT-\tau}}{AE_{iT} - AE_{iT-\tau}} \right)$  (i.e.  $m_i$  defined in equation 6). While the unrestricted model would not be consistent, this is still valid as an overidentification test. The results from this approach are presented in the sixth row of Table 6. As before, we fail to reject either specification. We have experimented with alternative assumptions and chose this one based on this test. For example increasing  $t^u$  to 10 leads to a rejection

of the potential experience model at the ten percent level. However, the basic result of little relationship between wage levels and growth is very robust.

## 7 Conclusions

The primary goal of this paper is to estimate the relationship between wage growth and wage levels for low skill workers. We do so by taking advantage of panel data in a random coefficient Mincer style wage model. We allow both the intercept and slope coefficient to be arbitrarily correlated with actual experience. Our results here generally confirm our previous results that there is little relationship between skill level and wage growth. Our point estimates suggest that a one standard deviation increase in permanent ability reduces the return to annual potential experience by 0.61 of a percentage point. The analogous point estimate for actual experience is a reduction of 0.87 of a percentage point. These estimates indicate that the rate of wage growth with work experience should be at least as high for low skill workers as for medium skill workers. Neither of these effects is statistically significant at conventional levels.

There are a number of caveats to keep in mind. First, we are measuring the effects of experience on log wage increases, not level increases. If low wage workers have similar levels of log wage growth, then their gains in wage levels would be smaller than for higher wage workers. Second, the magnitude of these effects is not huge. A full year of labor force experience leads to wage gains of approximately 4%. This is nontrivial, but it would not have a huge

effect on earnings inequality or poverty rates. Finally, the workers we study are not necessarily representative of the type of workers who would respond to particular policy changes. Sample size (and endogeneity) does not allow us to condition on welfare mothers as one might like.

While no particular specification that we use is perfect, our results are robust across a number of different underlying assumptions. Measured by covariances, wage growth among low skill workers is similar to or perhaps somewhat higher than wage growth among higher skill workers.

## References

- Abowd, J and Card, D. 1989. On the Covariance Structure of Earnings and Hours Changes. *Econometrica* 57: 411-445.
- Baker, M. 1997. Growth Rate Heterogeneity and the Covariance Structure of Life-Cycle Earnings. *Journal of Labor Economics* 15: 338-375.
- Baker, M and Solon, G. 1998. Earnings Dynamics and Inequality among Canadian Men 1976-1992: Evidence from Longitudinal Income Tax Records. *Journal of Labor Economics* Vol 21,no. 2: 289-321.
- Ben-Porath, Y. 1967. The Production of Human Capital and the Life Cycle of Earnings. *Journal of Political Economy* 75 : 352-65.
- Berry, S, Levinsohn, J and Pakes A. 1995. Automobile Prices in Market Equilibrium. *Econometrica* 63: 841-890.
- Burtless, G. 1994. Employment Prospects for Welfare Recipients. In *The Work Alternative, Welfare Reform and the Realities of the Job Market*. Nightingale DS, Haveman RH (eds). Urban Institute Press: Washington DC.
- Cameron, S and Heckman, J. 1993. The Nonequivalence of High School Equivalentents. *Journal of Labor Economics* 11: 1-47.
- Connolly, H, Gottschalk, P. 2000. Returns to Tenure and Experience Revisited: Do Less-Educated Workers Gain Less from Work Experience?

Economics department, Boston College, working paper.

Corcoran, M and Loeb, S. 2001. Welfare, Work Experience, and Economic Self-Sufficiency. *Journal of Policy Analysis and Management*. 20: 1-20.

French, E, Mazumder, B and Taber, C. 2006. The Changing Pattern of Wage Growth for Low Skilled Workers. Unpublished manuscript.

Fullerton, D and Rogers, D. 1993. Who Bears the Lifetime Tax Burden. Brookings Institute Press: Washington D.C.

Gladden, T and Taber, C. 2000. Wage Progression Among Less Skilled Workers. in *Finding Jobs: Work and Welfare Reform*. Card D, Blank R (eds). Russell Sage Foundation: New York.

Gladden, T and Taber, C. 2004. The relationship between Wage Growth and Wage Levels, unpublished manuscript.

Gokhale, J and Kotlikoff, L. 2002. The Impact of Social Security and Other Factors on the Distribution of Wealth. In *The Distributional Aspects of Social Security and Social Security Reform*. Feldstein M, Liebman J (eds). University of Chicago Press: Chicago.

Heaton, J and Lucas, D. 1996. Evaluating the Effects of Incomplete Markets on Risk Sharing and Asset Pricing. *Journal of Political Economy* 104: 443-487.

- Heckman, J. 1997. Instrumental Variables: A Study of Implicit Behavioral Assumptions in One Widely Used Estimator . *Journal of Human Resources* 32: 1–40.
- Heckman, J and Lochner, L. 2002. Distinguishing Between On-the-job Training and Learning-by-doing. Unpublished manuscript, .
- Heckman, J, Lochner, L and Taber, C. 1998. Explaining Rising Wage Inequality: Explorations With A Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents. *Review of Economic Dynamics* 1: 1-55.
- Heckman, J and Robb, R. 1985. Alternative Methods for Evaluating the Impact of Interventions. In *Longitudinal Analysis of Labor Market Data*. Heckman J, Singer B(eds). Cambridge University Press: Cambridge.
- Heckman, J and Vytlacil, E. 1998. Instrumental Variables for the Correlated Random Coefficient Model: Estimating the Average Rate of Return to Schooling when the Return is Correlated with Schooling. *Journal of Human Resources* 33: 974-1002.
- Imbens, G and Angrist J. 1994. Identification and Estimation of Local Average Treatment Effects. *Econometrica*. 62: 467-475
- Katz, L and Autor, D. 1999. Changes in the Wage Structure and Earnings Inequality. In *Handbook of Labor Economics Vol 3A*. Ashenfelter O, Card D (eds). Amsterdam: Elsevier-North Holland.

Lillard, L and Reville, R. 1999. New Testable Implications of the Life Cycle Human Capital Investment Model. unpublished manuscript, .

Moffitt, R and Rangarajan, A. 1989. The Effect of Transfer Programmes on Work Effort and Human Capital Formation: Evidence from the US. In *The Economics of Social Security* Dilnot A, Walker I (eds). Oxford University Press: Oxford.

Woodridge, J. 1997. On Two Staged Least Squares Estimation of the Average Treatment Effect in Random Coefficient Models. *Economic Letters* 56: 129-133.