J Popul Econ (2009) 22:463–499
DOI 10.1007/s00148-007-0180-5

ORIGINAL PAPER

Work experience as a source of specification error in earnings models: implications for gender wage decompositions

Tracy L. Regan · Ronald L. Oaxaca

Received: 4 August 2006 / Accepted: 5 November 2007 /
Published online: 16 February 2008

Abstract This paper models the bias from using potential vs actual experience in log wage models. The nature of the problem is best viewed as specification error as opposed to classical errors-in-variables. We correct for the discrepancy between potential and actual work experience and create a predicted measure of work experience. We use the 1979 National Longitudinal Survey of Youth and the Panel Study of Income Dynamics and extend our findings to the Integrated Public Use Microdata Sample. Our results suggest that potential experience biases the effects of schooling and the rates of return to labor market experience. Using such a measure in earnings models underestimates the explained portion of the male–female wage gap. We are able to separately identify the decomposition biases associated with incorrect experience measures and biased parameter estimates.

Keywords Experience · Decomposition · Specification error

JEL Classification C81 · J24 · J31

Responsible editor: Deborah Cobb-Clark

T. L. Regan (✉)
Department of Economics, University of Miami, P.O. Box 248126,
Coral Gables FL 33124-6550, USA
e-mail: tregan@miami.edu

R. L. Oaxaca
IZA, Bonn, Germany

R. L. Oaxaca
Department of Economics, University of Arizona, McClelland Hall #401,
P.O. Box 210108, Tucson, AZ 85721-0108, USA
e-mail: rlo@u.arizona.edu

1 Introduction

This paper models the bias inherent in the use of potential, as opposed to actual, work experience in human capital earnings models. The implications of such bias for gender wage decompositions are explored. In a departure from the previous literature, our analysis identifies how much of the specification bias for the explained and the unexplained components of gender wage decompositions can be associated with biased levels of mean work experience and how much can be associated with bias in the parameter estimates. The paper also considers alternative ways of attempting to measure actual work experience when such data are available. At a minimum, traditional (Mincerian) log wage equations employ one's completed schooling, ability (if available), prior work experience, and its square. The latter two variables capture the concave relationship that exists between labor market experience and its pecuniary rewards.[1] Many data sets do not contain actual work histories, however. The 1979 National Longitudinal Survey of Youth (NLSY79) and the Panel Study of Income Dynamics (PSID) are two exceptions and are thus preferable in many instances. Because of their relatively small size and unrepresentativeness, such data sets are not always desirable (Moulton 1986). Thus, one is often forced to proxy for actual work experience with potential work experience—measured as the time elapsed since leaving school (which in turn is often proxied by age minus schooling minus five or six). Such a measure assumes continuous work histories and abstracts away from employment status (over-, full-, or part-time) and multiple-job holding.

Most researchers seem content using potential work experience measures for males, as it is not unreasonable to assume that they have been in the labor force continuously since leaving school. The use of such a measure for females is viewed less favorably and is a potential problem when making comparisons between working men and women. While the Mincer proxy for work experience has become standard practice, we argue that the use of such a measure may still be problematic for males. Male workers, like their female counterparts, experience employment lapses. Such lapses take two different forms—namely, an active job search while unemployed or a withdrawal from the labor force. It is unreasonable to assume that one's labor market experience is affected in the same way by these two different forms of employment lapses. Furthermore, one would not expect the "return" to unemployed labor force experience to be the same as that of employed labor force experience.

[1]Murphy and Welch (1990) make the case for specifying work experience as a quartic. Because much of the literature continues to use a quadratic specification (see Blau and Kahn 1996 and Antecol and Bedard 2002, 2004), we confine our attention to the quadratic case, leaving for future work an examination of the quartic specification.

This paper employs data collected from the NLSY79 and the PSID to address specification error in work experience and to examine the implications for gender wage decompositions. The standard errors-in-variables framework is violated here because the discrepancy between potential and actual work experience does not have a mean of zero and is found to be correlated with actual measures of work experience. As an alternative to viewing the problem as one of classical measurement error, we believe it is more fruitful to think of the problem as one of specification error. We investigate the extent to which actual experience can be predicted from other variables and extend our predicted work experience measures to a data set in which actual measures of work experience are not available—specifically, the 1990 wave of the Integrated Public Use Microdata Sample (IPUMS). The potential work experience measure overstates the effects of schooling and underestimates the rates of return to labor market experience. Our predicted work experience measures generally lead to a substantial reduction in the bias on the schooling and experience coefficients. The single exception arises from the estimated coefficient on experience squared for the NLSY79 females. Furthermore, more of the male–female wage gap is explained when actual or our predicted work experience measures are used in lieu of potential experience.

The paper proceeds in the following fashion: Section 2 provides the background and literature review. Section 3 discusses the conceptual framework that underlies the analysis. Section 4 discusses the data used in the analysis. Section 5 presents and discusses the results. Section 6 concludes.

2 Background and literature review

Measurement error is a problem commonly faced in applied work. For practical purposes, measurement error in the endogenous variable is not problematic because it is usually assumed to be uncorrelated with the regressors. The R^2 of the regression is smaller, however, because of the additional noise contained in the random error term. Conversely, measurement error in regressors does pose serious problems because it leads to biased and inconsistent OLS estimates. When faced with such a problem, the tendency is to assume that the measurement error is classical, in the sense that the true regressors and their measurement error are uncorrelated, the random disturbance term and the measurement errors are asymptotically uncorrelated, and the measurement errors are uncorrelated among themselves. These assumptions lead naturally to instrumental variables (IV) estimation as the preferred method of bias correction.

The standard assumptions placed on the measurement error typically arise from convenience and are not usually supported by empirical evidence (e.g., Black et al. 2000). Duncan and Hill (1985) and Rodgers et al. (1993) use administrative records from a large manufacturing firm to verify workers' responses to questions pertaining to earnings and hours worked, while Bound and Krueger (1991) and Bound et al. (1994) examine measurement error in longitudinal earnings data. Bollinger (1998) also examines measurement error in panel data but uses a nonparametric methodology. Lee and Sepanski (1995) offer a computationally and analytically simpler method to the nonparametric methodology in consistently estimating regression models with measurement error in the dependent and/or independent variables when validation data are available. In sum, all of these validation studies confirm measurement error in survey data and their findings contradict many of the assumptions made in and implications drawn from classical measurement error models.

There is a limited body of research that attempts to address the measurement error associated with the use of potential experience. The issue has been probed, mainly for females, in hopes of addressing the bias that accrues to the OLS estimates when potential work experience measures are employed instead of reported actual experience. Garvey and Reimers (1980) argue that the use of the traditional age minus schooling minus six measure is biased because people do not always complete one grade per year (either due to acceleration or retention) and that one accrues nonwork time during one's life. Antecol and Bedard (2002, 2004) note that potential experience is less accurate for individuals who are less attached to the labor market (e.g., blacks, Mexicans, and women). Antecol and Bedard (2002, 2004) address the implications associated with the use of potential vs actual experience in explaining black/white and Mexican/white wage gaps for men and women separately. Garvey and Reimers (1980), Moulton (1986), and Filer (1993) construct predicted measures of work experience as alternatives to the standard Mincerian measure. Below, we discuss how our approach differs from those found in these studies.

3 Conceptual framework

As is evident from the previous discussion, much of the literature on measurement error in human capital models has focused on measurement error in a linear term (e.g., schooling). Several researchers (e.g., Ashenfelter and Rouse 1998; Behrman and Rosenzweig 1999) have noted the complications introduced by measurement error in a quadratic variable but few have tackled the problem. It is our intention to examine this issue more closely in the context of specification error. Note that we abstract away from any measurement errors that may arise with respect to age and schooling in the potential experience variable and away from measurement error in the direct measures of actual work experience. The objective here is to determine how the use of potential experience vs directly measured actual work experience affects parameter estimates in earnings models and inferences about gender wage inequality from wage decompositions. Actual work experience, as directly measured, is the standard against which potential and predicted work experience effects are compared.

Our discussion of specification error will be framed in the simplest of models—a traditional (Mincerian) log wage equation,

$$Y_i = \beta_0 + \beta_1 S_i + \beta_2 X_i^* + \beta_3 X_i^{*2} + \sum_{i=1}^{K} \alpha_i H_i + \varepsilon_i,\ i = 1, ..., N, \tag{1}$$

where Y is the natural log of the hourly wage, S is the schooling level, X^* is actual work experience, H is a set of K other control variables, ε is a random error term, i indexes the individual, and N represents the sample size. More compactly, we can express Eq. 1 as,

$$Y = W^* \gamma + \varepsilon,$$

where Y and ε are $(N \times 1)$ vectors, W^* is the $(N \times (K+4))$ observation matrix, and γ is the $((K+4) \times 1)$ coefficient vector. Taking the probability limit of the OLS estimator yields:

$$plim(\widehat{\gamma}) = \gamma + \Sigma_{W^* W^*}^{-1} \Sigma_{W^* \varepsilon},$$

which is consistent only if $plim(N^{-1} W^{*'} \varepsilon) = \Sigma_{W^* \varepsilon} = 0$. Thus, the regressors, specifically schooling and experience, must be exogenously determined (i.e., uncorrelated with ε).[2]

Now, suppose that actual work experience, X^*, is unobserved. Instead, one observes X, which can be thought of as potential work experience. For simplicity, we can express the relationship between potential and actual work experience as:

$$X_i = X_i^* + v_i, \tag{2}$$

where v is the discrepancy between the experience measures. At this point, we will allow that v may be correlated with X^* and that the mean of v may not be, and most probably is not, zero.[3] As is traditionally the case, we will, however, assume that there is no correlation between v and ε.

The nature of the model misspecification problem we are considering can be seen by substituting Eq. 2 into Eq. 1, yielding,

$$Y_i = \beta_0 + \beta_1 S_i + \beta_2 X_i + \beta_3 X_i^2 + \sum_{i=1}^{K} \alpha_i H_i + \varepsilon_i^*, \tag{3}$$

where $\varepsilon_i^* = \varepsilon_i - \beta_2 v_i - 2\beta_3 X_i^* v_i - \beta_3 v_i^2$.

[2]Several researchers have noted the endogenous nature of schooling (e.g., Bound and Solon 1999; Black et al. 2000).

[3]The mean of v could be positive if potential work experience overstates actual work experience, which is likely the case for many females. If, however, potential work experience understates one's actual work experience, which is more likely for those who work overtime or who hold multiple jobs, $E(v)$ would be negative.

More compactly, Eq. 3 can be expressed as,

$$Y = W\gamma + \varepsilon^*,$$

where W is the $(N \times (K+4))$ new observation matrix and ε^* is the new $(N \times 1)$ error vector. The error vector ε^* may be expressed as,

$$\varepsilon^* = \varepsilon - v\beta_2 - 2\left[X^* \odot v\right]\beta_3 - [v \odot v]\,\beta_3,$$

where $X^* \odot v$ and $v \odot v$ are Hadamard products (i.e., element-by-element multiplication between X^* and v and between v and v , respectively).[4]

The probability limit of the OLS estimates is:

$$\begin{aligned} plim(\widehat{\gamma}) &= \gamma + \Sigma_{WW}^{-1}\Sigma_{W\varepsilon} - \Sigma_{WW}^{-1}\Sigma_{Wv}\beta_2 \\ &\quad -2\Sigma_{WW}^{-1}\Sigma_{W,X^*\odot v}\beta_3 - \Sigma_{WW}^{-1}\Sigma_{W,v\odot v}\beta_3 \\ &= \gamma - \Sigma_{WW}^{-1}\Sigma_{Wv}\beta_2 - 2\Sigma_{WW}^{-1}\Sigma_{W,X^*\odot v}\beta_3 - \Sigma_{WW}^{-1}\Sigma_{W,v\odot v}\beta_3, \end{aligned}$$

assuming $\Sigma_{WW}^{-1}\Sigma_{W\varepsilon} = 0$. Now, with specification error associated with substitution of X for X^*, the asymptotic bias in $\widehat{\gamma}$ consists of three distinct terms.

Our approach to correcting for specification error consists of modeling actual experience as a stochastic regressor generated from a semi-log model:

$$\ell n\left(X_i^*\right) = Z_i\gamma_1 + \psi_{1i}, \tag{4}$$

where Z is a set of regressors that includes the regressors in Eq. 1 (i.e., S, H) and a set of identifying variables (i.e., a respondent's age, a set of occupational dummy variables, and the number of children for females), and ψ_{1i} satisfies the standard assumptions without any particular distributional assumption.[5]

The semi-log specification bounds X_i^* away from zero. Our proposed correction procedure uses a predicted measure of actual work experience constructed in the following fashion:

$$\widehat{X_i^*} = \hat{\delta}_1 \exp\left(Z_i\hat{\gamma}_1\right),$$

where $\hat{\gamma}_1$ is obtained from OLS estimation of Eq. 4 and $\hat{\delta}_1$ is a scale factor that forces the predicted mean to match the sample mean: (see Oaxaca and Ransom 2003 and Sarnikar et al. 2007)

$$\hat{\delta}_1 = \frac{\sum_i X_i^*}{\sum_i \exp\left(Z_i\hat{\gamma}_1\right)}.$$

[4]In the present case, $X \odot v = \left[d(X)\right] v$ and $v \odot v = \left[d(v)\right] \odot v$, where $d(X)$ and $d(v)$ are $(N \times N)$ diagonal matrices formed by arraying the elements of vectors X and v along the principal diagonal. See Ding and Engle (2001) and Styan (1973) for more details.

[5]Note that the occupational dummies could have been included in the log wage equation. Some would argue that the inclusion of occupational controls biases downward estimates of discrimination.

While our procedure for predicting experience resembles instrumental variables, its motivation does not depend on endogeneity issues. Our motivation is simply to apply the correction model to data sets lacking information on actual experience.[6]

Our method of predicting work experience is more general and comprehensive than other methods used in the literature. We construct actual experience from hours worked, unlike Filer (1993) and Antecol and Bedard (2002, 2004), who use weeks worked.[7] Our measure is more inclusive, as it captures experience from multiple jobs and all hours worked (e.g., over-, full-, part-time). The gender wage decompositions address the implications associated with the notable differences in actual and potential work experience. Antecol and Bedard (2002, 2004) find substantial differences in the fraction of the black/white and Mexican/white wage gap that can be explained when actual, rather than potential, experience is used for men and women. We too observe these differences but also offer an alternative to the standard Mincerian proxy—namely, a predicted experience measure. Garvey and Reimers (1980), Moulton (1986), and Filer (1993) construct predicted measures of work experience using a linear model. We, however, use a semi-log specification, which bounds predicted experience away from zero. Garvey and Reimers (1980) address the nonnegativity of experience with the use of a Tobit model. The predicted measures Garvey and Reimers (1980) and Filer (1993) construct are birth cohort- and occupation-specific, respectively. Our predicted measure is more general and thus can be applied more readily to other data sets (e.g., IPUMS). This paper uses three data sets, while the other studies mentioned here rely exclusively on the NLSY. While we do not focus on racial/ethnic issues, we do consider the issues associated with potential experience for both women and men. With the exception of Antecol and Bedard (2004), who investigate the issue for white, black, and Mexican men, these other studies focus on women.

[6]Literally, IV would require that actual experience and its square be separately regressed on Z_i rather than simply regressing the log of actual experience on Z_i. On the other hand, an earnings model with $\ell n(X_i^*)$ as a regressor could not include $\ell n(X_i^*)^2$ because of perfect multicollinearity. The model could include $\left[\ell n(X_i^*)\right]^2$. In any event, the inclusion of log experience and/or the square of log experience would be a quite different specification from the standard Mincerian earnings model.

[7]The correlations between our measure of actual experience, based on the hours worked per week in the past calendar year, and those used by Filer (1993) and Antecol and Bedard (2002, 2004), based on weeks worked in the past calendar year, are 0.7980 and 0.8520 for the NLSY79 males and females used in our sample, respectively. Correspondingly for the PSID, the correlation between the actual experience measures we construct from the family files and those from the individual files are 0.9874 and 0.9630 for the males and females.

Our empirical implementation of Eq. 3 includes completed schooling, marital status, industry dummies, regional dummies, and standard metropolitan statistical area (SMSA) dummies as the set of control variables, H.[8]

A particular focus of this paper is the implications of misspecification of work experience for gender wage gap decomposition. Without loss of generality, we will adopt the estimated wage structure for males as the comparison standard. Accordingly, the standard decomposition is expressed as,

$$\begin{aligned}\bar{Y}_m - \bar{Y}_f &= \left(\bar{X}^{m,a} - \bar{X}^{f,a}\right)\widehat{\beta}^{m,a} + \bar{X}^{f,a}\left(\widehat{\beta}^{m,a} - \widehat{\beta}^{f,a}\right)\\ &= \left(\bar{X}^{m,j} - \bar{X}^{f,j}\right)\widehat{\beta}^{m,j} + \bar{X}^{f,j}\left(\widehat{\beta}^{m,j} - \widehat{\beta}^{f,j}\right),\end{aligned}$$

where m and f denote males and females, a denotes actual experience, j denotes predicted or potential experience, $\bar{Y}$ is the mean log wage, $\bar{X}$ is the mean characteristics vector, and $\widehat{\beta}$ is the estimated parameter vector. The effects of experience specification bias on the endowment (explained) component of the wage decomposition can be decomposed into parameter bias and mean experience measure bias:

$$\begin{aligned}&\left(\bar{X}^{m,a}-\bar{X}^{f,a}\right)\widehat{\beta}^{m,a}-\left(\bar{X}^{m,j}-\bar{X}^{f,j}\right)\widehat{\beta}^{m,j}\\ &\quad= \left(\bar{X}^{m,a}-\bar{X}^{f,a}\right)\left(\widehat{\beta}^{m,a}-\widehat{\beta}^{m,j}\right) + \left[\left(\bar{X}^{m,a}-\bar{X}^{f,a}\right)-\left(\bar{X}^{m,j}-\bar{X}^{f,j}\right)\right]\widehat{\beta}^{m,j}.\end{aligned} \tag{5}$$

The first term on the right-hand side (rhs) of Eq. 5 represents the difference in the explained wage gap component that arises because of differences in the estimated parameters. The second term on the rhs of Eq. 5 represents the difference in the explained wage gap component that arises because of mean differences in the measures of experience. The effects of experience specification bias on the discrimination (unexplained) component of the wage decomposition can also be decomposed into parameter bias and mean experience measure bias:

$$\begin{aligned}&\bar{X}^{f,a}\left(\widehat{\beta}^{m,a} - \widehat{\beta}^{f,a}\right) - \bar{X}^{f,j}\left(\widehat{\beta}^{m,j} - \widehat{\beta}^{f,j}\right)\\ &\quad= \bar{X}^{f,j}\left[\left(\widehat{\beta}^{m,a}-\widehat{\beta}^{f,a}\right)-\left(\widehat{\beta}^{m,j}-\widehat{\beta}^{f,j}\right)\right]+\left(\bar{X}^{f,a}-\bar{X}^{f,j}\right)\left(\widehat{\beta}^{m,a}-\widehat{\beta}^{f,a}\right).\end{aligned} \tag{6}$$

[8]Mincer and Polachek (1974) provide a different approach than that used in our paper but not an entirely different model. They argue that the simple Mincer earnings function, which uses potential work experience, is biased due to omitted variables. Mincer and Polachek (1974) decompose potential work experience into actual (or instrumented) work experience and "home time" (i.e., time spent out of the labor force) and derive a segmented earnings function instead. This omitted-variables approach may explain why our specification error term is nonclassical and one-sided. We thank Solomon Polachek for these comments and insights. For further discussion of the life-cycle division of labor within a family unit, see Polachek (1975).

The first term on the rhs of Eq. 6 represents the difference in the unexplained wage gap component that arises because of differences in the estimated parameters. The second term on the rhs of Eq. 6 represents the difference in the unexplained wage gap component that arises because of mean differences in the measures of experience. Note that the only differences in the mean characteristics vectors between actual and potential or predicted experience stem from the differences between mean actual experience and its square and mean potential or predicted experience and its square. On the other hand, all of the parameter estimates can differ between specifications that use actual experience and those using either potential or predicted experience.

4 Data

The data used in this paper come from the NLSY79, the PSID, and the IPUMS. We focus on 1990 because it is a Census year common to all of our data sets and permits analysis of relatively young cohorts as well as a more broadly defined age grouping. The NLSY79 consists of 12,686 young men and women, living in the USA, who were between the ages of 14 and 22 when the first wave of the survey was conducted in 1979. The PSID is a longitudinal study that began in 1968. There were 4,800 families included in 1968, and the largest amount and most detailed information is collected for the head of the household. For this reason, amongst others, we restrict our sample to heads of household who are between the ages of 18 and 55 in 1990. The IPUMS is a collection of 25 cross-sectional samples spanning the 1850–2000 US Census years. To ensure comparability of our results across the data sets, we divided the IPUMS into two subsamples: (1) individuals between the ages of 25 and 33 in 1990 and (2) heads of household between the ages of 18 and 55 in 1990. The former construction most closely parallels the NLSY79 and the latter the PSID. In this paper, we abstract away from racial/ethnic issues by restricting our attention to whites. Although it would be interesting to extend our analysis to racial/ethnic comparisons, we leave this to future work.

The dependent variable used in the log wage equations is the hourly wage. We construct this measure by dividing the total income from wages and salary by the annual hours worked. For the NLSY79 and the PSID, the construction of the annual hours worked will be discussed below. For the IPUMS, this measure is just the product of the weeks worked last year and the usual hours worked per week. The control variables, H, are defined as follows: The schooling variable corresponds to the highest grade completed as of 1990. $MARRIED$ takes on a value of "1" if the individuals were married in 1990 and "0" otherwise (i.e., single, separated, annulled, divorced, or widowed). The industry dummies refer to the 1970 Census of the Population's Industry Classification System. The left-out reference group is public administration. The regional dummies correspond to the Northeast, North Central, West, and South. The omitted regional dummy is West. The SMSA dummies are: (1) not living in a SMSA, (2) living in a SMSA that is not a central city, (3) living in a

SMSA where the central city is unknown, and (4) living in a SMSA with central city known. The omitted reference group is not living in a SMSA.[9]

The actual work experience measures correspond to the "years" of fulltime equivalent (FTE) work experience accumulated as of the 1990 interview. The analysis considers only those who report at least 1 year of FTE work experience. In constructing this measure for the NLSY79, we used the "hours worked on all jobs" each week for a given calendar year. We summed these figures for 1977 through 1989 and then divided the total by 2,080 (40 *hours per week* $\times$ 52 *weeks per year*) to obtain a measure of FTE work experience.[10]

Information contained in the PSID family files on the "head's annual hours working for money" in 1967–1982 and the "head's total annual work hours" for 1983–1989 was used to construct actual work experience measures. Again, we summed the annual hours worked and divided through by 2,080. Doing so produced figures that were implausibly low.[11] Consequently, we redefined the actual work experience measures for the PSID in the following fashion: First, we referenced the information contained in response to the question asking how many years an individual worked since the age of 18 (inclusive). While the mean value of this variable is reasonable, this question elicited some unrealistic responses. For instance, there were several cases in which an individual reported having worked 98 years! Because this was obviously impossible, due to the fact that the oldest respondent considered was 55 years old, we identified such individuals and assigned them the maximum allowable calendar years of work experience possible since the age of 18. Next, we summed the number of years in which annual hours were reported. If,

$$\begin{aligned}&\#\, of\ years\ in\ which\ annual\ hours\ are\ reported\\ &> max.\ allowable\ calendar\ years\ of\ work\ experience\ since\ the\ age\ of\ 18,\end{aligned}$$

then the work experience measure used was simply the cumulative hours reported divided by 2,080. If the above inequality did not hold, then we constructed the work experience measure as:

$$\left(\frac{(cumulative\ hours\ reported)/2080}{\#\, of\ years\ in\ which\ annual\ hours\ are\ reported}\right)\times$$
$$(max.\ allowable\ calendar\ years\ of\ work\ experience\ since\ the\ age\ of\ 18)\,,$$

[9]For the PSID, information on a respondent's SMSA is available only until 1986. Alternatively, we considered using city size indicators and rural/urban status. The IPUMS does not contain comparable measures for the city size indicators, however, and including a respondent's rural/urban status does not change the results significantly. Thus, the regressions corresponding to the heads of household, ages 18–55, do not include such measures.

[10]Note that, most often, these "years" of work experience do not coincide with calendar years and actual work experience values would differ somewhat with alternative definitions of FTE.

[11]Part of the difficulty in using the PSID data is that the coding scheme does not allow one to distinguish between valid skips (i.e., missing information) and 0 values.

which simply amounts to assigning the average annual work hours to those years in which such information was missing, plus the annual work hours as reported in the PSID. Doing so yielded much more reasonable estimates of an individual's work history as of 1990.[12]

The potential work experience measure is calculated as follows:

$$\textit{Potential Work Experience} = \textit{Age} - \textit{Schooling} - 6. \tag{7}$$

Our sample consists of white males and females with at least 1 year of actual (and potential) work experience accumulated as of 1990. We omit individuals who were missing information on any of the aforementioned variables and exclude military personnel as well. In addition, we also exclude farmers and farm managers for the female heads of household regressions.

5 Estimation and results

Table 1 provides the descriptive statistics on the work experience measures across the data sets. The males, on average, have more work experience accumulated as of 1990 than the females. The NLSY79 males report 1.7 more years of FTE work experience than the females and the PSID males report 3.7 more years. For both data sets, the potential work experience measures overstate time actually spent working. The differences are slight for the males but quite pronounced for the females. The problems are less severe for the NLSY79 because this data set contains a relatively young set of respondents. Overall, the most dramatic difference between the potential and actual work experience measures is for the PSID females; there is a 5.4-year discrepancy between the potential measure (17.8 years) and the actual work experience measure (12.4 FTE years).

The descriptive statistics on the other control variables used in the log wage regressions were not published but are available upon request. With few exceptions, the sample characteristics were very nearly the same as between the NLSY79 and the IPUMS for the 25–33-year-old group and between the

[12]Both the NLSY79 and the PSID provide alternative measures of work experience. Specifically, the NLSY79 work history files contain hours worked in the past calendar year and the number of weeks worked per year. The PSID individual files also contain measures of the hours worked in a given year. Additionally, the PSID has information on the number of years a respondent has worked (full-time) since the age of 18. The number of weeks worked per year and the number of years worked since the age of 18 are obviously better than proxied work experience (i.e., potential measures) but they contain a fair amount of measurement error and do not allow for differences in employment status or multiple-job holding. Blank (1988) finds that, while being simultaneously determined, the hours worked per week decision is independent of the weeks worked per year decision. In spite of this, we did, however, re-estimate Eq. 3 using these alternative definitions and the results do not differ much.

Table 1 Work experience measures—descriptive statistics

	White males, ages 25–33				White males, heads of household, ages 18–55			
	NLSY79 ($N = 2789$)		IPUMS ($N = 3540$)		PSID ($N = 2892$)		IPUMS ($N = 9098$)	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Actual experience	9.186	3.168	–	–	16.063	9.099	–	–
(Actual experience)^2	94.423	61.396	–	–	340.781	373.201	–	–
Predicted experience	9.186	1.982	9.186	2.174	16.063	9.754	16.063	9.888
(Predicted experience)^2	88.314	38.279	89.114	42.291	353.134	458.350	355.777	434.100
Potential experience	9.715	3.265	10.057	3.504	17.129	8.813	18.612	9.293
(Potential experience)^2	105.027	64.892	113.423	76.242	371.046	350.949	432.748	379.450
	White females, ages 25–33				White females, heads of household, ages 18–55			
	NLSY79 ($N = 2386$)		IPUMS ($N = 3062$)		PSID ($N = 516$)		IPUMS ($N = 2579$)	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Actual experience	7.495	2.813	–	–	12.385	8.037	–	–
(Actual experience)^2	64.080	42.469	–	–	217.845	266.113	–	–
Predicted experience	7.495	1.648	7.495	1.793	12.385	7.357	12.385	7.421
(Predicted experience)^2	58.885	25.995	59.384	28.899	207.405	254.661	208.434	270.629
Potential experience	9.424	3.120	9.610	3.432	17.760	10.435	17.610	9.898
(Potential experience)^2	98.539	60.642	104.129	70.671	424.093	425.393	408.048	384.718

Actual and predicted experience are constructed from hours worked per week for the NLSY79 and from total annual work hours as reported in the family files for the PSID. Source of data: 1990 survey of the NLSY79, the PSID, and the IPUMS

PSID and the IPUMS for the 18–55-year-old-heads-of-household group. The main discrepancy was in the average hourly wage. The average hourly wage was consistently higher in the IPUMS data sets, ranging from only \$0.13 higher than the NLSY79 for the young female group to \$2.79 higher than the PSID for the 18–55 male heads of household group. The IPUMS hourly wage variable exhibited considerably more variation as well. Because other sample characteristics match up quite closely, these discrepancies can be attributed to differences in how hourly wages had to be constructed because of differential data restrictions between the IPUMS and the other data sets. A mitigating factor is that we are focusing on estimation and wage decomposition differences that arise from using actual and predicted experience compared with potential experience, *within* a given data set.

Based on the log wage regressions (available upon request from the authors), one can assess the extent of the bias correction for the estimated coefficients on schooling and work experience by referencing the percent differences. The percent difference between the estimated coefficients for the log wage regressions using either the predicted or the potential (j) experience measures vs actual work (a) experience is

$$\frac{\widehat{\beta_k}^{j} - \widehat{\beta_k}^{a}}{\widehat{\beta_k}^{a}}.$$

Regan et al. (2007) shows that, in the Mincerian simple schooling model, the coefficient on schooling is not identified because this model (1) is not based on an optimization framework, (2) does not control for ability (amongst other variables), and (3) is not concave in schooling. With respect to this last point, the simple schooling model is estimated with data representing tangencies between iso-present value lines and concave earnings/schooling production frontiers. Consequently, one cannot strictly interpret the coefficient on schooling as an internal rate of return. Nevertheless, for our purposes, we will follow past convention and treat the coefficient on schooling as the effect of schooling on wages.

Table 2 reports the effects on schooling and experience coefficients from using predicted vs potential experience instead of actual experience. In the log wage regressions for the NLSY79 white males, we find that the effect of schooling is overestimated by 1.3 vs 37.8 percent when predicted work experience is used vs potential work experience in lieu of actual work experience. See columns 1 and 2, respectively. Similarly, the estimated coefficients on the linear and quadratic experience terms are underestimated by 14.5 vs 61.7 and 50.4 vs 95.7 percent, respectively, when predicted vs potential work experience is used in place of actual work experience. By comparing columns 1 and 2, one sees that the discrepancy on the schooling and experience coefficients are much larger (in absolute value) when using potential work experience, instead of our predicted measure, in the log wage regression. For the NLSY79 females, the effect of schooling (relative to when actual work experience is used) is underestimated by 3.2 percent when predicted work experience is used and overestimated by 11 percent when potential experience

Table 2 Percent differences for estimated coefficients for schooling, experience, and experience^2

	NLSY79				PSID			
	White males, ages 25–33		White females, ages 25–33		White males, heads of household, ages 18–55		White females, heads of household, ages 18–55	
	Predicted vs actual (1) (%)	Potential vs actual (2) (%)	Predicted vs actual (3) (%)	Potential vs actual (4) (%)	Predicted vs actual (5) (%)	Potential vs actual (6) (%)	Predicted vs actual (7) (%)	Potential vs actual (8) (%)
% diff for schooling	1.3	37.8	−3.2	11.0	−0.2	19.0	0.9	15.3
% diff for experience	−14.5	−61.7	113.7	−115.7	0.8	−8.3	7.0	−37.2
% diff for experience^2	−50.4	−95.7	372.9	−161.0	−14.9	−27.6	13.0	−50.2

Actual and predicted experience are constructed from hours worked per week for the NLSY79 and from total annual work hours as reported in the family files for the PSID. Source of data: 1990 survey of the NLSY79 and the PSID is used. The pattern is similar to that of males in that the schooling estimates are off by a larger percentage in absolute terms when using potential rather than predicted work experience. However, for females, the use of predicted experience overstates and the use of potential experience understates the effects of both linear and quadratic experience. In absolute terms, the percent differences in returns to linear experience are approximately the same for predicted and potential experience. However, using our predicted measure in lieu of potential experience creates a significantly larger departure (in absolute terms) from the estimated coefficient on actual experience squared.

Differences in the estimated rates of return to experience occasioned by the use of either potential or predicted experience can be captured by the simple relationship:

$$\frac{\partial Y}{\partial X^j} - \frac{\partial Y}{\partial X^a} = \left(\hat{\beta}_1^j - \hat{\beta}_1^a\right) + 2\left(\hat{\beta}_2^j - \hat{\beta}_2^a\right) X, \tag{8}$$

where X is any given level of experience. On the basis of the negative percentage differences for both linear and quadratic experience appearing in columns 2 and 4 of Table 2, one can infer that both terms on the rhs of Eq. 8 are negative. This means that, for both males and females in the NLSY79, the use of potential rather than actual experience results in flatter estimated log earnings/experience profiles. On the other hand, this pattern holds for predicted vs actual experience only for males (column 1). For females in the NLSY79, the positive percentage entries for linear and quadratic experience coefficient estimates (column 3) imply that both terms on the rhs of Eq. 8 are positive. In other words, the use of predicted experience in place of actual experience results in a steeper estimated log earnings/experience profile for females. The peculiarities associated with the rates of return to work experience for the NLSY79 females may largely be stemming from the fact that this young sample has not accrued enough work experience to reach the concave portion of their wage–experience profile; the experience squared term only gains statistical significance when we use our predicted measure of work experience.

One can apply similar interpretations to the remaining figures in Table 2 corresponding to the PSID samples. We find that the log wage regressions using potential work experience overstate the effects of schooling relative to those using predicted work experience. With respect to the estimated rates of return to work experience, we again find that the log wage regressions using our predicted measures of work experience perform much better in absolute terms (relative to our actual experience measures) than those regressions using potential work experience. Overall, we can take these findings as continued evidence for the need to employ better proxies than time elapsed since leaving school for female and male work histories when such information is lacking.

Tables 3–8 report our results on the effects of different measures of work experience on the male–female wage gap and its decomposition. We separate out the contributions to the wage decomposition of gender differences in the

Table 3 Wage decomposition: NLSY79 white males and females, ages 25–33 (unadjusted male/female wage differential = 0.224)

	Endowment effects			Discrimination		
	Actual (1)	Predicted (2)	Potential (3)	Actual (4)	Predicted (5)	Potential (6)
Schooling	−0.032	−0.033	−0.045	−0.048	0.001	0.232
Experience	0.174	0.149	0.011	0.345	−0.254	0.456
Experience^2	−0.107	−0.051	−0.001	−0.153	0.212	−0.083
Other variables	0.034	0.030	0.040	0.011	0.171	−0.388
Total	0.070	0.095	0.006	0.154	0.129	0.218

Actual and predicted experience are constructed from hours worked per week for the NLSY79. Components may not add to totals due to rounding. Source of data: 1990 survey of the NLSY79

means and estimated coefficients associated with schooling, experience, and experience squared. The remaining contributions pertain to the constant term and the indicator variables and are summed and listed as "other variables." In reporting the estimates of discrimination (or the unexplained differential), the separate discriminatory components corresponding to the constant term and indicator variables are omitted because the estimated coefficients associated with these variables are not invariant to the choice of omitted reference groups (Oaxaca and Ransom 1999). Although the detailed endowment effects corresponding to the groups of indicator variables are invariant with respect to the choice of left-out reference groups, they are not reported separately because the differences across the alternative measures of work experience were negligible.

Table 3 reports the decomposition results for the 25–33 age group from the NLSY79 sample. The unadjusted male–female wage differential is 0.224 log points. For the log wage regression using actual work experience, the difference in average endowments account for 0.07 log points of the unadjusted wage differential. The partial contributions of schooling and experience to the endowment effect are −0.032 and 0.067, respectively, and are reported in

Table 4 Wage decomposition: IPUMS white males and females, ages 25–33 (unadjusted male/female wage differential = 0.223)

	Endowment effects		Discrimination	
	Predicted (1)	Potential (2)	Predicted (3)	Potential (4)
Schooling	−0.024	−0.033	−0.079	−0.096
Experience	0.197	0.009	0.041	0.138
Experience^2	−0.106	0.006	0.027	−0.014
Other variables	0.022	0.032	0.145	0.182
Total	0.089	0.013	0.134	0.210

Predicted experience is constructed from hours worked per week for the NLSY79. Components may not add to totals due to rounding. Source of data: 1990 survey of the IPUMS

Table 5 Wage decomposition: PSID white males and females, heads of household, ages 18–55 (unadjusted male/female wage differential = 0.297)

	Endowment effects			Discrimination		
	Actual (1)	Predicted (2)	Potential (3)	Actual (4)	Predicted (5)	Potential (6)
Schooling	0.004	0.004	0.005	−0.005	−0.016	0.031
Experience	0.160	0.162	−0.025	0.012	−0.020	0.235
Experience^2	−0.100	−0.101	0.031	−0.010	0.037	−0.088
Other variables	0.162	0.147	0.156	0.074	0.085	−0.048
Total	0.226	0.211	0.167	0.072	0.086	0.130

Actual and predicted experience are constructed from total annual work hours as reported in the family files for the PSID. Components may not add to totals due to rounding. Source of data: 1990 survey of the PSID

column 1. Similarly, for column 2, which corresponds to the log wage regression using predicted work experience, the endowment effects explain about 0.095 log points of the unadjusted wage differential. The schooling and experience endowment effects when using predicted experience correspond very closely to those estimated with actual experience. When work experience is measured as the time elapsed since leaving school (column 3), the endowment effects are 0.006. Thus, the use of potential experience would imply that virtually all of the gender wage gap is unexplained. Columns 4–6 tell the same story. Using actual or even predicted experience yields lower estimates of discrimination compared with potential experience. We note that the separate estimated effects of the discrimination components differ significantly between actual and predicted experience (columns 4 and 5), although the overall sums do not differ very much. This has largely to do with the problem of capturing concavity of the wage–experience profile for this relatively young group of workers.

Table 4 reports the decomposition results for the 25–33 age group from the IPUMS sample. Actual work experience is missing from the IPUMS data, so the comparison is between using predicted work experience and potential experience. The unadjusted gender wage gap of 0.223 log points is virtually

Table 6 Wage decompositions: IPUMS white males and females, heads of household, ages 18–55 (unadjusted male/female wage differential = 0.330)

	Endowment effects		Discrimination	
	Predicted (1)	Potential (2)	Predicted (3)	Potential (4)
Schooling	−0.016	−0.019	−0.186	−0.128
Experience	0.170	0.041	0.195	0.261
Experience^2	−0.109	−0.015	−0.046	−0.056
Other variables	0.079	0.087	0.243	0.157
Total	0.123	0.095	0.206	0.235

Predicted experience is constructed from total annual work hours as reported in the family files for the PSID. Components may not add to totals due to rounding. Source of data: 1990 survey of the IPUMS identical to that from the NLSY79 sample. The endowment effects account for 0.089 log points of the unadjusted gap when using predicted experience (column 1). On the other hand, the use of potential experience would imply that virtually none of the unadjusted gap is the result of endowment differences (column 2). Columns 3 and 4 show that the corresponding estimate of discrimination is much smaller when using predicted experience. As was the case with the NLSY79, the use of potential experience would imply that virtually the entire gender wage gap for the 25–33-year-olds is the result of discrimination or, at best, is unexplained.

Turning next to the broader group of workers aged 18 to 55 who are heads of households, we report the decomposition results for the PSID sample in Table 5. The unadjusted gender wage gap is 0.297 log points. Actual and predicted experience yield very similar estimates of the contribution of endowments at 0.226 and 0.211 log points, respectively (columns 1 and 2). On the other hand, potential experience yields a much lower estimate of the endowment effect (0.167 log points). Accordingly, actual and predicted experience yield very similar estimates of discrimination at around 0.072 and 0.086 log points (columns 4 and 5). The use of potential experience implies an estimate of discrimination or of the unexplained gap that is very much higher (column 6) than that obtained from using actual or predicted experience.

Table 6 reports the decomposition results for the IPUMS sample of workers aged 18 to 55 who are heads of households. The unadjusted gender wage gap for this sample is 0.330 log points. Again, actual work experience is missing from the IPUMS data set. The decomposition results follow the same pattern as found in the other samples. Endowment effects are larger and discrimination effects are correspondingly smaller when using predicted rather than potential experience.

Because the NLSY79 and PSID data sets permit the construction of actual work experience, Tables 3 and 5 are able to document the gender wage gap decomposition specification bias arising from the use of predicted or potential experience in lieu of actual experience. We now seek to examine how much of the biases in the endowment and discrimination components of the decomposition can be attributed to mean differences in the experience measures and differences in the estimated parameters. By construction, the total specification biases in the endowment and discrimination components are equal in absolute value and of opposite signs.

Table 7 reports the decomposition of the specification biases for the NLSY79 sample. The total specification bias is smaller when using predicted rather than potential experience (0.025 vs 0.064). Relative to actual experience, the use of predicted experience reduces the estimated endowment effect (explained gap) by 0.025 log points. Virtually all of this bias stems from the bias in the estimated parameters of the wage model (−0.022 log points), leaving no effects from mean differences in the experience measures. The case for potential experience is exactly opposite. Most of the estimated endowment effect bias of 0.064 log points stems from mean differences between potential and actual experience (0.052 log points). When considering the 0.025 log

Table 7 Decomposition of the experience specification bias: NLSY79 white males and females, ages 25–33

	Endowment effects						Discrimination					
	Total		Differences in estimated parameters		Differences in experience measures		Total		Differences in estimated parameters		Differences in experience measures	
	Predicted (1)	Potential (2)	Predicted (3)	Potential (4)	Predicted (5)	Potential (6)	Predicted (7)	Potential (8)	Predicted (9)	Potential (10)	Predicted (11)	Potential (12)
Schooling	0.000	0.012	0.000	0.012	0.000	0.000	−0.049	−0.281	−0.049	−0.281	0.000	0.000
Experience	0.025	0.163	0.025	0.108	0.000	0.055	0.599	−0.111	0.599	−0.023	0.000	−0.089
Experience^2	−0.055	−0.106	−0.054	−0.102	−0.002	−0.004	−0.364	−0.070	−0.352	−0.152	−0.012	0.082
Other variables	0.005	−0.006	0.005	−0.006	0.000	0.000	−0.160	0.398	−0.160	0.398	0.000	0.000
Total	−0.025	0.064	−0.023	0.012	−0.002	0.052	0.025	−0.064	0.037	−0.057	−0.012	−0.006

Actual and predicted experience are constructed from hours worked per week for the NLSY79. Components may not add to totals due to rounding. Source of data: 1990 survey of the NLSY79

Table 8 Decomposition of the experience specification bias: PSID white males and females, heads of household, ages 18–55

	Endowment effects						Discrimination					
	Total		Differences in estimated parameters		Differences in experience measures		Total		Differences in estimated parameters		Differences in experience measures	
	Predicted (1)	Potential (2)	Predicted (3)	Potential (4)	Predicted (5)	Potential (6)	Predicted (7)	Potential (8)	Predicted (9)	Potential (10)	Predicted (11)	Potential (12)
Schooling	0.000	−0.001	0.000	−0.001	0.000	0.000	0.011	−0.035	0.011	−0.035	0.000	0.000
Experience	−0.001	0.186	−0.001	0.013	0.000	0.172	0.032	−0.223	0.032	−0.217	0.000	−0.005
Experience^2	0.001	−0.132	−0.015	−0.028	0.016	−0.104	−0.047	0.078	−0.046	0.068	0.000	0.009
Other variables	0.015	0.005	0.015	0.005	0.000	0.000	−0.011	0.122	−0.011	0.122	0.000	0.000
Total	0.014	0.059	−0.002	−0.010	0.016	0.068	−0.014	−0.059	−0.014	−0.063	0.000	0.004

Actual and predicted experience are constructed from total annual work hours as reported in the family files for the PSID. Components may not add to totals due to rounding. Source of data: 1990 survey of the PSID point bias in estimating the unexplained or discriminatory wage gap with predicted experience, we see that bias in parameter estimation contributes 0.037 log points to the bias. This is partly offset by the modest −0.012 log point bias arising from mean differences in the predicted and potential experience measures. By contrast, most of the −0.064 underestimate of the unexplained gap when using potential experience is accounted for by negative bias in the estimated parameters (−0.057 log points), leaving little bias from mean differences in potential and actual experience.

The decomposition specification biases for the PSID sample are reported in Table 8. The overall decomposition bias when using predicted experience is rather modest at 0.014 log points. Most or all of the bias stems from mean differences in experience measures in the case of the endowment effect specification bias and from estimated parameter differences in the case of the unexplained gap specification bias. Specification bias is more substantial at 0.059 log points when potential experience is used. Mean differences in potential experience and actual experience increase the endowment effect bias by 0.068 log points which is partially offset by the −0.01 log point bias reduction from differences in parameter estimates. Virtually all of the −0.059 log point underestimate of discrimination when using potential experience is generated by differences in the estimated parameters.

6 Concluding remarks

This paper employs data from the NLSY79, the PSID, and the IPUMS in investigating the bias inherent in human capital models that utilize potential, as opposed to actual, work experience measures. We address the issue in a broader sense by not confining the analysis to women, where the problems of error in potential work experience are well known and broadly accepted, but by expanding the discussion to include males who also experience lapses of employment. The NLSY79 and the PSID allow us to measure actual work experience and to construct predicted work experience measures that we apply to the IPUMS—a data set lacking individual work histories.

A series of log wage regressions are run utilizing the various measures of work experience—actual, predicted, and potential (as proxied by age – schooling – 6). On the basis of these findings, we conclude that specification error in work experience not only biases its coefficient but also that of schooling as well; potential work experience overestimates the effects of schooling and underestimates the rates of return to labor market experience. Based on the figures in Table 2, the bias on the estimated coefficient of schooling is larger when potential work experience is used instead of our predicted work experience measure in human capital models. The overall mean absolute percent differences (linear and quadratic) reveal a substantial reduction in the bias on the estimated coefficients on experience when our predicted measures are used. The only exception to this is for the NLSY79 females.

Although we find that the average discrepancy in potential vs actual work experience is positive, there are cases in which potential work experience actually *understates* actual work histories due to overtime work, multiple-job holding, and our imposition of a FTE status. The discontinuous nature of female work patterns in particular, coupled with unemployment spells that affect both genders, and different labor market experiences (e.g., employment status) provide evidence that demands better proxies than potential work experience when actual work histories are lacking. Tables 3–8 support this claim. The wage decompositions suggest that more of the male–female wage gap is explained by the difference in average qualifications when our predicted measures of work experience are used in lieu of potential measures. In fact, the use of predicted experience as opposed to potential experience increases the explained proportion of the gender wage gap anywhere from 9 percentage points for the IPUMS heads of household to 41 percentage points for the NLSY79. The use of actual experience as opposed to potential experience increases the explained proportion of the gender wage gap by 20 percentage points for the PSID heads of household and 29 percentage points for the NSLY79. The decomposition biases associated with misspecification of experience were always much smaller when using predicted experience as compared with potential experience. Mean differences between potential and actual experience accounted for virtually all of the endowment effect specification bias, while differences in the estimated parameters accounted for virtually all of the discrimination specification bias.

Instrumental variables (IV) is the traditional approach taken to correct classical measurement error. We have conducted simple statistical tests that demonstrate that the apparent measurement error in work experience is nonclassical in the sense that the mean is not zero and there is correlation between the measurement error and actual work experience. These tests are not reported here. Matters are further complicated because experience enters log wage equations linearly and quadratically as well. While Kelejian (1971) offers an alternative estimation strategy (i.e., nonlinear 2SLS), identifying a set of unique instruments is not a trivial task. Basically, the problem with instrumenting potential experience (and its square) is that this assumes that the correct model specification requires potential experience but that, in a given data set, potential experience is measured with error. Thus, instrumenting potential experience would not solve the model misspecification problem. IV applied to potential experience produces biased wage decomposition components in both coefficient estimates and in the predicted mean work experience.

Acknowledgements We are grateful for comments received at the 2005 SOLE/EALE Annual Meeting; at the 2005 American Economic Association Annual Meeting; at the 2004 Western Economic Association International 79th Annual Conference; and from seminar participants at the University of Aberdeen, the University of Oxford, the University of Mississippi, the University of Miami, and the University of Arizona. We would also like to thank Tim Barmby, Solomon Polachek, and three anonymous referees.

Appendix

Table 9 Descriptive statistics for white males

	Ages 25–33				Heads of household, ages 18–55			
	NLSY79 ($N = 2789$)		IPUMS ($N = 3540$)		PSID ($N = 2892$)		IPUMS ($N = 9098$)	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Individual characteristics								
Age	28.635	2.270	29.062	2.564	36.094	8.386	37.958	9.015
Schooling	12.920	2.587	13.005	2.550	12.965	2.870	13.346	2.812
Married	0.576	0.494	0.590	0.492	0.801	0.399	0.791	0.407
Total annual income from wages and salary	24447.730	17490.510	24008.880	17352.150	30792.400	29136.030	33389.110	27629.730
Hourly wage	11.354	11.669	13.239	49.602	13.839	13.080	16.626	39.445
Industry								
Agriculture, forestry, fishing	0.038	0.191	0.031	0.173	0.029	0.169	0.026	0.158
Mining	0.014	0.119	0.012	0.107	0.012	0.109	0.015	0.120
Construction	0.134	0.341	0.136	0.343	0.102	0.303	0.116	0.321
Manufacturing	0.256	0.437	0.243	0.429	0.268	0.443	0.260	0.439
Transportation, communications, other public utilities	0.086	0.281	0.073	0.260	0.103	0.304	0.086	0.281
Wholesale and retail trade	0.183	0.387	0.204	0.403	0.173	0.379	0.170	0.376
Finance, insurance, real estate	0.044	0.206	0.052	0.221	0.043	0.203	0.053	0.223
Business, repair services	0.077	0.266	0.064	0.244	0.061	0.239	0.056	0.229
Personal services	0.019	0.135	0.013	0.112	0.016	0.126	0.013	0.111
Entertainment, recreation services	0.014	0.117	0.018	0.132	0.009	0.094	0.014	0.119
Professional, related services	0.090	0.286	0.100	0.300	0.107	0.309	0.120	0.325
Public administration	0.044	0.205	0.056	0.230	0.076	0.266	0.072	0.258
Occupation								
Professional, technical, and kindred	0.153	0.360	0.170	0.376	0.190	0.393	0.204	0.403
Managers, officials, and proprietors	0.143	0.351	0.135	0.342	0.166	0.372	0.181	0.385
Sales workers	0.055	0.228	0.062	0.242	0.053	0.224	0.061	0.240
Clerical and kindred	0.066	0.248	0.062	0.240	0.051	0.220	0.057	0.231
Craftsmen, foremen, and kindred	0.226	0.418	0.218	0.413	0.214	0.410	0.202	0.402
Operatives and kindred	0.112	0.315	0.124	0.329	0.110	0.312	0.109	0.312
Transportation, equipment operators	0.069	0.254	0.050	0.217	0.066	0.248	0.050	0.218

Table 9 (continued)

	Ages 25–33				Heads of household, ages 18–55			
	NLSY79 ($N = 2789$)		IPUMS ($N = 3540$)		PSID ($N = 2892$)		IPUMS ($N = 9098$)	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Laborers, except farm	0.080	0.271	0.085	0.279	0.054	0.226	0.052	0.222
Farmers and farm managers	0.003	0.057	0.004	0.060	0.006	0.079	0.005	0.070
Farm laborers and foreman	0.013	0.111	0.011	0.104	0.009	0.094	0.009	0.097
Service workers, except private household	0.080	0.272	0.080	0.271	0.064	0.245	0.069	0.253
Private household	0.001	0.027	0.000	0.017	0.001	0.026	0.000	0.015
Region								
Northeast	0.149	0.356	0.217	0.412	0.177	0.381	0.205	0.403
North central	0.284	0.451	0.246	0.431	0.226	0.419	0.259	0.438
South	0.338	0.473	0.302	0.459	0.365	0.482	0.326	0.469
West	0.228	0.420	0.236	0.424	0.232	0.422	0.211	0.408
SMSA								
Not in SMSA	0.244	0.430	0.251	0.434	–	–	–	–
SMSA, not central city	0.341	0.474	0.319	0.466	–	–	–	–
SMSA, central city unknown	0.304	0.460	0.273	0.446	–	–	–	–
SMSA, central city	0.111	0.314	0.157	0.364	–	–	–	–

Source of data: 1990 survey of the NLSY79, the PSID, and the IPUMS

Table 10 Descriptive statistics for white females

	Ages 25–33				Heads of household, ages 18–55			
	NLSY79 (N = 2386)		IPUMS (N = 3062)		PSID (N = 516)		IPUMS (N = 2579)	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Individual characteristics								
Age	28.749	2.262	28.955	2.570	36.674	9.611	37.150	9.384
Schooling	13.326	2.295	13.345	2.315	12.915	2.883	13.541	2.534
Married	0.609	0.488	0.637	0.481	0.006	0.076	0.195	0.397
Number of children	1.067	1.106	1.074	1.147	1.000	0.000	1.435	1.535
Total annual income from wages and salary	16345.260	12365.200	15745.160	12760.770	19713.420	13152.670	20040.240	16000.510
Hourly wage	9.566	14.520	9.700	10.497	9.898	5.930	11.348	14.106
Industry								
Agriculture, forestry, fishing	0.010	0.100	0.012	0.108	0.010	0.098	0.007	0.086
Mining	0.003	0.054	0.001	0.036	0.004	0.062	0.004	0.065
Construction	0.017	0.128	0.016	0.127	0.008	0.088	0.015	0.121
Manufacturing	0.157	0.364	0.153	0.360	0.178	0.383	0.161	0.367
Transportation, communications, other public utilities	0.047	0.211	0.040	0.196	0.041	0.198	0.039	0.194
Wholesale and retail trade	0.212	0.409	0.213	0.409	0.174	0.380	0.205	0.404
Finance, insurance, real estate	0.091	0.288	0.109	0.312	0.081	0.274	0.080	0.271
Business, repair services	0.060	0.238	0.056	0.230	0.033	0.179	0.061	0.239
Personal services	0.053	0.225	0.045	0.207	0.072	0.258	0.035	0.185
Entertainment, recreation services	0.011	0.104	0.018	0.132	0.012	0.107	0.014	0.116
Professional, related services	0.302	0.459	0.296	0.456	0.326	0.469	0.326	0.469
Public administration	0.037	0.189	0.042	0.200	0.062	0.241		
Occupation								
Professional, technical, and kindred	0.219	0.414	0.247	0.431	0.231	0.422	0.259	0.438
Managers, officials, and proprietors	0.112	0.316	0.106	0.308	0.118	0.323	0.132	0.339
Sales workers	0.051	0.220	0.060	0.237	0.043	0.202	0.060	0.237
Clerical and kindred	0.318	0.466	0.321	0.467	0.256	0.437	0.292	0.455
Craftsmen, foremen, and kindred	0.021	0.145	0.018	0.134	0.031	0.174	0.018	0.134
Operatives and kindred	0.087	0.282	0.077	0.266	0.105	0.306	0.073	0.261
Transportation, equipment operators	0.005	0.068	0.008	0.090	0.012	0.107	0.012	0.111

Table 10 (continued)

	Ages 25–33				Heads of household, ages 18–55			
	NLSY79 ($N = 2386$)		IPUMS ($N = 3062$)		PSID ($N = 516$)		IPUMS ($N = 2579$)	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Laborers, except farm	0.012	0.110	0.016	0.127	0.012	0.107	0.012	0.109
Farmers and farm managers	0.000	0.020	0.002	0.040	–	–	0.000	0.000
Farm laborers and foreman	0.002	0.046	0.004	0.060	0.006	0.076	0.003	0.056
Service workers, except private household	0.159	0.366	0.136	0.342	0.167	0.373	0.134	0.340
Private household	0.013	0.113	0.006	0.076	0.016	0.124	0.004	0.065
Region								
Northeast	0.141	0.348	0.216	0.411	0.203	0.403	0.220	0.414
North central	0.273	0.446	0.249	0.433	0.180	0.385	0.231	0.422
South	0.368	0.482	0.310	0.463	0.380	0.486	0.314	0.464
West	0.218	0.413	0.225	0.417	0.236	0.425	0.234	0.424
SMSA								
Not in SMSA	0.235	0.251	0.434	–	–	–	–	–
SMSA, not central city	0.344	0.475	0.332	0.471	–	–	–	–
SMSA, central city unknown	0.329	0.470	0.279	0.448	–	–	–	–
SMSA, central city	0.092	0.289	0.138	0.345	–	–	–	–

Source of data: 1990 survey of the NLSY79, the PSID, and the IPUMS

Table 11 Log wage regression for white males, ages 25–33

Variables	NLSY79			IPUMS	
	(1)	(2)	(3)	(4)	(5)
Schooling	0.080^{a} (0.042×10^{-1})	0.081^{a} (0.043×10^{-1})	0.110^{a} (0.059×10^{-1})	0.072^{a} (0.040×10^{-1})	0.098^{a} (0.053×10^{-1})
Actual experience	0.103^{a} (0.013)	–	–	–	–
(Actual experience)^2	-0.035×10^{-1a} (0.067×10^{-2})	–	–	–	–
Predicted experience	–	0.088^{b} (0.043)	–	0.116^{a} (0.033)	–
(Predicted experience)^2	–	-0.017×10^{-1} (0.022×10^{-1})	–	0.035×10^{-1b} (0.017×10^{-1})	–
Potential experience	–	–	0.039^{a} (0.014)	–	0.019^{c} (0.011)
(Potential experience)^2	–	–	-0.015×10^{-2} (0.067×10^{-2})	–	0.062×10^{-2} (0.050×10^{-2})
Married	0.116^{a} (0.021)	0.081^{a} (0.023)	0.149^{a} (0.021)	0.068^{a} (0.021)	0.130^{a} (0.019)
Agriculture, forestry, fishing	-0.321^{a} (0.070)	-0.359^{a} (0.071)	-0.310^{a} (0.071)	-0.518^{a} (0.066)	-0.481^{a} (0.066)
Mining	0.144 (0.095)	0.122 (0.097)	0.152 (0.097)	−0.053 (0.094)	−0.028 (0.094)
Construction	0.062 (0.055)	0.049 (0.056)	0.086 (0.056)	-0.099^{b} (0.046)	−0.068 (0.046)
Manufacturing	0.012 (0.051)	−0.003 (0.052)	0.039 (0.052)	−0.055 (0.043)	−0.022 (0.043)
Transportation, communications, other public utilities	0.015 (0.058)	−0.014 (0.059)	0.054 (0.059)	−0.078 (0.052)	−0.019 (0.051)
Wholesale and retail trade	-0.185^{a} (0.053)	-0.210^{a} (0.054)	-0.148^{a} (0.054)	-0.253^{a} (0.044)	-0.204^{a} (0.044)
Finance, insurance, real estate	0.126^{c} (0.067)	0.119^{c} (0.067)	0.139^{b} (0.068)	-0.111^{b} 0.056	−0.090 (0.056)
Business, repair services	−0.092 (0.059)	-0.112^{c} (0.060)	−0.076 (0.060)	-0.153^{a} (0.053)	-0.124^{b} (0.053)

Table 11 (continued)

Variables	NLSY79			IPUMS	
	(1)	(2)	(3)	(4)	(5)
Personal services	-0.251^{a} (0.087)	-0.256^{a} (0.088)	-0.248^{a} (0.088)	-0.198^{b} (0.090)	-0.190^{b} (0.090)
Entertainment, recreation services	-0.291^{a} (0.096)	-0.298^{a} (0.097)	-0.304^{a} (0.097)	-0.186^{b} (0.079)	-0.189^{b} (0.079)
Professional, related services	-0.134^{b} (0.058)	-0.128^{b} (0.059)	-0.143^{b} (0.059)	-0.218^{a} (0.049)	-0.246^{a} (0.049)
Northeast	0.053 (0.034)	0.051 (0.034)	0.064^{c} (0.034)	0.068^{b} (0.027)	0.075^{a} (0.027)
North central	-0.051^{c} (0.028)	-0.049^{c} (0.028)	-0.051^{c} (0.028)	-0.023×10^{-1} (0.027)	-0.029×10^{-1} (0.027)
South	-0.113^{a} (0.027)	-0.114^{a} (0.027)	-0.102^{a} (0.027)	-0.060^{b} (0.025)	-0.052^{b} (0.026)
SMSA, not central city	0.159^{a} (0.027)	0.151^{a} (0.027)	0.173^{a} (0.027)	0.236^{a} (0.026)	0.253^{a} (0.026)
SMSA, central city unknown	0.095^{a} (0.028)	0.098^{a} (0.028)	0.102^{a} (0.028)	0.141^{a} (0.026)	0.146^{a} (0.026)
SMSA, central city	0.106^{a} (0.037)	0.115^{a} (0.038)	0.088^{b} (0.038)	0.129^{a} (0.031)	0.100^{a} (0.031)
Constant	0.528^{a} (0.099)	0.512^{b} (0.217)	0.335^{b} (0.138)	0.582^{a} (0.172)	0.650^{a} (0.112)
R^2	0.243	0.224	0.219	0.200	0.196
adj. R^2	0.237	0.218	0.213	0.195	0.191
N	2789	2789	2789	3540	3540

Source of data: The data in parentheses are standard error. Actual and predicted experience are constructed from hours worked per week and potential experience= age−schooling−6

[a]Significant at the 1% level

[b]Significant at the 5% level

[c]Significant at the 10% level

Table 12 Log wage regression for white males, heads of household, ages 18–55

Variables	PSID			IPUMS	
	(1)	(2)	(3)	(4)	(5)
Schooling	0.076^{a} (0.038×10^{-1})	0.076^{a} (0.038×10^{-1})	0.090^{a} (0.041×10^{-1})	0.083^{a} (0.024×10^{-1})	0.097^{a} (0.025×10^{-1})
Actual experience	0.044^{a} (0.037×10^{-1})	–	–	–	–
(Actual experience)^2	-0.082×10^{-2a} (0.089×10^{-3})	–	–	–	–
Predicted experience	–	0.044^{a} (0.039×10^{-1})	–	0.046^{a} (0.025×10^{-1})	–
(Predicted experience)^2	–	-0.069×10^{-2a} (0.082×10^{-3})	–	-0.074×10^{-2a} (0.056×10^{-3})	–
Potential experience	–	–	0.040^{a} (0.043×10^{-1})	–	0.041^{b} (0.027×10^{-1})
(Potential experience)^2	–	–	-0.059×10^{-2a} (0.011×10^{-2})	–	-0.059×10^{-2a} (0.066×10^{-3})
Married	0.179^{a} (0.026)	0.158^{a} (0.026)	0.165^{a} (0.026)	0.093^{a}(0.015)	0.102^{a} (0.015)
Agriculture. forestry, fishing	-0.480^{a} (0.070)	-0.469^{a} (0.070)	-0.475^{a} (0.069)	-0.481^{a} (0.045)	-0.483^{a} (0.044)
Mining	0.131 (0.098)	0.139 (0.098)	0.150 (0.098)	0.168^{a} (0.055)	0.194^{a} (0.055)
Construction	-0.037×10^{-1} (0.049)	0.015 (0.049)	-0.086×10^{-1} (0.048)	-0.058×10^{-1} (0.029)	-0.078×10^{-1} (0.029)
Manufacturing	0.054 (0.042)	0.059 (0.041)	0.058 (0.041)	0.047×10^{-1} (0.026)	0.014 (0.026)
Transportation, communications, other public utilities	0.098^{b} (0.048)	0.095^{b} (0.048)	0.095^{b} (0.048)	0.035 (0.031)	0.044 (0.031)

Table 12 (continued)

Variables	PSID			IPUMS	
	(1)	(2)	(3)	(4)	(5)
Wholesale and retail trade	−0.139[a] (0.044)	−0.138[a] (0.044)	−0.135[a] (0.043)	−0.136[a] (0.027)	−0.123[a] (0.027)
Finance, insurance, real estate	−0.045 (0.060)	−0.045 (0.060)	−0.048 (0.060)	0.057 (0.035)	0.072[b] (0.035)
Business, repair services	−0.080 (0.055)	−0.062 (0.054)	−0.066 (0.055)	−0.067[c] (0.035)	−0.069[b] (0.035)
Personal services	−0.264[a] (0.087)	−0.272[a] (0.087)	−0.289[a] (0.086)	−0.186[a] (0.059)	−0.202[a] (0.059)
Entertainment, recreation services	0.056 (0.112)	0.050 (0.111)	0.031 (0.111)	−0.072 (0.156)	−0.104[c] (0.156)
Professional, related services	−0.026 (0.048)	−0.021 (0.048)	−0.045 (0.047)	−0.117[a] (0.029)	−0.134[a] (0.029)
Northeast	0.116[a] (0.032)	0.115[a] (0.032)	0.108[a] (0.032)	0.051[a] (0.019)	0.051[a] (0.019)
North central	−0.049 (0.030)	−0.049[c] (0.030)	−0.057[c] (0.030)	−0.064[a] (0.018)	−0.066[a] (0.018)
South	−0.166[a] (0.027)	−0.168[a] (0.027)	−0.171[a] (0.026)	−0.107[a] (0.017)	−0.110[a] (0.017)
Constant	0.951[a] (0.073)	0.928[a] (0.074)	0.738[a] (0.078)	0.959[a] (0.047)	0.726[a] (0.049)
R^2	0.277	0.280	0.288	0.241	0.246
adj. R^2	0.272	0.276	0.288	0.240	0.245
N	2892	2892	2892	9098	9098

Source of data: Data in parentheses are standard error. Actual and predicted experience are constructed from total annual work hours as reported in the family files and potential experience=age−schooling−6

[a]Significant at the 1% level

[b]Significant at the 5% level

[c]Significant at the 10% level

Table 13 Log wage regression for white females, ages 25–33

Variables	NLSY79			IPUMS	
	(1)	(2)	(3)	(4)	(5)
Schooling	0.083[a] (0.055×10^{-1})	0.081[a] (0.058×10^{-1})	0.093[a] (0.075×10^{-1})	0.077[a] (0.047×10^{-1})	0.106[a] (0.060×10^{-1})
Actual experience	0.057[a] (0.019)	–	–	–	–
(Actual experience)^2	0.011×10^{-1} (0.012×10^{-1})	–	–	–	– –
Predicted experience	–	0.122[b](0.050)	–	0.111[a](0.038)	–
(Predicted experience)^2	–	-0.053×10^{-1}[c] (0.032×10^{-1})	–	-0.040×10^{-1}[c] 0.023×10^{-1}	–
Potential experience	–	–	0.090×10^{-1} (0.017)	–	0.049×10^{-1} (0.012)
(Potential experience)^2	–	–	-0.069×10^{-2} (0.086×10^{-1})	–	0.075×10^{-2} (0.058×10^{-2})
Married	0.026 (0.024)	0.023 (0.024)	0.025 (0.024)	0.018 (0.021)	0.014 (0.021)
Agriculture, forestry, fishing	−0.328[b] (0.130)	−0.324[b] (0.132)	−0.375[a] (0.133)	−0.238[b] (0.103)	−0.333[a] (0.104)
Mining	0.076 (0.222)	0.077 (0.225)	0.049 (0.226)	−0.058 (0.276)	−0.024 (0.277)
Construction	0.100 (0.108)	0.117 (0.110)	0.047 (0.110)	−0.084 (0.091)	−0.156[c] (0.091)
Manufacturing	-0.066×10^{-1} (0.067)	-0.046×10^{-1} (0.068)	−0.014 (0.069)	−0.014[b] (0.055)	−0.120[b] (0.055)
Transportation, communications, other public utilities	0.066 (0.081)	0.066 (0.082)	0.056 (0.082)	0.031 (0.069)	0.026 (0.069)
Wholesale and retail trade	−0.201[a] (0.066)	−0.197[a] (0.067)	−0.232[a] (0.067)	−0.277[a] (0.053)	−0.317[a] (0.053)
Finance, insurance, real estate	−0.072 (0.071)	−0.073 (0.072)	−0.069 (0.073)	−0.085 (0.057)	−0.071 (0.057)

Table 13 (continued)

Variables	NLSY79			IPUMS	
	(1)	(2)	(3)	(4)	(5)
Business, repair services	−0.005 (0.076)	0.002 (0.078)	−0.048 (0.078)	−0.084 (0.064)	−0.122[c] (0.064)
Personal services	−0.444[a] (0.079)	−0.442[a] (0.080)	−0.498[a] (0.080)	−0.359[a] (0.067)	−0.396[a] (0.067)
Entertainment, recreation services	−0.303[b] (0.126)	−0.299[b] (0.128)	−0.327[b] (0.128)	−0.222[b] (0.088)	−0.228[b] (0.089)
Professional, related services	−0.075 (0.064)	−0.072 (0.065)	−0.111[c] (0.065)	−0.094[c] (0.052)	−0.145[c] (0.052)
Northeast	-0.046×10^{-1} (0.040)	-0.067×10^{-1} (0.041)	0.011 (0.041)	0.073×10^{-1} (0.030)	0.027 (0.030)
North central	−0.087[a] (0.033)	−0.086[b] (0.034)	−0.076[b] (0.034)	−0.076[c] (0.029)	−0.061[b] (0.030)
South	−0.127[a] (0.032)	−0.130[a] (0.032)	−0.108[a] (0.032)	−0.108[a] (0.028)	−0.080[a] (0.028)
SMSA, not central city	0.148[a] (0.032)	0.140[a] (0.033)	0.168[a] (0.032)	0.218[a] (0.028)	0.255[a] (0.027)
SMSA, central city unknown	0.170[a] (0.032)	0.163[a] (0.033)	0.192[a] (0.033)	0.129[a] (0.028)	0.168[a] (0.027)
SMSA, central city	0.276[a] (0.046)	0.268[a] (0.047)	0.291[a] (0.047)	0.297[a] (0.034)	0.319[a] (0.035)
Constant	0.568[a] (0.113)	0.365[c] (0.207)	0.818[a] (0.171)	0.471[a] (0.160)	0.554[a] (0.130)
R^2	0.226	0.204	0.196	0.234	0.225
adj. R^2	0.220	0.197	0.189	0.229	0.220
N	2386	2386	2386	3062	3062

Source of data: Data in parentheses are standard error. Actual and predicted experience are constructed from hours worked per week and potential experience=age−schooling−6

[a]Significant at the 1% level

[b]Significant at the 10% level

[c]Significant at the 5% level

Table 14 Log wage regression for white females, heads of household, ages 18–55

Variables	PSID (1)	PSID (2)	PSID (3)	IPUMS (4)	IPUMS (5)
Schooling	0.076^{a}	0.077^{a}	0.088^{a}	0.096^{a}	0.107^{a}
	(0.079×10^{-1})	(0.080×10^{-1})	(0.092×10^{-1})	(0.047×10^{-1})	(0.051×10^{-1})
Actual experience	0.043^{a}	–	–	–	–
	(0.084×10^{-1})				
$(\text{Actual experience})^2$	0.077×10^{-2a}	–	–	–	–
	(0.025×10^{-2})				
Predicted experience	–	0.046^{a}	–	0.030^{a}	–
		(0.010)		(0.049×10^{-1})	
$(\text{Predicted experience})^2$	–	-0.087×10^{-2a}	–	-0.052×10^{-2a}	–
		(0.029×10^{-2})		0.013×10^{-2}	
Potential experience	–	–	0.027^{a}	–	0.026^{a}
			(0.074×10^{-1})		(0.042×10^{-1})
$(\text{Potential experience})^2$	–	–	-0.038×10^{-2b}	–	0.045×10^{-2a}
			(0.019×10^{-2})		(0.011×10^{-2})
Married	−0.316 (0.266)	−0.310 (0.271)	−0.348 (0.272)	0.012 (0.029)	−0.037 (0.08)
Agriculture, forestry, fishing	-0.441^{b} (0.224)	-0.439^{c} (0.228)	-0.430^{c} (0.229)	-0.403^{a} (0.140)	-0.375^{a} (0.138)
Mining	0.037×10^{-1}	0.056×10^{-2}	−0.088 (0.339)	0.197 (0.177)	0.185 (0.176)
	(0.332)	(0.338)			
Construction	0.016 (0.243)	0.011 (0.248)	0.038×10^{-1} (0.249)	−0.072 (0.103)	−0.058 (0.103)
Manufacturing	-0.216^{b} (0.096)	-0.204^{b} (0.098)	-0.255^{a} (0.097)	−0.076 (0.056)	-0.113^{b} (0.056)
Transportation, communications, other public utilities	0.039 (0.128)	0.077 (0.131)	0.032 (0.131)	0.068 (0.074)	0.055 (0.074)

Table 14 (continued)

Variables	PSID			IPUMS	
	(1)	(2)	(3)	(4)	(5)
Wholesale and retail trade	−0.492[a] (0.097)	−0.480[a] (0.098)	−0.546[a] (0.097)	−0.290[a] (0.055)	−0.328[a] (0.054)
Finance, insurance, real estate	−0.046 (0.107)	−0.028 (0.110)	−0.053 (0.109)	−0.019 (0.062)	−0.013 (0.063)
Business, repair services	−0.079 (0.138)	−0.066 (0.140)	−0.099 (0.0140)	−0.098 (0.066)	−0.122[b] (0.066)
Personal services	−0.554[a] (0.114)	−0.527[a] (0.117)	−0.665[a] (0.115)	−0.333[a] (0.078)	−0.416[a] (0.077)
Entertainment, recreation services	−0.631[a] (0.205)	−0.615[a] (0.211)	−0.683[a] (0.210)	−0.099 (0.107)	−0.138 (0.107)
Professional, related services	−0.207[b] (0.089)	−0.198[b] (0.090)	−0.243[a] (0.090)	−0.148[a] (0.052)	−0.178[a] (0.052)
Northeast	0.094 (0.061)	0.099 (0.063)	0.078 (0.063)	0.089[a] (0.033)	0.077[b] (0.033)
North central	0.016 (0.063)	0.073×10^{-1} (0.065)	−0.012 (0.065)	−0.023 (0.033)	−0.040 (0.033)
South	−0.075 (0.053)	−0.078 (0.055)	−0.062 (0.055)	−0.085[a] (0.030)	−0.076[b] (0.030)
Constant	1.039[a] (0.156)	0.995[a] (0.168)	0.977[a] (0.168)	0.774[a] (0.094)	0.668[a] (0.098)
R^2	0.420	0.400	0.395	0.242	0.244
adj. R^2	0.400	0.378	0.373	0.237	0.239
N	516	516	516	2579	2579

Source of data: Data in parentheses are standard error. Actual and predicted experience are constructed from total annual work hours as reported in the family files and potential experience=age-schooling-6

[a]Significant at the 1% level

[b]Significant at the 5% level

[c]Significant at the 10% level

Table 15 First-stage log actual experience regression

Variables	NLSY79	PSID	NLSY79	PSID
	White males		White females	
	Ages 25–33	Heads of household, ages 18–55	Ages 25–33	Heads of household, ages 18–55
	(1)	(2)	(3)	(4)
Schooling	−0.017[a] (0.033×10^{-1})	-0.097×10^{-2} (0.030×10^{-1})	-0.059×10^{-1} (0.048×10^{-1})	0.052×10^{-1} (0.010)
Married	0.122[a] (0.014)	0.096[a] (0.018)	0.074[a] (0.018)	−0.208[a] (0.295)
Age	0.074[a] (0.030×10^{-1})	0.065[a] (0.085×10^{-2})	0.069[a] (0.039×10^{-1})	0.064[a] (0.027×10^{-1})
Number of children	–	–	−0.152[a] (0.089×10^{-1})	−0.103[a] (0.018)
Agriculture, forestry, fishing	0.082 (0.053)	−0.155[b] (0.067)	−0.158 (0.111)	0.528 (0.373)
Mining	0.055 (0.065)	0.078 (0.070)	−0.022 (0.163)	−0.048 (0.368)
Construction	0.047 (0.039)	−0.091[b] (0.039)	−0.210[a] (0.080)	−0.027 (0.269)
Manufacturing	0.076[b] (0.036)	−0.000 (0.034)	0.042 (0.052)	−0.260[b] (0.116)
Transportation, communications, other public utilities	0.120[a] (0.040)	0.021 (0.037)	−0.026 (0.060)	−0.121 (0.144)
Wholesale and retail trade	0.071[c] (0.036)	−0.025 (0.035)	−0.061 (0.049)	−0.270[c] (0.108)
Finance, insurance, real estate	0.004 (0.045)	−0.047 (0.045)	0.008 (0.053)	0.008 (0.123)
Business, repair services	0.049 (0.040)	−0.066 (0.041)	−0.113[b] (0.056)	−0.201 (0.154)
Personal services	−0.016 (0.059)	−0.125[c] (0.060)	−0.022 (0.062)	−0.590[a] (0.130)
Entertainment, recreation services	−0.010 (0.064)	−0.169[b] (0.078)	−0.005 (0.093)	−0.341 (0.227)
Professional, related services	−0.014 (0.039)	−0.103[a] (0.036)	−0.064 (0.048)	−0.188[c] (0.100)
Professional, technical, and kindred	−0.089 (0.252)	−0.126[b] (0.062)	0.084 (0.088)	0.001 (0.175)
Managers, officials, and proprietors	0.088 (0.252)	0.047 (0.062)	0.184[b] (0.088)	0.097 (0.179)
Sales workers	0.028 (0.253)	−0.042 (0.068)	0.022 (0.094)	−0.059 (0.203)
Clerical and kindred	−0.078 (0.252)	−0.124[c] (0.065)	0.152[c] (0.086)	−0.007 (0.171)
Craftsmen, foremen, and kindred	0.007 (0.252)	0.064 (0.062)	0.103 (0.104)	0.227 (0.211)

Table 15 (continued)

Variables	NLSY79	PSID	NLSY79	PSID
	White males		White females	
	Ages 25–33	Heads of household, ages 18–55	Ages 25–33	Heads of household, ages 18–55
	(1)	(2)	(3)	(4)
Operatives and kindred	−0.083 (0.252)	−0.065 (0.064)	0.011 (0.091)	0.078 (0.192)
Transportation, equipment operators	−0.019 (0.252)	−0.053 (0.066)	−0.001 (0.153)	0.340 (0.268)
Laborers, except farm	−0.148 (0.252)	−0.162[b] (0.067)	0.011 (0.117)	0.129 (0.267)
Farmers and farm managers	0.203 (0.278)	0.387[a] (0.121)	−0.031 (0.436)	–
Farm laborers and foreman	−0.030 (0.261)	0.078 (0.111)	0.133 (0.228)	−0.519 (0.494)
Service workers, except private household	−0.041 (0.251)	−0.072 (0.062)	−0.022 (0.084)	−0.007 (0.168)
Northeast	0.019 (0.022)	−0.017 (0.022)	0.027 (0.030)	−0.120[c] (0.068)
North central	−0.002 (0.019)	−0.023 (0.020)	0.041[c] (0.025)	−0.111 (0.071)
South	0.017 (0.018)	−0.026 (0.018)	0.061[a] (0.023)	0.038 (0.059)
SMSA, not central city	0.042[b] (0.018)	–	0.064[a] (0.024)	–
SMSA, central city unknown	0.012 (0.019)	–	0.067[a] (0.024)	–
SMSA, central city	−0.057[b] (0.025)	–	0.016(0.034)	–
Constant	0.121(0.270)	0.274[a] (0.075)	0.017(0.157)	0.231(0.246)
R^2	0.282	0.707	0.228	0.589
adj. R^2	0.274	0.704	0.218	0.566
N	2789	2892	2386	516

Source of data: Data in parentheses are standard error

[a]Significant at the 1% level

[b]Significant at the 5% level

[c]Significant at the 10% level

References

Antecol H, Bedard K (2002) The relative earnings of Young Mexican, Black, and White Women. Ind Labor Relat Rev 56(1):122–135

Antecol H, Bedard K (2004) The racial wage gap: the importance of labor force attachment differences across Black, Mexican, and White Men. J Hum Resour 39(2):564–583

Ashenfelter O, Rouse C (1998) Income, schooling, and ability: evidence from a new sample of identical twins. Q J Econ 113(1):253–284

Behrman JR, Rosenzweig MR (1999) Ability biases in schooling returns and twins: a test and new estimates. Econ Educ Rev 18:159–167

Black DA, Berger MC, Scott FA (2000) Bounding parameter estimates with nonclassical measurement error. Am Stat Assoc 95(451):739–748

Blank R (1988) Simultaneously modeling the supply of weeks and hours of work among female household heads. J Labor Econ 6(2):177–204

Blau FD, Kahn LM (1996) Wage structure and gender earnings differentials: an international comparison. Economica 63(250):S29–S62

Bollinger CR (1998) Measurement error in the current population survey: a nonparametric look. J Labor Econ 16(3):576–594

Bound J, Krueger AB (1991) The extent of measurement error in longitudinal earnings data: do two wrongs make a right? J Labor Econ 9(1):1–24

Bound J, Duncan GJ, Rodgers WL (1994) Evidence on the validity of cross-sectional and longitudinal labor market data. J Labor Econ 12(3):345–368

Bound J, Solon G (1999) Double trouble: on the value of twins-based estimation of the return to schooling. Econ Educ Rev 18(2):169–182

Ding Z, Engle RF (2001) Large scale conditional covariance matrix modeling, estimation and testing. Acad Econ Pap 29(2):157–184

Duncan GJ, Hill DH (1985) An investigation of the extent and consequences of measurement error in labor-economic survey data. J Labor Econ 3(4):508–532

Filer RK (1993) The usefulness of predicted values for prior work experience in analyzing labor market outcomes for women. J Hum Resour 28(3):519–537

Garvey N, Reimers C (1980) Predicted vs. potential work experience in an earnings function for young women. In: Ehrenberg R (ed) Research in labor economics 3. JAI, Greenwich, pp 99–127

Kelejian HH (1971) Two-stage least squares and econometric systems linear in parameters but nonlinear in the endogenous variables. J Am Stat Assoc 66(334):373–374

Lee L, Sepanski JH (1995) Estimation of linear and nonlinear errors-in-variables. J Am Stat Assoc 90(429):130–140

Mincer J, Polachek S (1974) Family investments in human capital: earnings of women. J Polit Econ 82(2):S76–S108

Moulton BR (1986) An analysis of female work experience data derived from social security records. J Econ Soc Meas 14(1):66–75

Murphy KM, Welch F (1990) Empirical age-earnings profiles. J Labor Econ 8(2):202–229

Oaxaca RL, Ransom MR (2003) Using econometric models for intrafirm equity salary adjustments. J Econ Inequality 1(3):221–249

Oaxaca RL, Ransom MR (1999) Identification in detailed wage decompositions. Rev Econ Stat 81(1):154–157

Polachek S (1975) Potential biases in measuring male-female discrimination. J Hum Resour 10(2):205–229

Regan TL, Oaxaca RL, Burdhardt G (2007) A human capital model of the effects of abilities and family background on optimal schooling levels. Econ Inq 45(4):721–738

Rodgers WL, Brown C, Duncan GJ (1993) Errors in survey reports of earnings, hours worked, and hourly wages. J Am Stat Assoc 88(424):1208–1218

Sarnikar S, Oaxaca RL, Sorensen T (2007) Did females receive lenient sentences despite the federal sentencing guidelines. Working paper

Styan G (1973) Hadamard products and multivariate statistical analysis. Linear Algebra Appl 6:217–240