

A HUMAN CAPITAL MODEL OF THE EFFECTS OF ABILITY AND FAMILY BACKGROUND ON OPTIMAL SCHOOLING LEVELS

TRACY L. REGAN, RONALD L. OAXACA and GALEN BURGHARDT*

This paper develops a theoretical model of optimal schooling levels where ability and family background are the central explanatory variables. We derive schooling demand and supply functions based on individual wealth maximization. Using the National Longitudinal Survey of Youth 1979 data, we stratify our sample into 1-yr full-time equivalent (FTE) work experience cohorts for 1985–1989. The estimated Mincerian “overtaking” cohort (the years of work experience at which individuals’ observed earnings approximately equal what they would have been based on schooling and ability alone) corresponds to 13 FTE years of experience, yielding on average a rate of return of 10.3% and an average (optimal) 11.4 yr of schooling. (JEL J24, J31, J22)

I. INTRODUCTION

Human capital investments are of wide ranging interest because they can be used to explain income disparities across people, geography, and time. According to Becker (1962), human capital investments are activities that affect future real income streams through the embedding of resources in people. Examples include schooling, on-the-job training (OJT), migration, job search, that is, anything that increases one’s stock of human capital or the value of one’s existing stock. A vast literature supports the social and intellectual interest in income inequality, primarily attributed to differing schooling levels. Schooling is a unique type of investment in that it affects not only current consumption but also future

earnings potential as well. Optimizing individuals choose to invest in schooling until their marginal rate of return equals their discounting rate of interest. Equivalently, they choose their schooling levels so as to maximize their expected (discounted) future earnings stream.

This paper specifies and estimates a human capital model that is based on individual wealth maximization along the lines of the original Austrian problem (Blaug, 1962, pp. 506–507). We use an earnings-schooling relationship to identify individual marginal rates of return to schooling and discounting rates of interest. From these, we can identify and estimate supply and demand functions for

*We would like to thank the workshop participants at the University of Arizona, the IZA/SOLE Summer 2003 Conference, and two anonymous referees for their helpful comments and insights. Special thanks to Price Fishback and Alfonso Flores-Lagunes. We also appreciate the research assistance provided by Laura Martinez.

Regan: Department of Economics, University of Miami, P.O. Box 248126, Coral Gables, FL 33124-6550. Phone (305) 284-5540, Fax (305) 284-2985, E-mail: tregan@miami.edu

Oaxaca: IZA, Bonn, Germany and Department of Economics, University of Arizona, McClelland Hall #401, P.O. Box 210108, Tucson, AZ 85721-0108. Phone (520) 621-4135, Fax (520) 621-8450, E-mail: rlo@email.arizona.edu

Burghardt: Calyon Financial, 550 West Jackson Boulevard, Suite 500, Chicago, IL 60661. Phone (312) 762-1140, Fax (312) 762-1148, E-mail: Galen.burghardt@gmail.com

ABBREVIATIONS

AFQT: Armed Forces Qualification Test
AIC: Akaike’s Information Criterion
FTE: Full-Time Equivalent
IQ: Intelligence Quotient
IV: Instrumental Variables
LM: Lagrange Multiplier
NLOLS: Nonlinear OLS (NLSUR/NLOLS)
NLSUR: Nonlinear Seemingly Unrelated Regressions
NLSY79: National Longitudinal Survey of Youth 1979
OJT: On-the-Job Training
OLS: Ordinary Least Squares
PC: Amemiya’s Prediction Criterion
SC: Schwarz Criterion
SEE: Estimated Residual Standard Error

schooling investment. In this framework, the emphasis on rates of return to schooling is misplaced. Emphasis is properly placed on the optimal level of schooling investment. We ultimately arrive at an optimal level of schooling equation that incorporates permanent family income, family size, and ability. Our estimation strategy borrows from Mincer (1974) and involves disaggregating a sample of white males into 1-yr full-time equivalent (FTE) work experience cohorts for 1985–1989. We estimate a log earnings equation to identify the work experience cohort for whom the estimated residual standard error (SEE) is minimized as well as three other model selection criteria, namely the Akaike's information criterion (AIC), the Schwarz criterion (SC), and Amemiya's prediction criterion (PC). This procedure should help reduce the biases associated with omitted variables, measurement error, and "discount rate." Once identified, the remaining estimation proceeds with the "overtaking" work experience cohort. We employ the following estimation strategies in this paper: ordinary least squares (OLS), nonlinear seemingly unrelated regressions/nonlinear OLS (NLSUR/NLOLS), and two-stage least squares.

The paper is organized as follows: Section II provides the background and literature review. Section III discusses the conceptual framework that underlies the analysis. Section IV discusses the data used in the analysis. Section V presents the results, while Section VI discusses them and provides alternative estimation strategies as well. Finally, Section VII concludes.

II. BACKGROUND AND LITERATURE REVIEW

A substantial portion of the economics literature has been devoted to studying human capital investments and the economic rates of return, particularly in relation to education. Researchers have exploited the models and theories developed by Mincer (1974) and Becker (1962) in their attempts to obtain better estimates of the rates of return.

A variety of modifications to the traditional Mincerian log earnings regression endeavor to correct the potential measurement error bias and omitted variables bias that afflict OLS estimates. Early work addressing the OLS bias includes Griliches (1976, 1977). Behrman and Birdsall (1983), like Card and Krueger (1992),

incorporate a quality of schooling variable into the log earnings regression to correct the omitted variables bias, while Altonji and Dunn (1996), Ashenfelter and Zimmerman (1997), Lang and Ruud (1986), and Agnarsson and Carlin (2002) instead include a family background variable. The twins-based study of Ashenfelter and Krueger (1994) not only addresses the omitted variables bias but also addresses measurement error in schooling through the creative use of both the self- and the twin-reported education levels.¹ While Ashenfelter and Krueger's (1994) large, measurement error-adjusted rates of return to education (i.e., 12%–16%) are now considered an anomaly of the data, their paper laid the foundation for subsequent work (e.g., Ashenfelter and Rouse, 1998; Rouse, 1999; Neumark, 1999; Behrman and Rosenzweig, 1999). Later work has uncovered rates of return (e.g., 9%) that are more reasonable and consistent with the earlier findings (e.g., Willis and Rosen, 1979). The consensus reached by researchers is that omitted variables bias the rates of return upward, whereas measurement error in schooling biases the rates downward. While fixed effects or instrumental variables (IV) are often used to remedy such problems, Griliches (1979) warns that first differencing can exacerbate measurement error in schooling. Card (1995) provides a survey of this work.

In addition to the biases mentioned above, a recent literature has investigated another source of bias in human capital models, specifically that stemming from the heterogeneity in students' access to credit markets for educational decisions. Lang (1993) and Card (1995, 2000) refer to this bias as "discount rate bias." They argue that this bias can help explain the large IV estimates of the rates of return to schooling. Using data from the National Longitudinal Survey of Youth 1979 (NLSY79), Cameron and Taber (2004) find no evidence of credit constraints when they instrument schooling with foregone earnings and the direct costs of schooling. Kling (2001) adopts a Becker (1975) supply and demand model of schooling to examine the types of biases summarized by Card (1995). Generally speaking, Kling (2001) argues that the choice of instrument for schooling may

1. Twin- and sibling-studies predating Ashenfelter and Krueger (1994) include Behrman and Taubman (1976) and Chamberlain and Griliches (1975).

have effects that differ by individuals/groups. IV estimates of rates of return to schooling are interpreted as weighted averages of individual-specific causal effects.

This paper takes a step back and abstracts from some of the issues occupying researchers' attention in recent years. We return to Mincer's (1974) earlier work where he introduces the notion of an overtaking year of work experience in which an individual's observed earnings are most reflective of his investment in school (and innate ability). According to Becker (1962), human capital investments lower observed earnings in the early part of one's working life because observed earnings are net of the costs of investment. However, as an individual ages, his observed earnings rise as he reaps the benefits of the investments.² At the overtaking year of work experience, observed earnings are equal to earnings based on schooling (and ability). The distortion from postschooling investments (e.g., OJT) is minimized because the returns on an individual's prior OJT investment equal the cost of current OJT investment. Thus, an individual's earnings at this point provide the best test of the simple schooling model.

Murphy and Welch (1990) investigate the (in)appropriateness of the quadratic experience term in Mincer's (1974) human capital earnings function. Murphy and Welch (1990) is one of the few studies that address the quadratic experience term; much of the prior research was concerned with the form of the dependent variable. Specifically, they ask how do wages vary with age and consider the confounding effects of experience on earnings. Their empirical findings lend support for Mincer's (1974) emphasis on experience, not age. They note that the severity of problems associated with the quadratic term will depend on how much the variables of interest vary within the experience levels.

III. CONCEPTUAL FRAMEWORK

We posit the existence of an earnings transformation function for the overtaking

2. Using data from the 1982 Employment Opportunity Pilot Program Survey and the 1992 Small Business Administration Survey, Barron, Berger, and Black (1998) find that OJT is associated with a small reduction in one's starting wage. OJT has a larger impact on one's productivity growth than on one's wage growth.

work experience cohort and define it as follows:^{3,4}

$$(1) \quad Y = F(S, A).$$

This function relates an individual's annual earnings, Y , to his years of schooling, S , and to his natural ability, A . For the earnings function to exhibit the conventional positive but diminishing marginal returns to schooling and positive returns to ability, we need the following inequalities to be satisfied:

$$(2) \quad F_S, F_A > 0 \text{ and } F_{SS} < 0.$$

One might also expect more able people to reap greater returns to schooling:⁵

$$(3) \quad F_{SA} = F_{AS} > 0.$$

In the analysis that follows, it is more convenient to think of the earnings transformation function in its log form:

$$(4) \quad \ln Y = \ln F(S, A).$$

Let the marginal rate of return to schooling, r , be defined as follows:

$$(5) \quad r = \partial \ln F(S, A) / \partial S.$$

In order for the marginal rate of return to schooling to increase with ability (and hence for the demand for schooling to increase with

3. As Rosen (1974) points out, the transformation function is derived from a production function of knowledge whose arguments are schooling and ability. The units of knowledge (human capital) are multiplied by a constant market rental rate on human capital to yield earnings. The production function itself is derived from a learning function that governs the rate at which knowledge can be produced from prior schooling and ability. Certainly, other reasonable explanatory variables could be included in this functional form, at the expense of parsimony. However, unless these variables are interacted with schooling, they do not affect the theoretical model because the optimal level of schooling is determined by differentiating with respect to schooling. Furthermore, experience does not appear separately as it is implicitly controlled for in the overtaking model described in Section V.

4. Lazear (1977) frames his discussion of education in the context of a production function.

5. For an early, general discussion of the effects of schooling and ability (and their interaction) on log earnings, see Hause (1972).

ability), we need the following inequality to be satisfied:

$$(6) \quad FF_{SA} > F_A F_S.$$

(See the Appendix for the proof.) Next, we assume that all relevant costs are foregone earnings and that an individual seeks to maximize the present value of his lifetime earnings over an infinite horizon subject to the constraint imposed by Equation (1).⁶ Formally, we can represent an individual's maximization problem as:

$$(7) \quad \underset{S}{Max} V = \int_S^{\infty} Y e^{-it} dt$$

Subject to $Y = F(S, A)$

where V is the present value of lifetime earnings, i is a fixed discounting rate of interest, and t is the index of integration. We assume that there are no borrowing constraints. Work by Lang (1993), Card (1995, 2000), and Cameron and Taber (2004) supports this assumption.

We simplify the present value of lifetime earnings expressed in Equation (7) and take the log of the resulting expression to obtain:

$$(8) \quad \ln V = \ln Y - iS - \ln i.$$

Taking derivatives with respect to S , we arrive at the following first-order condition:

$$(9) \quad r = i.$$

Hence, the optimal level of schooling for an individual occurs at the point where his marginal rate of return to schooling exactly equals his discounting rate of interest as noted by Becker (1962).

The above analysis can be couched in a supply and demand framework. Taking the derivative of the log transformation function as defined in Equation (4) with respect to schooling yields an individual's inverse demand function for schooling,

$$(10) \quad r = r(S, A),$$

which is equivalently expressed as:

6. This infinite horizon is imposed for mathematical simplicity. An infinite horizon model has been used by numerous other researchers as well (e.g., Lang and Ruud, 1986).

$$S^d = S^d(i, A),$$

where S^d is the level of schooling demanded at each discounting rate of interest for an individual with a given (fixed) ability level A .

An individual's supply function for schooling investment can be derived using the present value function as defined in Equation (8). Simple manipulation of this expression yields:

$$(11) \quad \ln Y = \ln(iV) + iS.$$

Differentiating this expression with respect to S , for a given V , yields i which indexes an individual's supply curve thereby establishing the relationship between the supply of schooling and the discounting rate of interest. An individual's discounting rate of interest, i , is uniquely fixed and does not vary with the level of schooling. However, since i can also be interpreted as the marginal opportunity cost of an additional year of school, i can vary across individuals. For example, the discounting rate of interest would likely be higher for children from poorer families than that for children from wealthier families. The same could be said of children from larger families as compared to children from smaller families. Hence, we express i as a function of an individual's family characteristics:

$$(12) \quad i = i(X),$$

where X denotes a vector of family background variables. In the analysis, these include family size and permanent family income.⁷ There are a number of models where family background is central to the analysis. At the present, we have chosen to take a parsimonious view and choose to incorporate family background through i .

By combining Equations (9), (10), and (12), the optimal level of schooling, S^* , is obtained as:

$$(13) \quad S^* = f(X, A).$$

7. Certainly, X could include other variables to address the possibility that i varies across a child's age, birth order, and number of siblings. For example, the number of minutes a parent reads to his/her child could influence i and is probably related to the child's birth order and spacing between siblings. For the most part, such detailed information is not contained in the NLSY79. All we can determine is if a respondent is the oldest child.

In our case, the optimal level of schooling can be graphically illustrated using a supply and demand framework and a framework involving the log earnings functions. Becker and Chiswick (1966) give a very general discussion of how human capital investment can be nested in the context of a supply-and-demand-curve analysis. This can be seen in Figure 1. The top graph relates the log earnings transformation function to the log earnings present value functions as defined in Equation (11). The log earnings transformation function is a concave curve reflecting the positive but diminishing marginal returns to schooling. The log earnings iso-present value functions are represented by a set of parallel lines relating $\ln Y$ and S at a given i . S^* occurs at the point of tangency between these two curves—the point at which discounted lifetime earnings are maximized. Similarly, the bottom graph relates the downward sloping demand function, as defined in Equation (10), to the infinitely elastic supply curve, as defined in Equation (12). The intersection of these two curves corresponds to the point S^* where the discounting rate of interest exactly equals the marginal rate of return to schooling (i.e., the equilibrium as defined in Equation (9)). These two frameworks graphically establish the solution to the maximization problem as defined in Equation (7).

Figure 2 allows A and i to vary across individuals. Fitting a line through the set of tangency points in the top graph parallels the development of Mincer's (1974) simple schooling model:

$$(14) \quad \ln Y_j = \beta_0 + \beta_1 S_j + u_j,$$

for individual j .⁸

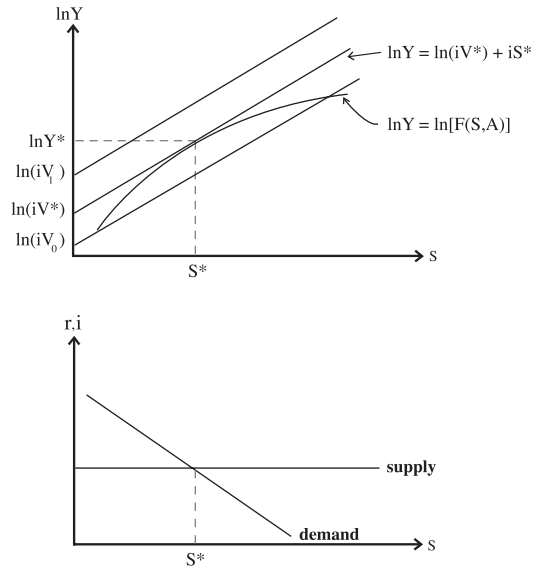
A stochastic approximation to the transformation function as defined in Equation (4) is:

$$(15) \quad \ln Y_j = \beta_0 + \beta_1 S_j + \beta_2 A_j S_j + \beta_3 S_j^2 + \beta_4 A_j + u_{1j}$$

where u_1 is iid(0, σ_1^2). This is a standard human capital functional form that is consistent with the literature. To maintain the

8. Note that the model is not identified. Thus, β_1 has no economic meaning. However, its interpretation as an average rate of return to schooling is maintained throughout the analysis.

FIGURE 1
Graphical Derivation of S^* for an Individual



restrictions corresponding to Equations (2) and (3), we require:

$$(16) \quad \beta_1, \beta_2, \beta_4 > 0 \text{ and } \beta_3 < 0.$$

Differentiating Equation (15) with respect to S yields the schooling investment demand function:

$$(17) \quad r_j = \beta_1 + \beta_2 A_j + 2\beta_3 S_j.$$

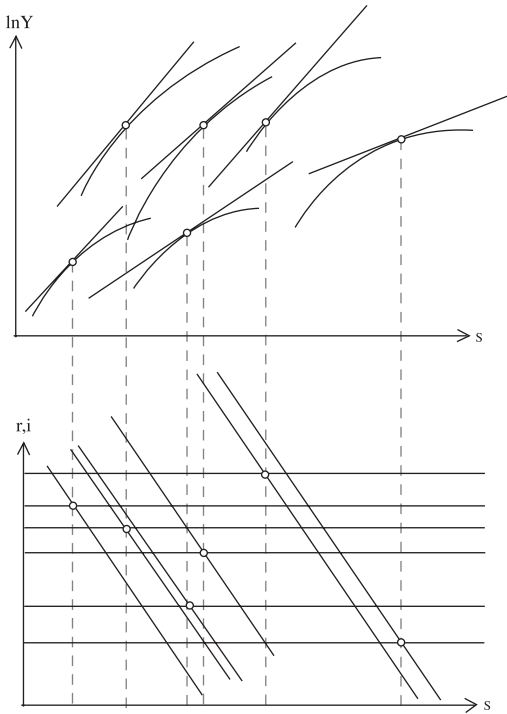
We specify the schooling investment supply function to be a linear function of various family background variables. Consider:

$$(18) \quad i_j = \theta_0 + \theta_1 S_{fj} + \theta_2 S_{mj} + \theta_3 (S_{fj} + S_{mj}) + \theta_4 DVS_{fj} + \theta_5 DVS_{mj} + (\theta_6 + \theta_7) N_j + u_{2j},$$

where S_f is father's schooling, S_m is mother's schooling, N is family size, and u_2 is iid $N(0, \sigma_2^2)$. Permanent family income is proxied with the schooling levels of an individual's parents.⁹ So as to not lose observations and

9. We considered several other proxies for permanent family income, namely, the Duncan socioeconomic index and variations of the parental schooling levels—the average, maximum, and head of household's. Such alternatives were not pursued because we lost too many observations due to missing information.

FIGURE 2
Graphical Derivation of S^* across Individuals



to maintain a constant sample size across regressions for the NLSUR estimations, we assigned an education level of “0” yr for any respondent’s parent whose education level was missing and created dummy variables to indicate whether or not such a value was imposed.¹⁰ Hence, $DVS_{f(m)}$ takes on a value of “1” if we replaced a missing value for the respondent’s father’s (mother’s) education level with a “0.”

The coefficients in Equation (18) are nicely interpreted. θ_1 and θ_2 capture the pure wealth effects of family income on an individual’s discounting rate of interest. We would expect these two coefficients to be negative because an individual’s discounting rate of interest (marginal opportunity cost of an additional year of schooling) decreases with his family wealth (i.e., the individual has the luxury to postpone earnings for more schooling). It is intended that θ_3 captures the effect of family wealth on potential financial aid. Since finan-

cial aid offices base their decisions purely on family wealth, not on individual parental contributions, we sum these two variables together and expect their common parameter, θ_3 , to be positive. While there are no theoretical predictions concerning the expected sign on θ_4 or θ_5 , a positive estimate clearly means that an individual’s discounting rate of interest is higher once we have made the imputation for a missing level of parental schooling. Children from wealthier families have a decreased likelihood of receiving financial aid which raises their discounting rate of interest. The effects of family size on an individual’s marginal opportunity cost of an additional year of schooling can be decomposed into two separate effects: θ_6 captures the pure income effect of family size and θ_7 captures the indirect effect via financial aid considerations. We would expect θ_6 to be positive because individuals from larger families likely have increased opportunity costs to additional schooling. However, the larger a family, the more widely the (financial) resources are spread and hence the greater the opportunity for financial aid assistance. Thus, θ_7 would be negative.

Of course, the individual coefficients are not identified in the above specification, so we collect terms to arrive at:

$$(19) \quad i_j = \alpha_0 + \alpha_1 S_{fj} + \alpha_2 S_{mj} + \alpha_3 DVS_{fj} \\ + \alpha_4 DVS_{mj} + \alpha_5 N_j + u_{2j},$$

where:

$$(20) \quad \alpha_1 = \theta_1 + \theta_3, \\ \alpha_2 = \theta_2 + \theta_3,$$

and

$$\alpha_5 = \theta_6 + \theta_7.$$

The stochastic approximation to Equation (12) is a simple linear model that identifies the differential parental contributions on wealth effects, aside from the financial aid effects since $\alpha_1 - \alpha_2 = \theta_1 - \theta_2$. We recognize that if family size is endogenous in the discounting rate of interest (supply equation), this will ultimately lead to endogeneity of family size in the reduced-form schooling equation below. However, we do not model family fertility

10. Using the NLSY79, Lang and Zagorsky (2001) examine the effects of growing up in a single-parent home on a variety of outcome variables.

decisions because of a desire to focus on the explanatory power of a conceptually straightforward Beckerian schooling demand and supply framework.

The reduced-form optimal level of schooling equation is obtained by substituting Equations (17) and (19) into the individual-specific equilibrium condition:

$$(21) \quad r_j = i_j.$$

Solving for S :

$$(22) \quad S_j = \gamma_0 + \gamma_1 S_{fj} + \gamma_2 S_{mj} + \gamma_3 DVS_{fj} + \gamma_4 DVS_{mj} + \gamma_5 N_j + \gamma_6 A_j + u_{3j},$$

where:

$$(23) \quad \begin{aligned} \gamma_0 &= (\alpha_0 - \beta_1)/2\beta_3, \quad \gamma_1 = \alpha_1/2\beta_3, \\ \gamma_2 &= \alpha_2/2\beta_3, \quad \gamma_3 = \alpha_3/2\beta_3, \\ \gamma_4 &= \alpha_4/2\beta_3, \quad \gamma_5 = \alpha_5/2\beta_3, \\ \gamma_6 &= -\beta_2/2\beta_3, \quad u_{3j} = u_{2j}/2\beta_3, \end{aligned}$$

and

$$\sigma_3^2 = \sigma_2^2/4\beta_3^2.$$

The coefficients' signs establish the net effect of the direct and indirect effects of wealth on schooling. However, γ_6 can be unambiguously signed since more able people reap greater rewards from increased schooling levels. Thus, γ_6 should be positive.

Because an individual's discounting rate of interest and marginal rate of return to schooling are not directly observable, they must be estimated in order to identify the supply and demand functions. In determining an individual's marginal rate of return to schooling, \hat{r}_j , we use the estimated parameters $\hat{\beta}_1$, $\hat{\beta}_2$, and $\hat{\beta}_3$ obtained from OLS estimation of Equation (15). Specifically, $\hat{r}_j = \hat{\beta}_1 + \hat{\beta}_2 A_j + 2\hat{\beta}_3 S_j$. Imposing the equilibrium condition as defined in Equation (21) generates an estimated discounting rate of interest, \hat{i}_j , so that $\hat{i}_j = \hat{r}_j$. We use these estimated marginal rates of return and discounting rates of interest as the dependent variables in the demand of and supply for schooling investment functions, respectively. Note that Equation (19) is estimated explicitly using \hat{i}_j as the dependent variable and Equation (17) is directly constructed from the OLS estimates of Equation (15).

Our empirical strategy follows Mincer's (1974) estimation of the simple schooling model of Equation (14). Mincer's (1974) post-schooling investment model relates earnings to experience and education. Arguably, one might be concerned about potential endogeneity with work experience in a postschooling investment regression. This problem should be mitigated with Mincer's (1974) notion of an overtaking year of work experience in which an individual's observed earnings are most reflective of his investment in school (and innate ability). Hence, experience is no longer a regressor in the log earnings equation. At the point of overtaking, the distortion from postschooling investments (OJT) is minimized since observed earnings approximate the earnings based on schooling (and ability) alone.

The empirical implementation involves stratifying our sample into 1-yr FTE work experience cohorts and running Equation (15) separately for each cohort. This strategy should at least reduce, if not entirely eliminate, the biases typically plaguing log earnings models. Such a procedure allows for a full interaction of each explanatory variable with experience, thus minimizing the aforementioned bias.¹¹ Once the overtaking cohort is identified, based on a series of goodness-of-fit measures, Equations (17), (19), and (22) are estimated.¹²

Goodness-of-Fit Measures

To identify the overtaking year of work experience, we considered five separate "goodness-of-fit" measures for the model described in Equation (15). The most typical and singular way of gauging the "goodness of fit" of an OLS regression is the R^2 measure. Although the number of regressors in Equation (15) does not vary, the degrees of freedom do vary because sample sizes differ for each experience cohort. The \bar{R}^2 measure adjusts for degrees of

11. For comparison's sake, pooling the experience cohorts and including X and X^2 as explicit regressors produce the usual results—the coefficient estimate on X is positive (and statistically significant) and the coefficient estimate on X^2 is negative (and statistically significant). Doing so, however, produces statistically insignificant coefficient estimates on all the other variables (and that on A is negative as well).

12. This of course is a problem if our actual work experience variable is incorrectly measured, but using actual work experience is superior to the use of potential work experience measures (see Regan and Oaxaca, 2006).

freedom, but arguably even this measure does not impose a harsh enough penalty for the loss in degrees of freedom. The next three measures attempt to correct this problem by minimizing the mean-squared error of prediction (Greene, 2000; Kennedy, 1998; Maddala, 2001; Judge et al., 1988).

PC seeks to minimize:

$$(24) \quad \text{PC} = \text{SSE}(1 + k/N)/(N - k) \\ \approx \hat{\sigma}_1^2(1 + k/N),$$

where SSE denotes the total sum of squared errors, k is the number of regressors (including the constant term), N refers to the sample size, and $\hat{\sigma}_1^2$ is the estimated variance of u_1 .

AIC minimizes:

$$(25) \quad \text{AIC} = \ln(\text{SSE}/N) + 2k/N \\ \approx \ln(\hat{\sigma}_1^2) + 2k/N,$$

while the SC seeks to minimize:

$$(26) \quad \text{SC} = \ln(\text{SSE}/N) + k \ln(N)/N \\ \approx \ln(\hat{\sigma}_1^2) + k \ln(N)/N.$$

The PC, AIC, and SC criteria are usually nested in discussions of regressor selection. Typically, researchers test different models using the same data set. We, however, test a common model using different samples to identify the work experience cohort for which the schooling model best explains earnings.

The last “goodness-of-fit” measure we consider is the estimated standard error of the regression. We seek to minimize the estimated residual variance:

$$(27) \quad \hat{\sigma}_1^2 = \text{SSE}/(N - k),$$

(or alternatively its square root, SEE).

IV. DATA

The data used in this study are from the NLSY79. The NLSY79 consists of 12,686 young men and women living in the United States who were between the ages of 14 and 22 when the first survey was conducted in 1979.

The demographic variables were collected from the 1979 interview. We limit our analysis

to white males who are not enrolled in school, currently or during the remainder of the survey, and who earn at least \$500/yr in nominal terms. We also omit anyone who attended school after 1989 to ensure that the wages we observe are truly reflective of the final schooling choices.¹³ Measures of a respondent’s family background/income level include the family size and the highest grade completed by the mother and the father. The NLSY79 provides three measures of a respondent’s ability—the intelligence quotient (IQ), the knowledge of the world of work, and the armed forces qualification test (AFQT). Following most of the literature, we focus on the AFQT measure. For a very early discussion of the use of AFQT in the log earnings function, see Griliches and Mason (1972).

The dependent variable in the log earnings regression is the log of a respondent’s total income from wages and salary in the respective year. Using the consumer price index for all urban consumers, as published by the Bureau of Labor Statistics, we deflated the income figures and express them in terms of 1985 dollars.

The variables used in the construction of the work experience measures were collected from the supplementary NLSY79 work history file. Due to this detailed collection of actual work experience, we do not have to use less precise, potential work experience measures. We calculate a respondent’s FTE work experience for a given year by summing the hours worked in that and all prior years (since 1979) and then divide through by 2,080 (40 h/wk \times 52 wk/yr). Taking account of the fact that many of our respondents were older than age 18 (the usual age that one graduates from high school in the United States) and had potentially been working for several years prior to the first survey, we constructed a variable to approximate their work experience prior to 1979. This variable is calculated as follows:

$$(28) \quad \text{FTE work experience}_{\text{prior to 1979}} \\ = (\text{age}_{1979} - \text{schooling}_{1979} - 6) \\ \times (\text{work experience}_{1979}/2,080).$$

13. The term “final schooling” is used somewhat loosely here because we can only observe individual schooling choices/enrollment through 1998, the most recent wave of the NLSY79 survey that we had at the time of our study. Beginning in 1994, the NLSY79 survey was conducted biannually.

This provides us with a measure of FTE work experience.¹⁴

Like Mincer (1974), we stratified our sample into 1-yr FTE work experience cohorts for 1985–1989.¹⁵ Equation (15) is estimated separately for each work experience cohort, which allows for work experience to fully interact with each coefficient. The earnings data in the model defined by Equations (15), (17), (19), and (22) reflect not only ability and schooling investment decisions but post-schooling investments (e.g., work experience, OJT) as well. Unfortunately, the NLSY79 does not provide adequate information to capture school quality. Ignoring the potential correlation between schooling and work experience in cross-sectional rate of return to schooling models biases OLS. By stratifying our sample into work experience cohorts, we purge the model of any postschooling investment decisions. Thus, there exists an overtaking year in which an individual's earnings are most reflective of his natural ability and schooling levels alone. We reasonably assume that this overtaking year varies across individuals, even within a given work experience cohort. Thus, we stratify our sample into 1-yr FTE work experience cohorts for 1985–1989 to best identify the group whose earnings are on average free of OJT effects.

V. ESTIMATION AND RESULTS

As mentioned above, our statistical estimation pertains to white males who nominally

14. Note that most often these "years" of work experience do not coincide with calendar years and that the composition of the cohorts would differ somewhat with alternative definitions of FTE. Furthermore, implicit in the construction of this measure is the assumption that children begin school at age 6 and complete one grade per year (i.e., no acceleration or retention in schooling) and that the fraction of hours worked in 1979 is proportional to that in the previous years.

15. We chose to confine our attention to 1985–1989 for a couple of reasons. First, Mincer (1974) finds that the correlation between log earnings and education is strongest in the first decade of work experience. The NLSY79 began in 1979 and a decade later corresponds to 1989. Second, Mincer (1974) finds that the overtaking year occurs 8 yr after an individual has left school and has acquired 7–9 "yr" of work experience. In the first year of the survey, our respondents are between ages 14 and 22. Roughly, half are under age 18 and are most likely still enrolled in school. By 1985, the youngest respondents could reasonably have acquired 4 yr of work experience.

earned at least \$500 for a given survey year and who were not enrolled in school currently or any time after 1989. Table 1 provides the descriptive statistics for each variable used in the analysis, when such information is available. On average, our respondents are 18.3 yr old and have the equivalent of a high school education while their parent(s) appear to have completed their junior year of high school. The average household size is 3.8 persons.

A. Sample Stratification

As was previously mentioned, we stratified our sample into 1-yr FTE intervals of work experience for 1985–1989. Table 2 lists the number of people in each respective cohort and the corresponding percentage of the sample they comprise. The procedure for constructing the FTE work experience intervals worked as follows: For example, in constructing the 1-yr work experience cohort, we included individuals for whom we calculated having between 1 (inclusive) and 2

TABLE 1
Descriptive Statistics for White Males

	Mean	Standard Deviation	Nobs.
Age 1979	18.345	2.257	2,761
Nominal wage 1985	15,959.550	9,995.254	1,950
Nominal wage 1986	18,488.490	11,527.950	1,872
Nominal wage 1987	20,675.170	12,354.530	1,918
Nominal wage 1988	23,859.100	33,124.600	1,899
Nominal wage 1989	24,615.260	17,483.510	1,878
Schooling 1985	12.133	2.254	2,032
Schooling 1986	12.149	2.256	1,962
Schooling 1987	12.178	2.241	1,917
Schooling 1988	12.175	2.253	1,940
Schooling 1989	12.157	2.261	1,939
AFQT	46.045	28.657	2,539
Experience 1985	4.470	3.168	2,758
Experience 1986	5.187	3.435	2,758
Experience 1987	5.906	3.721	2,758
Experience 1988	6.641	4.044	2,758
Experience 1989	7.379	4.395	2,758
Mother's schooling	10.993	3.074	2,600
Father's schooling	11.041	3.829	2,494
Family size 1979	3.802	2.224	2,761

Notes: Sample is based on those individuals whose wages are \geq \$500 and who are not enrolled in school currently or any time after 1989. Source of data: NLSY79.

Nobs., Number of observations.

(not inclusive) years of work experience at any time between 1985 and 1989. Below, we describe how the experience calculations were done. In using such a decision rule, we encountered the possibility of individuals having, say, 1.2 yr of work experience in 1985 and 1.9 yr of work experience calculated for 1986. To ensure that an individual entered a particular work experience cohort only once, we manually identified those individuals who were double or even triple counted. For these individuals, we chose to use the most recent year in which their work experience fell within the specified range. Once this year was identified, we chose the individual's corresponding education and income levels. Of course, individuals can and do appear in more than one experience cohort over the period 1985–1989.

We performed similar procedures for all other relevant work experience cohorts and separately estimated the log earnings function specified in Equation (15) for each cohort, excluding ability as a separate regressor. Including ability as an independent regressor in Equation (15), as is often the standard practice, does not affect the overall fit of the model and hence our estimate of the overtaking cohort. In addition, the estimated coefficient on linear ability is never statistically significant except for the 2-yr cohort. For this cohort, only linear ability and linear schooling achieve statistical significance. Consequently, in the log earnings regressions that follow, we assume $\beta_4 = 0$.¹⁶

B. Thirteen-Yr Work Experience Cohort

As can be seen from Table 3, the AIC, SC, PC, and SEE are minimized for the 13-yr work experience cohort, while the R^2 is maximized for the 14-yr cohort. The 13-yr work experience cohort includes a larger sample, and the estimated coefficients are statistically significant and of the expected signs. While the estimated

16. We also controlled for "sheepskin effects" in our log earnings function by including a dummy variable indicating whether a respondent holds a high school diploma or a general educational development test. Cameron and Heckman (1993) find that the two are not equivalent. Research on sheepskin effects is a growing literature and includes work by Hungerford and Solon (1987), Frazis (1993), Kane and Rouse (1995), Jaeger and Page (1996), Arkes (1999), Ferrer and Riddell (2002), and Agnarsson and Carlin (2002). Augmenting Equation (15) with a high school diploma dummy does not affect our choice of the overtaking cohort.

TABLE 2
Work Experience Cohort Frequency
Distribution: 1985–1989

FTE Years of Work Experience 1985–1989	Nobs.	% of sample
<1	206	4.69
1	483	10.99
2	728	16.57
3	997	22.70
4	1,157	26.34
5	1,221	27.79
6	1,186	27.00
7	1,060	24.13
8	916	20.85
9	786	17.89
10	584	13.29
11	422	9.61
12	342	7.79
13	215	4.89
14	149	3.39

Notes: Sample is based on those individuals whose wages are $\geq \$500$ and who are not enrolled in school currently or any time after 1989. Source of data: NLSY79.

Nobs., Number of observations.

coefficients from the 14-yr work experience cohort are of the appropriate signs, the only statistically significant coefficient (at the 10% level) is the schooling-ability interaction term. Thus, our preferred estimate of the overtaking year is 13 FTE years of work experience.

As was previously noted, the AIC, SC, and PC criteria are typically nested in discussions of regressor selection. We, however, employ such criteria to determine which cohort (of varying sample sizes) best fits our proposed log earnings functional form (where the number of regressors is fixed). Thus, the differing degrees of freedom across our regressions are due to variations in the sample size as opposed to the number of explanatory variables. Holding other factors constant (i.e., $\hat{\sigma}_1^2$), the AIC, SC, and PC criteria would favor larger samples. Thus, the use of such criteria would bias our results toward finding earlier work experience cohorts as the overtaking year(s). Given that we estimate the overtaking year to be as high as 13 FTE years of work experience, we believe the bias to be negligible.

Table 4 lists the descriptive statistics for the 13-yr work experience cohort. On average,

TABLE 3
Log Earnings Function: Selection Criterion (Equation (15))

FTE Work Experience Cohort	Nobs.	AIC	SC	PC	SEE	R ²
0	206	0.267	0.332	1.307	1.132	0.046
1	483	-0.476	-0.442	0.621	0.785	0.174
2	728	-0.670	-0.644	0.512	0.713	0.193
3	997	-0.899	-0.879	0.407	0.637	0.228
4	1157	-0.963	-0.946	0.382	0.617	0.199
5	1221	-1.048	-1.032	0.351	0.591	0.183
6	1186	-1.051	-1.033	0.350	0.590	0.161
7	1060	-1.211	-1.192	0.298	0.545	0.175
8	916	-1.240	-1.219	1.2190	0.537	0.200
9	786	-1.123	-1.099	0.325	0.569	0.184
10	584	-1.201	-1.171	0.301	0.547	0.212
11	422	-1.107	-1.069	0.330	0.572	0.136
12	342	-0.474	-0.430	0.622	0.784	0.083
13	215	-1.461	-1.398	0.232	0.477	0.299
14	149	-1.317	-1.236	0.268	0.511	0.353

Notes: Bolded figures correspond to the minimum AIC, SC, PC, and SEE and the maximum R². Samples are based on those individuals whose wages are \geq \$500 and who are not enrolled in school currently or any time after 1989. Source of data: NLSY79.

Nobs., Number of observations.

the overtaking cohort is 28.7 yr old (at any point between 1985 and 1989) and earns a real (nominal) annual income of \$19,594.42 (\$25,417.10). The respondents have been out of school for 11.3 yr after completing their junior year of high school. Many of these individuals are either working overtime or are multiple-job holders because the average experience level is 13.5 yr. The NLSY79 reports the respondent's mother (father) completing 10.6 (10.1) yr of schooling on average. However, including zero for missing values lowers the average level by 1 yr. The mean family size is 3.7 persons.

The remainder of the estimation will be based on the 13-yr overtaking cohort. Table 5, Column 2, lists the OLS results for Equation (15). As theory predicts, the coefficients on schooling and the schooling-ability interaction are positive, while schooling squared is negative. The estimates are statistically significant.

Table 5, Column 1, lists the results from the simple schooling model (Equation (14)). As might be expected when ability is not controlled for, the rate of return to schooling estimated from the simple schooling model is greater than that estimated directly from Equation (15). The simple schooling model

predicts a rate of return of 14%, while the estimates from Equation (15) suggest a 9.7% rate of return.

TABLE 4
Descriptive Statistics 13-yr Work Experience Cohort

	Mean	Standard Deviation	Nobs.
Age	28.693	1.691	215
Age in 1980	20.693	1.414	215
Nominal wage	25,417.10	16,957.80	215
Log real wage	9.883	0.566	215
Schooling	11.377	2.044	215
AFQT	44.823	27.800	215
Experience	13.469	0.294	215
Years out of school	11.316	2.205	215
Mother's schooling	10.621	2.819	195
Father's schooling	10.130	3.750	192
Mother's schooling dummy	0.093	0.291	215
Father's schooling dummy	0.107	0.310	215
Family size 1979	3.693	1.864	215

Notes: Sample is based on those individuals whose wages are \geq \$500 and who are not enrolled in school currently or any time after 1989. Source of data: NLSY79.

Nobs., Number of observations.

TABLE 5
Estimated Schooling Model

Mode/Estimation Strategy	Unrestricted/OLS						
	13 FTE Years						
Cohort	14	15	17	19	22	Years of school completed	
Dependent Variables	ln(earnings)		Estimated r	Estimated i			
Constant	8.295 (43.832)***	7.619 (14.426)***	0.290 (2.965)***	0.151 (14.764)***	0.153 (4.075)***	7.317 (13.449)***	7.420 (2.440)**
AFQT \times schooling	0.140 (8.528)**	0.290 (2.965)***	-1.871E-02 (-2.056)**	—	—	—	—
AFQT	—	4.008E-04 (3.241)***	4.008E-04 (3.241)***	—	—	—	—
Schooling ²	—	-9.356E-03 (-2.056)**	—	—	—	—	—
Father's schooling	—	—	—	-2.442E-03 (-3.717)***	-1.975E-03 (-1.670)*	0.106 (2.862)***	0.130 (1.799)*
Mother's schooling	—	—	—	-3.092E-03 (-3.494)***	-2.757E-03 (-1.711)*	0.147 (3.088)***	0.165 (1.707)*
Father's schooling dummy	—	—	—	-2.499E-02 (-3.062)***	-2.499E-02 (-1.626)	1.335 (2.659)***	1.524 (0.858)
Mother's schooling dummy	—	—	—	-1.089E-02 (-0.945)	-9.796E-03 (-0.790)	0.523 (0.855)	0.582 (0.035)
Family size	—	—	—	-3.895E-05 (-0.035)	-5.542E-04 (-0.481)	2.962E-02 (0.495)	2.082E-03 (-0.481)
R^2	0.255	0.299	—	0.235	—	0.446	—
\bar{R}^2	0.251	0.289	—	0.216	—	0.430	—
SEE	0.490	0.477	—	2.912E-02	2.889E-02	1.544	1.556
Nobs.	215	215	215	215	215	215	215
Estimated at sample mean r, i	0.140	0.097	0.097	0.097	0.103	—	—
Estimated at sample mean optimal years of schooling	—	—	—	—	—	11.377	11.359

continued

TABLE 5
Continued

Model/Estimation Strategy	Restricted/NLSUR				
	13 FTE Years				
Equation	15	17	19	22	
Dependent Variable	ln(earnings)	Estimated <i>r</i>	Estimated <i>i</i>	Years of school completed	
Constant	(8) 7.865 (35.685)***	(9) 0.236 (6.051)***	(10) 0.139 (8.558)***	(11) 7.320 (14.617)***	
Schooling	0.236 (6.051)***	-0.01316714 (-2.759)***	—	—	
AFQT × schooling	3.981E-04 (2.941)***	—	—	—	
AFQT	—	3.981E-04 (2.941)***	—	3.085E-02 (5.980)***	
Schooling	-6.584E-03 (-2.759)***	—	—	—	
Father's schooling	—	—	-1.400E-03 (-1.940)*	0.106 (2.819)***	
Mother's schooling	—	—	-1.977E-03 (-2.094)**	0.150 (3.925)***	
Father's schooling dummy	—	—	-1.786E-02 (-1.963)**	1.356 (2.929)***	
Mother's schooling dummy	—	—	-6.832E-03 (-0.857)	0.519 (0.932)	
Family size	—	—	-3.580E-04 (-0.467)	2.719E-02 (0.459)	
R^2	0.297	—	—	0.446	
\bar{R}^2	—	—	—	—	
SEE	—	—	—	—	
Nobs.	215	215	215	215	
Estimated at sample mean <i>r</i> , <i>i</i>	0.103	0.103	0.103	—	
Estimated at sample mean optimal years of schooling	—	—	—	11.384	

Notes: Sample is based on individuals whose wages are ≥\$500 and who are not enrolled in school currently or any time after 1989. Source of data: NLSY79. (*t*-statistic). *, **, and *** are significant at the 10%, 5%, and 1% levels, respectively. Nobs., Number of observations.

The results from the schooling investment demand function are presented in Table 5, Column 3. Because the coefficients on the demand function are taken directly from Equation (15), the coefficient on ability is positive and that on schooling is negative.

VI. ESTIMATION STRATEGIES

A. Unrestricted OLS

Reduced-Form Optimal Level of Schooling. The initial estimation strategies are based on the assumption that A is uncorrelated with u_1 and u_3 (and hence u_2) and that S is also uncorrelated with u_1 . The first estimation strategy involves the direct estimation of the schooling investment supply function (Equation (19)) by OLS. Since our estimation procedure constrains the model to be in equilibrium, the marginal rates of return calculated from Equation (15) are directly imposed as the dependent variable for Equation (19) (i.e., the discounting rates of interest). Table 5, Column 4, lists these results. The negative coefficient estimates on the permanent family income proxies, the parental education levels, suggest that children from wealthier families have lower discounting rates of interest. This implies that the pure wealth effects of increased parental schooling levels outweigh the indirect effects that family wealth has on the likelihood of receiving financial aid. The estimated coefficients on the parental missing schooling dummies are negative but only statistically significant for the father. Thus, the marginal opportunity cost of an additional year of schooling is lower for those whose father's education level is missing. The coefficient estimate on family size is negative but statistically insignificant which implies that the pure wealth effects of family size completely offset the indirect wealth effects on financial aid. Alternatively, it could be the case that family size has no effect on the discounting rate of interest or that the parameter is imprecisely estimated. Shea (2000) finds that changes in parents' income due to luck have a negligible impact on their children's human capital except when the father has a low level of schooling.

The estimated coefficients from Equations (15) and (19), corresponding to Columns 2

and 4 in Table 5, are used to derive the parameters in Equation (22). Thus:

$$(29) \quad \begin{aligned} \tilde{\gamma}_0 &= (\hat{\alpha}_0 - \hat{\beta}_1) / 2\hat{\beta}_3, \tilde{\gamma}_1 = \hat{\alpha}_1 / 2\hat{\beta}_3, \\ \tilde{\gamma}_2 &= \hat{\alpha}_2 / 2\hat{\beta}_3, \tilde{\gamma}_3 = \hat{\alpha}_3 / 2\hat{\beta}_3, \\ \tilde{\gamma}_4 &= \hat{\alpha}_4 / 2\hat{\beta}_3, \tilde{\gamma}_5 = \hat{\alpha}_5 / 2\hat{\beta}_3, \\ \tilde{\gamma}_6 &= -\hat{\beta}_2 / 2\hat{\beta}_3, \end{aligned}$$

and

$$\tilde{\sigma}_3^2 = \hat{\sigma}_2^2 / 4\hat{\beta}_3^2.$$

Table 5, Column 7, lists these results. The standard errors, hence the t -statistics, have been computed using the delta method (Greene, 2000). It is assumed that $\text{cov}(\hat{\beta}, \hat{\alpha}) \approx 0$. The optimal level of schooling is higher for more able individuals from wealthier families. The optimal level of schooling based on these coefficients for this work experience cohort is 11.4 yr.

Derived Supply Equation. The second estimation strategy directly estimates the log earnings Equation (15) and the optimal level of schooling reduced-form Equation (23)—the two equations in which we observe the dependent variable—by OLS. We can derive consistent estimators of the parameters in the supply Equation (19) from:

$$(30) \quad \begin{aligned} \tilde{\alpha}_0 &= 2\hat{\beta}_3\hat{\gamma}_0 + \hat{\beta}_1, \tilde{\alpha}_1 = 2\hat{\beta}_3\hat{\gamma}_1, \\ \tilde{\alpha}_2 &= 2\hat{\beta}_3\hat{\gamma}_2, \tilde{\alpha}_3 = 2\hat{\beta}_3\hat{\gamma}_3, \\ \tilde{\alpha}_4 &= 2\hat{\beta}_3\hat{\gamma}_4, \tilde{\alpha}_5 = 2\hat{\beta}_3\hat{\gamma}_5, \end{aligned}$$

and

$$\tilde{\sigma}_2^2 = 4\hat{\beta}_3^2\hat{\sigma}_3^2.$$

Table 5, Column 6, lists the OLS results for Equation (23). The signs and magnitudes on the coefficients are similar, but not identical, to those derived above based on the OLS estimates of α and β because the system is over-identified. The estimated coefficients on the parental schooling levels and the associated dummies are smaller for direct OLS, while the coefficients on AFQT and family size are larger. The estimated coefficients on the

parental schooling levels and AFQT are statistically significant.

Table 5, Column 5, lists the derived results of Equation (19). Again, we use the delta method to calculate the standard errors of the estimates. While the signs on the estimated coefficients are identical to those based on the OLS estimates, the magnitudes differ somewhat.

B. Restricted/NLSUR

NLSUR. Another estimation strategy involves the following recursive, constrained system of equations:

$$(31) \quad \begin{aligned} S_j &= \gamma_0 + \gamma_1 S_{fj} + \gamma_2 S_{mj} + \gamma_3 DV_{fj} \\ &\quad + \gamma_4 DV_{mj} + \gamma_5 N_j + \gamma_6 A_j + u_{3j} \\ \ln Y_j &= \beta_0 + \beta_1 S_j + \beta_2 A_j S_j \\ &\quad + \beta_3 S_j^2 + u_{1j} \end{aligned}$$

subject to

$$\gamma_6 = -\beta_2/2\beta_3.$$

We used NLSUR to estimate this restricted recursive system (which requires the sample sizes to be equal). The equations were stacked with the OLS estimates providing the starting values for the iteration. We imposed two alternative variance-covariance matrices for the error terms, Σ , that allowed us to test the following hypothesis:

$$(32) \quad \begin{aligned} H_0 &: \Sigma \text{ is diagonal;} \\ H_1 &: \Sigma \text{ is not diagonal.} \end{aligned}$$

Under the null hypothesis, there is no correlation between the two errors, u_1 and u_3 , and each equation could be estimated separately by NLOLS. The estimated residual variances and covariances were obtained from the OLS estimates of Equations (15) and (22). We tested the null hypothesis using a Breusch Pagan lagrange multiplier (LM) test. The LM test is based on the restricted model where Σ has zero off-diagonal entries. Because the calculated test statistic is less than the critical $\chi_1^{2,.95}$, we cannot reject the null hypothesis and therefore assume that there is no covariance between the error terms. Consequently, each equation could have been estimated sep-

arately by NLOLS, producing consistent but biased results with no loss in efficiency.

Next, we turn to testing the cross-equation restriction:

$$(33) \quad \begin{aligned} H_0 &: \gamma_6 = -\beta_2/2\beta_3; \\ H_1 &: \gamma_6 \neq -\beta_2/2\beta_3. \end{aligned}$$

We were able to test the null hypothesis using a likelihood ratio test. We cannot reject the null hypothesis and thus conclude that the system of equations is in fact constrained but that there is no correlation between the error terms.

Table 5, Columns 8–11, provide the restricted NLSUR results for Equations (15), (17), (19), and (22). All the coefficient estimates from Equation (15), with the exception of that on schooling and the schooling-ability interaction term, increase in statistical significance because estimation of this set of equations by NLSUR imposes cross-equation restrictions that tighten the standard errors making the estimates more precise. Overall, the coefficient estimates decrease in magnitude. The coefficient estimates on Equation (17), derived from Equation (15), are of the expected signs and have similar statistical significance. The derived coefficient estimates on Equation (19), from Equations (15) and (22), are of the same sign as those from the unrestricted OLS estimates, but the magnitudes differ somewhat. The *t*-statistics are larger than those on the previous derived form (i.e., Column 5) but smaller than those when estimated directly (i.e., Column 4). The latter finding may be due to the fact that Equation (19) is not directly part of the constrained system of equations. The estimated coefficients on Equation (22) are nearly identical to those from unrestricted OLS. One could consider estimating a three-equation system (i.e., Equations (15), (19), and (22)) by NLSUR. However, this strategy is not feasible because the variance-covariance matrix is singular.

corr(A, u₃) ≠ 0? Measures of ability pose continuing problems for researchers. The importance of incorporating such a measure is well-documented in the literature; however, choosing an appropriate measure/proxy is a persistent challenge. “First, even our cognitive abilities as adults are heavily influenced by the social environment that we experienced during childhood, making it hard to

discern any influence of preexisting genetic differences. Second, tests of cognitive ability (like IQ tests) tend to measure cultural learning and not pure innate intelligence, whatever that is" (Diamond, 1999, p. 20). Some researchers (e.g., Ashenfelter and Krueger, 1994) have devised resourceful ways of overcoming such problems, but most are left using various potentially error-ridden proxies in their analyses.

Fortunately, the NLSY79 does provide some measures of ability; the question, however, remains as to what type of ability is actually being measured. It is reasonable to question just how well the AFQT score proxies for true, innate ability. The AFQT score comes from the Armed Services Vocational Aptitude Battery test, which was administered in 1980 and used by the armed forces to assess a respondent's measure of trainability. Thus, there are any number of reasons to think that $\text{corr}(A, u_3) \neq 0$, for example, simultaneity bias, omitted variables bias, and so on. In testing for the possible correlation between A and u_3 , we instrumented AFQT with the inverse of a respondent's age in 1980 (the year in which the test was administered) and a set of occupational dummies for the adult present in a respondent's home when he was age 14 along with the other predetermined variables.¹⁷ The inverse of the respondent's age in 1980 allows ability to be concave with respect to age. Thus, we expect ability to increase, but at a decreasing rate, with age conditional on family background characteristics. The positive relationship between a child's ability and a family's resources (financial and time equivalents) is well-known (e.g., Cameron and Heckman, 1998; Cameron and Taber, 2004).

The occupational dummies were constructed based on the respondent's answers to whom he lived with when he was age 14. If there was an adult male present in the household, we used this individual's occupation. If there was no adult male present but an adult female was present, we used her occupation instead. Individuals with other arrangements, those who lived by themselves, and those with no adults

present were coded as missing values. We lose 32 observations due to missing values. We constructed a set of occupational dummies based on the 1970 Census of the population's occupational classification system.

We tested for the potential correlation that exists between A and u_3 using a Hausman specification test (Greene, 2000). We tested the following hypothesis:

$$(34) \quad \begin{aligned} H_0 : \text{plim} (\hat{\gamma}^{\text{OLS}} - \hat{\gamma}^{\text{2SLS}}) &= 0; \\ H_1 : \text{plim} (\hat{\gamma}^{\text{OLS}} - \hat{\gamma}^{\text{2SLS}}) &\neq 0. \end{aligned}$$

The p -value for the Hausman χ^2 statistic is 0.22, so we cannot reject OLS. Thus, our ability proxy, AFQT, does not appear to be correlated with u_3 .

VII. CONCLUDING REMARKS

This paper develops a model of earnings and optimal schooling. The analysis and estimation strategy are inspired by the Mincerian (1974) schooling model. The estimated coefficient on schooling in the simple schooling model (Equation (14)) generally overstates the returns because it does not control for ability. In addition, the simple schooling model is subject to an identification problem if the data in log earnings-schooling space are generated by tangencies between concave earnings functions and linear iso-present value curves. We incorporate human capital investment (i.e., schooling) into a model based on individual wealth maximization while controlling for ability and work experience. The model incorporates the effects of family background on the individual's discounting rate of interest. From this model, we derive individual schooling supply and demand functions that determine optimal schooling levels from the equilibration of the marginal rate of return from an additional year of schooling to the individual's discounting rate of interest.

Using data from the NLSY79, we stratify our sample into 1-yr FTE work experience cohorts over the period 1985–1989 and estimate a log earnings model that incorporates both schooling and ability for each cohort. Our measures of work experience correspond to actual hours worked in past calendar years and allow for lapses in employment and differing employment statuses (i.e., part-, full-, or overtime). Because we impose an FTE status,

17. Cameron and Heckman (1998) address the spurious correlation that potentially exists between AFQT and schooling by conditioning on a subset of individuals who were between ages 14 and 17 when the test was administered and hence still in school. Doing so eliminates any causal effect of schooling on ability.

our measures of work experience do not necessarily correspond to an actual calendar year. Based on the estimates of Equation (15) and the “goodness-of-fit” measures, we conclude that the overtaking cohort corresponds to individuals with 13 FTE years of work experience (11 calendar years). The earnings of this cohort are most reflective of natural ability and schooling investments.

Based on our empirical findings, we conclude that we have a constrained system of equations relating earnings determination and optimal schooling. We assume that the error term in the log earnings function is normally distributed and determine that it is not correlated with the error term in the optimal level of schooling equation. According to a Hausman specification test, we cannot reject OLS and conclude that measured ability (AFQT) is uncorrelated with u_3 (and hence exogenous to the system). Thus, our most preferred set of estimates corresponds to Columns 8–11 of Table 5.

Since the schooling equation parameters vary across experience cohorts, the full interaction of experience with the schooling production function parameters in the overtaking cohort addresses the bias inherent in estimating a pooled earnings model with additive experience and its square. According to Mincer’s (1974) rule of thumb (1/overtaking year), 13 yr of FTE work experience corresponding to 11 yr beyond the completion of schooling yield approximate rates of return of 7.7% and 9.1%. Our model estimates that the (average) marginal rate of return to schooling is 10.3% and the optimal level of schooling is 11.4 yr. Our estimate of the rate of return to schooling is consistent with the past findings.

APPENDIX

Proof of $FF_{SA} > F_A F_S$.

$$\begin{aligned} r &= \partial \ln F(S, A) / \partial S = F_S / F \\ \Rightarrow \partial r / \partial A &= (FF_{SA} - F_S F_A) / F^2 > 0 \\ \Rightarrow FF_{SA} &> F_A F_S. \end{aligned}$$

REFERENCES

Agnarsson, S., and P. S. Carlin. “Family Background and the Estimated Return to Schooling.” *Journal of Human Resources*, 37, 2002, 680–692.

- Altonji, J. G., and T. A. Dunn. “The Effects of Family Characteristics on the Return to Education.” *Review of Economics and Statistics*, 78, 1996, 692–703.
- Arkes, J. “What Do Educational Credentials Signal and Why Do Employers Value Credentials?” *Economics of Education Review*, 18, 1999, 133–141.
- Ashenfelter, O., and A. Krueger. “Estimates of the Return to Schooling from a New Sample of Twins.” *American Economic Review*, 84, 1994, 1157–1173.
- Ashenfelter, O., and C. Rouse. “Income, Schooling, and Ability: Evidence from a New Sample of Identical Twins.” *Quarterly Journal of Economics*, 113, 1998, 253–284.
- Ashenfelter, O., and D. J. Zimmerman. “Estimates of the Returns to Schooling from Sibling Data: Fathers, Sons, and Brothers.” *Review of Economics and Statistics*, 97, 1997, 1–9.
- Barron, J. M., M. C. Berger, and D. A. Black. “Do Workers Pay for On-the-Job Training?” *Journal of Human Resources*, 34, 1998, 235–252.
- Becker, G. S. “Investment in Human Capital: A Theoretical Analysis.” *Journal of Political Economy*, 70, 1962, 9–49.
- . *Human Capital*. Chicago, IL: University of Chicago Press, 1975.
- Becker, G. S., and B. R. Chiswick. “Education and the Distribution of Earnings.” *American Economic Review*, 56, 1966, 358–369.
- Behrman, J., and P. Taubman. “Intergenerational Transmission of Wealth and Income.” *American Economic Review Papers and Proceedings*, 66, 1976, 436–440.
- Behrman, J. R., and N. Birdsall. “The Quality of Schooling: Quantity Alone is Misleading.” *American Economic Review*, 73, 1983, 928–946.
- Behrman, J. R., and M. R. Rosenzweig. “‘Ability’ Biases in Schooling Returns and Twins: A Test and New Estimates.” *Economics of Education Review*, 18, 1999, 159–167.
- Blaug, M. *Economic Theory in Retrospect*. Homewood, IL: Richard D. Irwin, Inc., 1962.
- Cameron, S., and C. Taber. “Estimation of Educational Borrowing Constraints Using Returns to Schooling.” *Journal of Political Economy*, 112, 2004, 132–182.
- Cameron, S. V., and J. J. Heckman. “The Nonequivalence of High School Equivalents.” *Journal of Labor Economics*, 11, 1993, 1–47.
- . “Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males.” *Journal of Political Economy*, 106, 1998, 262–333.
- Card, D. “Earnings, Schooling, and Ability Revisited,” in *Research in Labor Economics*, edited by S. Polachek. Greenwich, CT: JAI Press, 14, 1995, 23–48.
- . “The Causal Effect of Education on Earnings,” in *Handbook of Labor Economics*, edited by O. Ashenfelter and D. Card. Amsterdam, The Netherlands: Elsevier, 3A, 2000, 1801–1863.
- Card, D. and A. B. Krueger. “Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States.” *Journal of Political Economy*, 100, 1992, 1–40.
- Chamberlain, G., and Z. Griliches. “Unobservables with a Variance-Components Structure: Ability, Schooling, and the Economic Success of Brothers.” *International Economic Review*, 16, 1975, 422–449.

- Diamond, J. *Guns, Germs, and Steel: The Fates of Human Societies*. New York: W.W. Norton & Company, 1999.
- Ferrer, A. M., and W. C. Riddell. "The Role of Credentials in the Canadian Labour Market." *Canadian Journal of Economics*, 35, 2002, 879–905.
- Frazis, H. "Selection Bias and the Degree Effect." *Journal of Human Resources*, 28, 1993, 538–554.
- Greene, W. H. *Econometric Analysis*. 4th ed., Upper Saddle River, NJ: Prentice-Hall, Inc., 2000.
- Griliches, Z. "Wages of Very Young Men." *Journal of Political Economy*, 84, 1976, S69–S86.
- . "Estimating the Returns to Schooling: Some Econometric Problems." *Econometrica*, 45, 1977, 1–22.
- . "Sibling Models and Data in Economics: Beginnings of a Survey." *Journal of Political Economy*, 87, 1979, S37–S64.
- Griliches, Z., and W. M. Mason. "Education, Income, and Ability." *Journal of Political Economy*, 80, 1972, S74–S103.
- Hause, J. C. "Earnings Profile: Ability and Schooling." *Journal of Political Economy*, 80, 1972, S108–S138.
- Hungerford, T., and G. Solon. "Sheepskin Effects in the Returns to Education." *Review of Economics and Statistics*, 69, 1987, 175–177.
- Jaeger, D. A., and M. E. Page. "Degrees Matter: New Evidence on Sheepskin Effects in the Returns to Education." *Review of Economics and Statistics*, 78, 1996, 733–740.
- Judge, G. G., R. C. Hill, W. E. Griffiths, H. Lutkepohl, and T. C. Lee. *Introduction to the Theory and Practice of Econometrics*, 2nd Edition. New York: John Wiley & Sons, 1998.
- Kane, T. J., and C. E. Rouse. "Labor-Market Returns to Two- and Four-Year Colleges." *American Economic Review*, 85, 1995, 600–614.
- Kennedy, P. *A Guide to Econometrics*. 4th ed., Cambridge, MA: The MIT Press, 1998.
- Kling, J. R. "Interpreting Instrumental Variables Estimates of the Returns to Schooling." *Journal of Business & Economic Statistics*, 19, 2001, 358–364.
- Lang, K. "Ability Bias, Discount Rate Bias, and the Return to Education." Working Paper, Boston University, 1993.
- Lang, K., and P. A. Ruud. "Returns to Schooling, Implicit Discount Rates and Black-White Wage Differentials." *Review of Economics and Statistics*, 68, 1986, 41–47.
- Lang, K., and J. L. Zagorsky. "Does Growing Up with a Parent Absent Really Hurt?" *Journal of Human Resources*, 36, 2001, 253–273.
- Lazear, E. "Education: Consumption or Production?" *Journal of Political Economy*, 85, 1977, 569–598.
- Maddala, G. S. *Introduction to Econometrics*, 3rd ed., Chichester, England: John Wiley & Sons, Ltd., 2001.
- Mincer, J. *Schooling, Experience, and Earnings*. New York: National Bureau of Economic Research, 1974.
- Murphy, K. M., and F. Welch. "Empirical Age-Earnings Profiles." *Journal of Labor Economics*, 8, 1990, 202–229.
- Neumark, D. "Biases in Twin Estimates of the Return to Schooling." *Economics of Education Review*, 18, 1999, 143–148.
- Regan, T. L., and R. L. Oaxaca. "Work Experience as a Source of Specification Error in Earnings Models: Implications for Gender Wage Decompositions." *IZA Discussion Paper No. 1920*, 2006.
- Rosen, S. "Human Capital and the Internal Rate of Return." *Industrial Relations Research Association Series*, 1974, 243–250.
- Rouse, C. E. "Further Estimates of the Economic Return to Schooling from a New Sample of Twins." *Economics of Education Review*, 18, 1999, 149–157.
- Shea, J. "Does Parents' Money Matter?" *Journal of Public Economics*, 77, 2000, 155–184.
- Willis, R. J., and S. Rosen. "Education and Self-Selection." *Journal of Political Economy*, 87, 1979, S7–S36.