I Z A

IZA DP No. 1927

A Human Capital Model of the Effects of Abilities and Family Background on Optimal Schooling Levels

Tracy L. Regan
Galen Burghardt
Ronald Oaxaca

January 2006

A Human Capital Model of the Effects of Abilities and Family Background on Optimal Schooling Levels

Tracy L. Regan
University of Miami

Galen Burghardt
Calyon Financial

Ronald L. Oaxaca
*University of Arizona
and IZA Bonn*

ABSTRACT

A Human Capital Model of the Effects of Abilities and Family Background on Optimal Schooling Levels[*]

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 NLSY79 data we stratify our sample into one-year "FTE" work experience cohorts for 1985-1989. Mincer's (1974) "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 work experience yielding on average a rate of return of 9.6 percent and an average (optimal) 11.4 years of schooling.

JEL Classification: J24, J31, J22

Keywords: human capital, ability, family background, schooling, earnings

Corresponding author:

Ronald L. Oaxaca
Department of Economics
University of Arizona
P.O. Box 210108
Tucson, AZ 85721-0108
USA
Email: rlo@email.arizona.edu

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, i.e. 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. 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 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 levels, family size, and abilities. Our estimation strategy borrows from Mincer (1974) and involves disaggregating our sample of white males into one-year full time equivalent (FTE) work experience cohorts for 1985-1989. We estimate a log wage equation to identify the work experience cohort that minimizes the estimated residual standard error as well as three other model selection criteria, namely the Akaike information

1

criterion, the Schwarz criterion, and Amemiya's prediction criterion. Once identified, the remaining estimation proceeds with this work experience cohort. We employ the following estimation strategies in this paper: OLS, Nonlinear Seemingly Unrelated Regressions/Nonlinear OLS (NLSUR/NLOLS), and 2SLS.

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 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. Ashenfelter and Krueger's (1994) innovative twins-based study not only addresses the omitted variables bias, but also measurement error in schooling through the creative use of both the self- and twin-reported education levels. While Ashenfelter and Krueger's large, measurement error adjusted rates of return to edu-

2

cation (i.e. 12-16 percent) 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 percent) which are more reasonable and consistent with earlier findings (e.g., Willis and Rosen, 1979). The consensus reached by researchers is that omitted variables biases the rates of return upwards whereas the measurement error in schooling biases the rates downwards. While fixed effects or instrumental variables 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 instrumental variables estimates of the rates of return to schooling. Using NLSY79 data, Cameron and Taber (2004) find no evidence of credit constraints when they instrument with foregone earnings and the direct costs of schooling. Kling (2001) adopts a Becker supply and demand model of schooling to examine the types of biases summarized by Card (1995). Generally speaking, Kling argues that the choice of instrument for schooling may 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 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 human capital investments lower observed

3

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.[1] At the "overtaking" year of work experience observed earnings are equal to earnings based on schooling (and ability). The distortion from post-schooling 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 Mincer's human capital earnings function. Murphy and Welch's paper is one of the few studies that adress 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 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.

---

[1]Using data from the 1982 EOPP Survey and the 1992 SBA Survey, Barron et al. (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.

III. CONCEPTUAL FRAMEWORK

We posit the existence of an earnings transformation function and define it as follows,[2,3]

$$Y = F(S, A). \tag{1}$$

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,

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

Since we might also expect more able people to reap greater rewards (i.e. in the form of their resulting wage structure),[4]

$$F_{SA} = F_{AS} > 0. \tag{3}$$

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

---

[2] 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.

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

[4] For a general discussion of the effects of schooling and ability (and their interaction) on log earnings, see Hause (1972).

$$\ln Y = \ln F(S, A). \tag{4}$$

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

$$r = \frac{\partial \ln F(S, A)}{\partial S}. \tag{5}$$

In order for the marginal rate of return to schooling to increase with ability (and hence for the demand for schooling to increase with ability) we need the following inequality to be satisfied,

$$FF_{SA} > F_A F_S. \tag{6}$$

(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 (1).[5] Formally, we can represent an individual's maximization problem as,

$$\underset{S}{Max} V = \int_S^\infty Y e^{-it} dt \text{ subject to } Y = F(S, A), \tag{7}$$

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 there are no borrowing constraints. Work by Lang (1993), Card (1995, 2000), and, Cameron and Taber (2004) support this assumption.

We simplify the present value of lifetime earnings and take the log of the resulting expression to obtain,

$$\ln V = \ln Y - iS - \ln i. \tag{8}$$

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

---

[5] 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).

$$r = i. \tag{9}$$

Hence, the optimal level of schooling for an individual occurs at the point where his marginal rate of return to schooling exactly equals his/her 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 (4) with respect to schooling yields an individual's inverse demand for schooling,

$$r = r(S, A), \tag{10}$$

which is equivalently expressed as,

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

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. Thus, the demand for schooling is the marginal rate of return to schooling.

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

$$\ln Y = \ln(iV) + iS. \tag{11}$$

Differentiating this expression with respect to S produces 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

7

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 define i as a function of an individual's family characteristics,

$$i = i(X), \tag{12}$$

where X denotes a vector of family background variables. In the analysis these include family size and permanent family income levels.

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

$$S^* = f(X, A). \tag{13}$$

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 (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 a set of straight lines relating lnY and S at a given i. The optimal level of schooling, 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 (10), to the infinitely elastic supply curve, as defined in (12). The intersection of these two curves corresponds to the point where the discounting rate of interest exactly equals the marginal rate of return to schooling (i.e. the equilibrium as defined in (9)). This in turn establishes again, the optimal

level of schooling, S^*. These two frameworks graphically establish the solution to the maximization problem as defined in (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 simple schooling model,

$$\ln Y_j = \beta_0 + \beta_1 S_j + u_j, \tag{14}$$

for individual j.[6]

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

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

where u_1 is iid $N(0, \sigma_1^2)$. To maintain the restrictions corresponding to (2) and (3), we would require

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

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

$$r_j = \beta_1 + \beta_2 A_j + 2\beta_3 S_j. \tag{17}$$

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

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

---

[6]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.

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.[7] So as to not lose observations, and to maintain a constant sample size across regressions, for the NLSUR estimations, we assigned an education level of "0" years for any respondent's parent whose education level was missing and created dummy variables to indicate whether or not such a value was imposed.[8] 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 (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). However, θ_3 captures the effect of family wealth on potential financial aid. Since financial 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. Children from richer 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

---

[7]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 the household's level. Such alternatives were not pursued because we lost too many observations.

[8]Using the NLSY79, Lang and Zagorsky (2000) examine the effects of growing up in a single parent home on a variety of outcome variables.

10

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

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

where,

$$\begin{aligned}
\theta_1 + \theta_3 &= \alpha_1, \\
\theta_2 + \theta_3 &= \alpha_2,
\end{aligned} \qquad (20)$$

and

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

The above specification identifies the differential parental contributions on wealth effects, aside from the financial aid effects since $\alpha_1 - \alpha_2 = \theta_1 - \theta_2$.

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

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

and solving for S,

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

where

$$\gamma_0 = \frac{\alpha_0 - \beta_1}{2\beta_3}, \tag{23}$$

$$\gamma_1 = \frac{\alpha_1}{2\beta_3},$$

$$\gamma_2 = \frac{\alpha_2}{2\beta_3},$$

$$\gamma_3 = \frac{\alpha_3}{2\beta_3},$$

$$\gamma_4 = \frac{\alpha_4}{2\beta_3},$$

$$\gamma_5 = \frac{\alpha_5}{2\beta_3},$$

$$\gamma_6 = \frac{-\beta_2}{2\beta_3},$$

$$u_{3j} = \frac{u_{2j}}{2\beta_3},$$

and

$$\sigma_3^2 = \frac{\sigma_2^2}{4\beta_3^2}.$$

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 estimating an individual's marginal rate of return to schooling, $\hat{r}_j$, we use the estimated parameters obtained from the OLS estimation of (15). Imposing the equilibrium condition as defined in (21) generates an estimated discounting rate of interest, $\hat{\imath}_j$, so that $\hat{\imath}_j = \hat{r}_j$. 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.

Our empirical strategy follows Mincer's estimation of the simple schooling model

of equation (14). Mincer's post-schooling investment model relates earnings to experience and education. Arguably, one might be concerned about potential endogeneity with work experience in a post-schooling investment regression. This problem should be mitigated with Mincer introduction 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 post-schooling investments (i.e. OJT) is minimized since observed earnings approximate the earnings based on schooling (and ability) alone.

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 (15). The most typical and singular way of gauging the "goodness of fit" of a regression is the R^2. Although the number of regressors in (15) do not vary, the degrees of freedom do vary because sample sizes vary for each experience cohort. The $\overline{R}^2$ measure adjusts for degrees of 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 (MSE) of prediction,

$$E(\ln Y_f - \widehat{\ln Y}_f)^2, \tag{24}$$

where $\ln Y_f$ is the future value of $\ln Y$ and $\ln \widehat{Y}_f$ is the predicted future value [Greene, 2000; Kennedy, 1998; Maddala, 2001; Judge et al., 1988].

Amemiya's prediction criteria (PC) seeks to minimize,

$$PC = \frac{SSE(1 + \frac{k}{N})}{N - k} = \sigma_1^2(1 + \frac{k}{N}), \tag{25}$$

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 σ_1^2 is the variance of u_1.

Akaike's information criterion (AIC) minimizes,

$$AIC = \ln \frac{SSE}{N} + \frac{2k}{N} \approx \ln \sigma_1^2 + \frac{2k}{N}, \tag{26}$$

while the Schwarz (SC) criterion seeks to minimize,

$$SC = \ln \frac{SSE}{N} + \frac{k \ln N}{N} \approx \ln \sigma_1^2 + \frac{k \ln N}{N}. \tag{27}$$

The PC, AIC, and SC criterion are typically 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 whom 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,

$$\widehat{\sigma}_1^2 = \frac{SSE_j}{(N-k)}, \tag{28}$$

(or alternatively its square root, i.e. S.E.E.).

IV. DATA

The data used in this study are from the National Longitudinal Survey of Youth 1979 (NLSY79). The NLSY79 consists of 12,686 young men and women, living in the U.S., who were between the ages of 14 and 22 years when the first survey was conducted in 1979.

The demographic variables were collected from the 1979 interview. We restrict

our attention to white males and are left with 4,393 individuals. 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 (KWW), and the Armed Forces Qualification Test (AFQT). We used the latter two measures because an IQ score was not provided for many of the respondents.

Educational attainment and enrollment status were available for each survey year. We have measures indicating both the highest grade completed and current enrollment status. Our experience cohort samples do not include any respondent who is currently enrolled in school. We also omit anyone who attended school after 1989 to ensure that the wages we observe are truly reflective of the final schooling choices.[9]

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 CPI (Consumer Price Index) for all urban consumers, as published by the BLS (Bureau of Labor Statistics), we deflated the income figures and express them in terms of 1985 dollars. In our analysis, we only consider people who nominally earned at least $500 for a given year.

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 2080 (40 *hours per week* * 52 *weeks per year*). Taking account of the fact that many of our respondents were older than 18 years (the usual age

---

[9]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.

that one graduates from high school in the U.S.) and had potentially been working for several years prior to the first survey, we constructed a variable to measure their work experience prior to 1979. This variable is calculated as follows,

$$FTE\ Work\ Experience_{prior\ to\ 1979}$$

$$=\ (Age_{1979} - Schooling_{1979} - 6) * (Work\ Experience_{1979}/2080). \qquad (29)$$

This provides us with a measure of "full-time equivalent" (FTE) work experience.[10]

Like Mincer we stratified our sample into one-year FTE work experience cohorts for 1985-1989.[11] 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 (15), (17), (19), and (22) reflects not only ability and schooling investment decisions but post-schooling investments (e.g., work experience, on-the-job training) as well. Ignoring the potential correlation between schooling and work experience in cross-section return to schooling models biases OLS. By stratifying our sample into work experience cohorts we purge the model of any

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

[11]We chose to confine our attention to 1985-1989 for several 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 finds that the "overtaking" year occurs eight years after an individual has left school and has acquired seven to nine "years" of work experience. In the first year of the survey, our respondents are between 14 and 22 years old. Roughly, half are under 18 years and are most likely still enrolled in school. By 1985 the youngest respondents could reasonably have acquired four years of work experience. The work experience is more crudely measured in early years of the survey. Going late into the survey probably biases us towards finding a later "overtaking" year. As a rough rule of thumb the rate of return to education can be approximated by the inverse of the "overtaking" year of experience. This method predicts a 16.7 percent rate in 1985 and a 9.1 percent rate in 1989, both reasonable estimates.

post-schooling 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 one-year 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 earned at least $500 for a given survey year and who were not enrolled in school currently or anytime after 1989. Table 1 provides the descriptive statistics for each variable used in the analysis. On average, our respondents are 18 years 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 four persons. There is a positive time-trend in schooling levels, years of experience, and earnings.

Sample Stratification

As was previously mentioned, we stratified our sample into one-year FTE intervals of work experience for 1985-1989. Table 2 lists the number of people in each respective cohort and the corresponding percentage they comprise of the sample. The procedure for constructing the FTE work experience intervals worked as follows: For example, in constructing the one-year work experience cohort, we included individuals who reported having between one (inclusive) and two (not inclusive) years of work experience at any time between 1985 and 1989. In using such a decision rule we encountered the possibility of individuals who reported say, 1.2 years of work experience in 1985 and 1.9 years of work experience in 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 for each cohort (15), excluding ability as a separate regressor. Including ability as an independent regressor in (15), as is often 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-year cohort. For this cohort only linear ability and linear schooling achieve statistical significance. Consequently, in the log earnings regressions that follow we assume $\beta_4 = 0$.[12]

Table 3 lists the selection criterion (AIC, SC, PC, $\widehat{\sigma}_1$, and R^2) for each regression done utilizing AFQT as the ability proxy. Although the results are similar for AFQT and KWW, we obtain somewhat better results, in terms of sign and significance, when AFQT is used as the ability proxy. Thus, we do not include the results which use the KWW score instead.[13]

---

[12]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 G.E.D. 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 (15) with a high school diploma dummy does not affect our choice of the "overtaking" cohort, however, and hence such estimates are not included.

[13]For an earlier discussion of the use of AFQT in the log earnings function see Griliches and Mason (1972).

13-Year Work Experience Cohort

As can be seen from Table 3 the AIC, SC, PC, and estimated residual standard error are minimized for the 13-year work experience cohort while the R^2 is maximized for the 14-year cohort. The 13-year work experience cohort includes a larger sample and the estimated coefficients are statistically significant and of the expected signs. While the estimated coefficients from the 14-year work experience cohort are of the appropriate signs, the only statistically significant coefficient (at the 10 percent level) is the schooling-ability interaction term. Thus, we estimate the "overtaking" year as 13 FTE years of work experience.[14] The Bera and Jarque test supports our assumption of normality of the errors [Greene, 2000].

As was previously noted, the AIC, SC, and PC criterion are typically nested in discussions of regressor selection. We, however, employ such criterion 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. σ_1^2), the AIC, SC, and PC criterion would favor larger samples. Thus, the use of such criterion would bias our results towards 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-year work experience cohort. On average, the "overtaking" cohort is 28.7 years old (at any point between 1985 and

---

[14]One might have concern over our choice of the 13-year work experience cohort as the "overtaking" cohort because of the jumps in the estimated residual error that occur in prior cohorts. In order to test the robustness of the 9.7 percent rate of return to education, we re-estimated the model using the 10-, 11-, and 12-year work experience cohorts and found rates of return to be 9.9, 8.7, and 7.6 percent, respectively.

1989) and earn a real (nominal) annual income of \$19,594.42 (\$25,417.10). The respondents have been out of school for 11.3 years after completing their junior year of high school. Many of these individuals are either working over-time or are multiple-job holders because the average experience level is 13.5 years. The NLSY79, reports the respondents' parents completing their sophomore year of high school. However, adjusting these figures for missing values lowers the average level by one year. The average family size is 3.7 persons.

The remainder of the estimation will be based on the 13-year "overtaking" cohort. Table 5, column 2, lists the OLS results for (15). As theory predicts, the coefficients on schooling and schooling-ability interacted are positive while schooling squared is negative. The estimates are statistically significant at the five percent level.

Table 5, column 1, lists the results from the simple schooling model (14). As might be expected when ability is not controlled for, the rate of return estimated from the simple schooling model is greater than that estimated directly from (15). The simple schooling model predicts a rate of return of 14.0 percent while the estimates from (15) suggest a 9.5 percent rate of return.

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 (15), the coefficient on ability is positive and that on schooling is negative.

VI. ESTIMATION STRATEGIES

VI(i). Unrestricted/OLS

1. 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 (19) by OLS. Since our estimation procedure constrains the model to be in equilibrium, the marginal rates of return calculated from (15) are directly imposed as the dependent variables for (19) (i.e. the discounting rates of interest). Table 5, column 4, lists these results. The negative coefficient on the permanent family income proxies, the parental education levels, suggests 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. Shea (2000) finds that changes in parents' income due to luck have a negligable impact on their children's human capital except when the father has a low level of schooling.

The estimated coefficients from (15) and (19) corresponding to columns 2 and 4 in Table 5 are used to derive the parameters in (22). Thus,

$$\widetilde{\gamma}_0 = \frac{\hat{\alpha}_0 - \hat{\beta}_1}{2\hat{\beta}_3}, \ \widetilde{\gamma}_1 = \frac{\hat{\alpha}_1}{2\hat{\beta}_3}, \ \widetilde{\gamma}_2 = \frac{\hat{\alpha}_2}{2\hat{\beta}_3}, \ \widetilde{\gamma}_3 = \frac{\hat{\alpha}_3}{2\hat{\beta}_3}, \ \widetilde{\gamma}_4 = \frac{\hat{\alpha}_4}{2\hat{\beta}_3}, \ \widetilde{\gamma}_5 = \frac{\hat{\alpha}_5}{2\hat{\beta}_3}, \ \widetilde{\gamma}_6 = \frac{-\hat{\beta}_2}{2\hat{\beta}_3},$$

(30)

and

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

Table 5, column 7, lists these results. The standard errors, hence the t-statistics,

have been computed using the Delta Method. It is assumed that $cov(\beta, \alpha) \approx 0$. 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 years.

2. Derived Supply Equation

The second estimation strategy directly estimates the log earnings equation (15) and the optimal level of schooling reduced-form equation (22) (i.e. 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

$$\widetilde{\alpha}_0 = 2\hat{\beta}_3\hat{\gamma}_0 + \hat{\beta}_1, \ \widetilde{\alpha}_1 = 2\hat{\beta}_3\hat{\gamma}_1, \ \widetilde{\alpha}_2 = 2\hat{\beta}_3\hat{\gamma}_2, \ \widetilde{\alpha}_3 = 2\hat{\beta}_3\hat{\gamma}_3, \ \widetilde{\alpha}_4 = 2\hat{\beta}_3\hat{\gamma}_4, \ \widetilde{\alpha}_5 = 2\hat{\beta}_3\hat{\gamma}_5,$$

$$(31)$$

and

$$\widetilde{\sigma}_2^2 = 4\widehat{\beta}_3^2\widehat{\sigma}_3^2.$$

Table 5, column 6, lists the OLS results for (22). 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 overidentified. The estimated coefficients on the parental schooling levels, the parental missing schooling dummies, and the family size variables are smaller while the coefficient on AFQT is larger. All of the coefficients, except for that on the mother's missing schooling dummy and the family size, are statistically significant at the five percent level.

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

VI(ii). Restricted/NLSUR

3. NLSUR

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

$$
\begin{aligned}
S_j &= \gamma_0 + \gamma_1 S_{fj} + \gamma_2 S_{mj} + \gamma_3 DV_{fj} + \gamma_4 DV_{mj} + \gamma_5 N_j + \gamma_6 A_j + u_{3j} \quad (32)\\
\ln Y_j &= \beta_0 + \beta_1 S_j + \beta_2 A_j S_j + \beta_3 S_j^2 + u_{1j}
\end{aligned}
$$

subject to

$$
\gamma_6 = \frac{-\beta_2}{2\beta_3}.
$$

We used Nonlinear Seemingly Unrelated Regressions (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,

$$
H_0 : \Sigma \text{ is diagonal; } H_1 : \Sigma \text{ is not diagonal.}
$$

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 non-linear OLS (NLOLS). The estimated residual variances and covariances were obtained from the OLS estimates of (15) and (22).

We tested the null hypothesis using a Breusch Pagan, Lagrange multiplier (LM), test.

$$LM = N \sum_{m<j} \sum_{j=1}^{M} r_{mj}^2 \rightarrow \chi_{\frac{M(M-1)}{2}}^2, \tag{33}$$

where M represents the number of equations in the system and r^2 is the simple squared correlation between the residuals, u_1 and u_3. The LM test is based on the restricted model where Σ has non-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 separately by Nonlinear OLS (NLOLS) producing consistent but biased results with no loss in efficiency.

Next, we turn to testing the restriction. Specifically, we test,

$$H_0 : \gamma_6 = \frac{-\beta_2}{2\beta_3}; \ H_1 : \gamma_6 \neq \frac{-\beta_2}{2\beta_3}. \tag{34}$$

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

Table 5, columns 8-11, provide the restricted NLSUR results for (15), (17), (19), and (22). All of the estimates from (15), with the exception of the schooling and the schooling-ability interaction term, increase in 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 increase in magnitude. The coefficient estimates on (17), derived from (15), are of the expected signs and increase in statistical significance. The derived coefficient estimates on (19), from (15) and (22), are of the same sign as those from the unrestricted OLS estimates, but the magnitudes differ somewhat. The t-statistics suggest that some significance is lost and this may be due to the fact that (19) is not directly

part of the constrained system of equations. The estimated coefficients on (22) are nearly identical to those from unrestricted OLS. One could consider estimating a three-equation system (i.e. (15), (19), and (22)) by NLSUR. However, this strategy is not feasible because the variance/covariance matrix is singular.

4. $Corr(A, u_3) \neq 0$?

Measures of ability pose continuing problems for researchers and labor economists in particular. 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), pp. 20]. Some researchers (e.g., Ashenfelter and Krueger, 1994) have devised clever 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 KWW and AFQT scores used in this paper proxy for "true, innate" ability.[15] Both of these tests were administered in the teenage or early adult years of our respondents' lives and are also quite particular as to what they are testing. The AFQT score comes from the Armed Services Vocational Aptitude

_____

[15]Lazear (1977) attempts to purge the use of the KWW in his NLSY66 study by instrumenting for it with the following variables: schooling, schooling squared, parental education levels, a race dummy, and the median income for the father. In an earlier version of this paper, we similarly attempted to "purge" our ability measures of any outside influences. We did not pursue this avenue due to the poor results obtained using such a method.

Battery (ASVAB) 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 $corr(A, u_3) \neq 0$, e.g., simultaneity bias, omitted variables bias, etc. 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), the respondent's family size, and a set of occupational dummies for the adult present in a respondent's home when he was 14 years old.[16] 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 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 14 years old. 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 constructed a set of 12 occupational dummies based on the 1970 Census of the Population's Occupational Classification System. Regressing AFQT on these instruments yielded mainly statistically significant results; the estimated coefficients on inverse age and family size were negative. The estimated coefficients on the occupational dummies for the most part were statistically significant.

---

[16]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 14 and 17 years when the test was administered, and hence still in school. Doing so eliminates any causal effect of schooling on ability.

We tested for the potential correlation that exists between A and u_3 using a Hausman-Wu test [Greene, 2000]. We test the following hypothesis,

$$H_0 : \text{ plim} \left(\hat{\gamma}^{OLS} - \hat{\gamma}^{2SLS} \right) = 0; \ H_1 : \text{ plim} \left(\hat{\gamma}^{OLS} - \hat{\gamma}^{2SLS} \right) \neq 0. \tag{35}$$

Under the null hypothesis OLS and 2SLS produce consistent estimates of γ but OLS is asymptotically efficient. Under the alternative hypothesis, OLS is not consistent but 2SLS is. Wu (1973) suggests the use of a t-test, or an F-test in the multivariate case, as an alternative to the traditional Hausman test on the estimated coefficient of γ_7 in the augmented regression,

$$S_j = \gamma_0 + \gamma_1 S_{fj} + \gamma_2 S_{mj} + \gamma_3 DVS_{fj} + \gamma_4 DVS_{mj} + \gamma_5 N_j + \gamma_6 A_j + \gamma_7 \hat{A}_j + u_{3j}^*, \tag{36}$$

where $\widehat{A}$ is the fitted value of A from the first-stage instrumental variables regression. If $\widehat{\gamma}_7$ is statistically significant, then A is correlated with u_3 and 2SLS is appropriate. Based on the regression results $\hat{\gamma}_7$ is statistically insignificant. We conclude that our ability proxy, AFQT, is uncorrelated with u_3 (and hence exogenous to the system).

We also tested for the correlation of A and u_3 in the constrained system of equations. We test the statistical significance of $\widehat{\gamma}_7$ in,

$$
\begin{aligned}
S_j &= \gamma_0 + \gamma_1 S_{fj} + \gamma_2 S_{mj} + \gamma_3 DVS_{fj} + \gamma_4 DVS_{mj} + \gamma_5 N_j + \gamma_6 A_j + \gamma_7 \hat{A}_j + u_{3j}^* \\
\ln Y_j &= \beta_0 + \beta_1 S_j + \beta_2 A_j S_j + \beta_3 S_j^2 + u_{1j}
\end{aligned} \tag{37}
$$

subject to

$$\gamma_6 = \frac{-\beta_2}{2\beta_3}.$$

We again conclude that A is uncorrelated with u_3 (and hence exogenous to the

system), even in the context of a constrained system of equations, because $\widehat{\gamma}_7$ is not statistically significant.[17]

VII. CONCLUDING REMARKS

This paper develops a model of earnings and optimal schooling. The analysis and estimation strategy is inspired by the Mincerian schooling model. The coefficient on schooling in the simple model (14) generally overstates the returns because it does not control for ability. In addition the simple 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 abilities 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 collected from the NLSY79, we stratify our sample into one-year 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 over-

---

[17]Kelejian (1971) outlines an estimation procedure for structural equations that are linear in parameters but whose regressors are nonlinear functions of endogenous and predetermined variables. It would appear that the endogenous nature of S in (22) might warrant a closer look at the role it plays in (15). Kelejian's nonlinear 2SLS (N2SLS) requires S to be uncorrelated with u_1. However, because estimation strategy #3 indicates that u_1 and u_3 are uncorrelated, OLS is fine because we have a recursive, not a simultaneous, system of equations.

time). Because we impose a FTE status, our measures of work experience do not necessarily correspond to an actual calendar year. Based on the estimates of (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 abilities and schooling investments.

Based on our empirical findings we conclude that we have a constrained system of equations relating earnings determination and optimal schooling. The error term in the log earnings function is normally distributed and is not correlated with the error term in the optimal level of schooling equation. According to the Hausman-Wu test of exogeneity, we cannot reject the hypothesis that measured ability (AFQT) is exogenously determined and hence is uncorrelated with u_1 and u_3. Thus, our most preferred set of estimates correspond to columns 8-11 of Table 5.

According to Mincer's rule of thumb (1/overtaking year), 13 years of FTE work experience corresponding to 11 years beyond the completion of schooling yield approximate rates of return of 7.7 percent and 9.1 percent. Our model estimates that the (average) marginal rate of return to schooling is 9.6 percent and the optimal level of schooling is 11.4 years. Our estimate of the rate of return to schooling is consistent with past findings.

REFERENCES

Agnarsson, S. and P.S. Carlin, 2002, "Family Background and the Estimated Return to Schooling," *Journal of Human Resources*, 37, pp. 680-692.

Altonji, J.G. and T.A. Dunn, 1996, "The Effects of Family Characteristics on the Return to Education," *Review of Economics and Statistics*, 78, pp. 692-703.

Arkes, J., 1999, "What Do Educational Credentials Signal and Why Do Employers Value Credentials?" *Economics of Education Review*, 18, pp. 133-141.

Ashenfelter, O. and A. Krueger, 1994, "Estimates of the Return to Schooling from a New Sample of Twins," *American Economic Review*, 84, pp. 1157-1173.

Ashenfelter, O. and C. Rouse, 1998, "Income, Schooling, and Ability: Evidence from a New Sample of Identical Twins," *Quarterly Journal of Economics*, 113, pp. 253-284.

Ashenfelter, O. and D.J. Zimmerman, 1997, "Estimates of the Returns to Schooling from Sibling Data: Fathers, Sons, and Brothers," *Review of Economics and Statistics*, 97, pp. 1-9.

Barron, J.M, M.C. Berger, and D.A. Black, 1998, "Do Workers Pay for On-the-Job Training?" *Journal of Human Resources*, 34, pp. 235-252.

Becker, G.S., 1962, "Investment in Human Capital: A Theoretical Analysis," *Journal of Political Economy*, 70, pp. 9-49.

Becker, G.S. and B.R. Chiswick, 1966, "Education and the Distribution of Earnings," *American Economic Review*, 56, pp. 358-369.

Behrman, J.R. and N. Birdsall, 1983, "The Quality of Schooling: Quantity Alone is Misleading," *American Economic Review*, 73, pp. 928-946.

Behrman, J.R. and M.R. Rosenzweig, 1999, "Ability Biases in Schooling Returns and Twins: A Test and New Estimates," *Economics of Education Review*, 73, pp. 928-946.

Blaug, M., 1962, *Economic Theory in Retrospect* (Richard D. Irwin, Inc., Homewood, IL).

Cameron, S. and C. Taber, 2004, "Estimation of Educational Borrowing Constraints Using Returns to Schooling," *Journal of Political Economy*, 112, pp. 132-182.

Cameron, S.V. and J.J. Heckman, 1993, "The Nonequivalence of High School Equivalents," *Journal of Labor Economics*, 11, pp. 1-47.

Cameron, S.V. and J.J. Heckman, 1998, "Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males," *Journal of Political Economy*, 106, pp. 262-333.

Card, D., 1995, "Earnings, Schooling, and Ability Revisited." *Research in Labor Economics*, 14, pp. 23-48.

Card, D., 2000, "The Causal Effect of Education on Earnings," *Handbook of Labor Economics* (O. Ashenfelter and D. Card eds., North-Holland, Amsterdam).

Card, D. and A.B. Krueger, 1992, "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States," *Journal of Political Economy*, 100, pp. 1-40.

Diamond, J., 1999, *Guns, Germs, and Steel: The Fates of Human Societies* (W.W. Norton & Company, New York).

Ferrer, A.M. and W.C. Riddell, 2002, "The Role of Credentials in the Canadian Labour Market," *Canadian Journal of Economics*, 35, pp. 879-905.

Frazis, H., 1993, "Selection Bias and the Degree Effect," *Journal of Human Resources*, pp. 538-554.

Greene, W.H., 2000, *Econometric Analysis*, 4th edition, (Prentice-Hall, Inc., New Jersey).

Griliches, Z., 1976, "Wages of Very Young Men," *Journal of Political Economy*, 84, pp. S69-S86.

Griliches, Z., 1977, "Estimating the Returns to Schooling: Some Econometric Problems," *Econometrica*, 45, pp. 1-22.

Griliches, Z., 1979, "Sibling Models and Data in Economics: Beginnings of a Survey," *Journal of Political Economy*, 87, pp. S37-S64.

Griliches, Z. and W.M. Mason, 1972, "Education, Income, and Ability," *Journal of Political Economy*, 80, pp. S74-S103.

Hause, J.C., 1972, "Earnings Profile: Ability and Schooling," *Journal of Political Economy*, 80, pp. S108-S138.

Hungerford, T. and G. Solon, 1987, "Sheepskin Effects in the Returns to Education," *Review of Economics and Statistics*, 69, pp. 175-177.

Jaeger, D.A. and M.E. Page, 1996, "Degrees Matter: New Evidence on Sheepskin Effects in the Returns to Education," *Review of Economics and Statistics*, 78, pp. 733-740.

Judge, G.G., R.C. Hill, W.E. Griffiths, H. Lutkephol, and T.C. Lee, 1988, *Introduction to the Theory and Practice of Econometrics*, 2nd edition, (John Wiley & Sons, New York).

Kane, T.J. and C.E. Rouse, 1995, "Labor-Market Returns to Two- and Four-Year Colleges," *American Economic Review*, 85, pp. 600-614.

Kelejian, H.H., 1971, "Two-Stage Least Squares and Econometric Systems Linear in Parameters but Nonlinear in the Endogenous Variables," *Journal of the American Statistical Association*, 66, pp. 373-374.

Kennedy, P., 1998, *A Guide to Econometrics*, 4th edition, (The MIT Press, Cambridge, MA).

Kling, J.R., 2001, "Interpreting Instrumental Variables Estimates of the Returns to Schooling," *Journal of Business & Economic Statistics*, 19, pp. 358-364.

Lang, K., 1993, "Ability Bias, Discount Rate Bias, and the Return to Education," *Boston University Working Paper*.

Lang, K. and J.L. Zagorsky, 2000, "Does Growing Up with a Parent Absent Really Hurt?" *Journal of Human Resources*, 36, pp. 253-273.

Lang, K. and P.A. Ruud, 1986, "Returns to Schooling, Implicit Discount Rates and Black-White Wage Differentials," *Review of Economics and Statistics*, 68, pp. 41-47.

Lazear, E., 1977, "Education: Consumption or Production?" *Journal of Political Economy*, 85, pp. 569-598.

Maddala, G.S., 2001, *Introduction to Econometrics*, 3rd edition, (John Wiley & Sons, Ltd., Chichester, England).

Mincer, J., 1974, *Schooling, Experience, and Earnings*, (National Bureau of Economic Research, New York).

Murphy, K.M. and F. Welch, 1990, "Empirical Age-Earnings Profiles," *Journal of Labor Economics*, 8, pp. 202-229.

Neumark, D., 1999, "Biases in Twin Estimates of the Return to Schooling," *Economics of Education Review*, 18, pp. 143-148.

Rosen, S., 1974, "Human Capital and the Internal Rate of Return," *Industrial Relations Research Association Series*, pp. 243-250.

Rouse, C.E., 1999, "Further Estimates of the Economic Return to Schooling from a New Sample of Twins," *Economics of Education Review*, 18, pp. 149-157.

Shea, J., 2000, "Does Parents' Money Matter?" *Journal of Public Economics*, 77, pp. 155-184.

Willis, R.J. and S. Rosen, 1979, "Education and Self-Selection," *Journal of Political Economy*, 87, no. 5, pp. S7-S36.

Wu, D., 1973, "Alternative Tests of Independence between Stochastic Regressors and Disturbances," *Econometrica*, 41, pp. 733-750.

Appendix

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

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

lnY = ln(iV*) + iS*

lnY = ln[F(S,A)]

lnY*

$\ln(iV_l)$

$\ln(iV^*)$

$\ln(iV_0)$

S

S*

lnY

s

d

S

S*
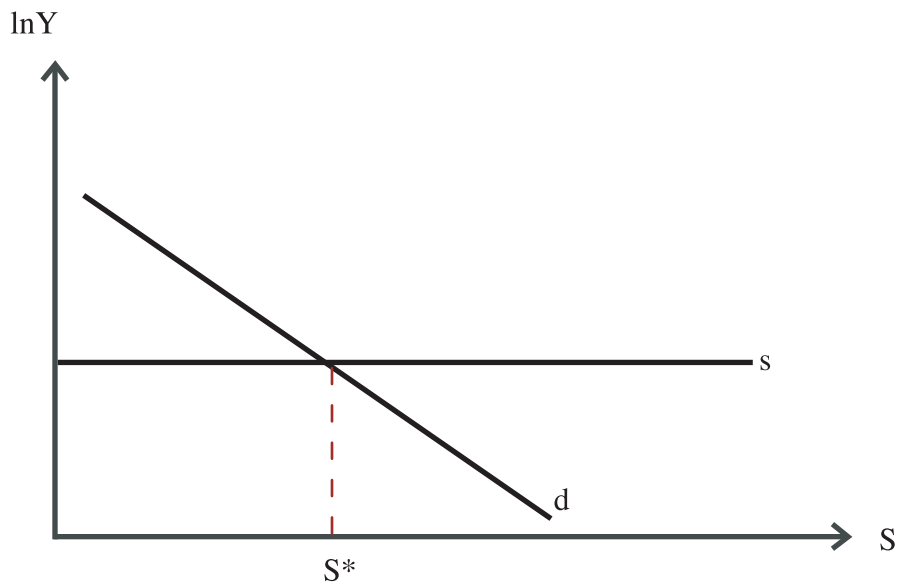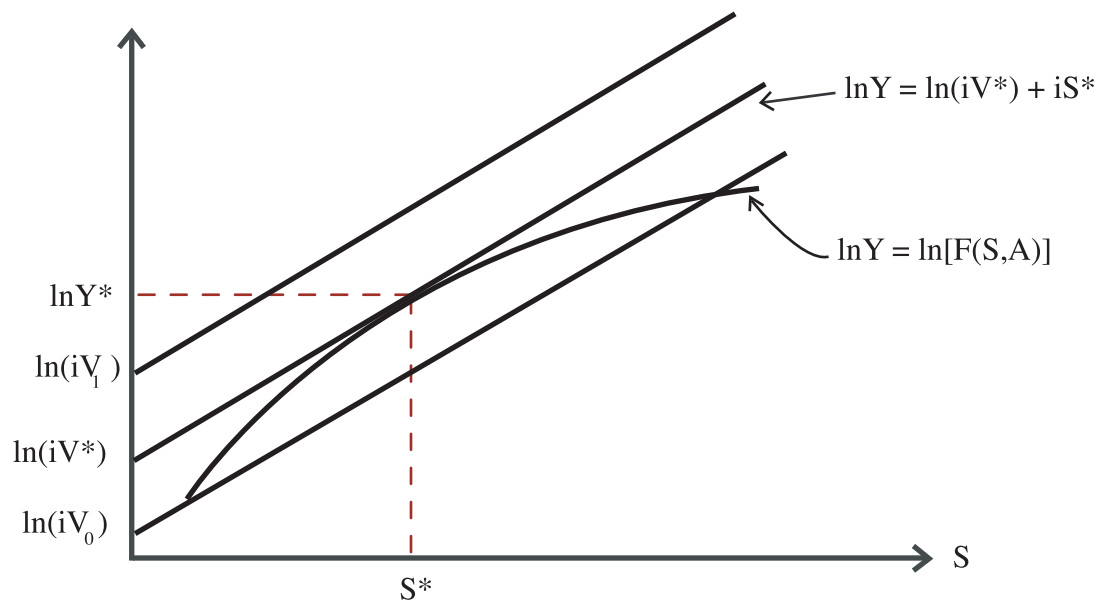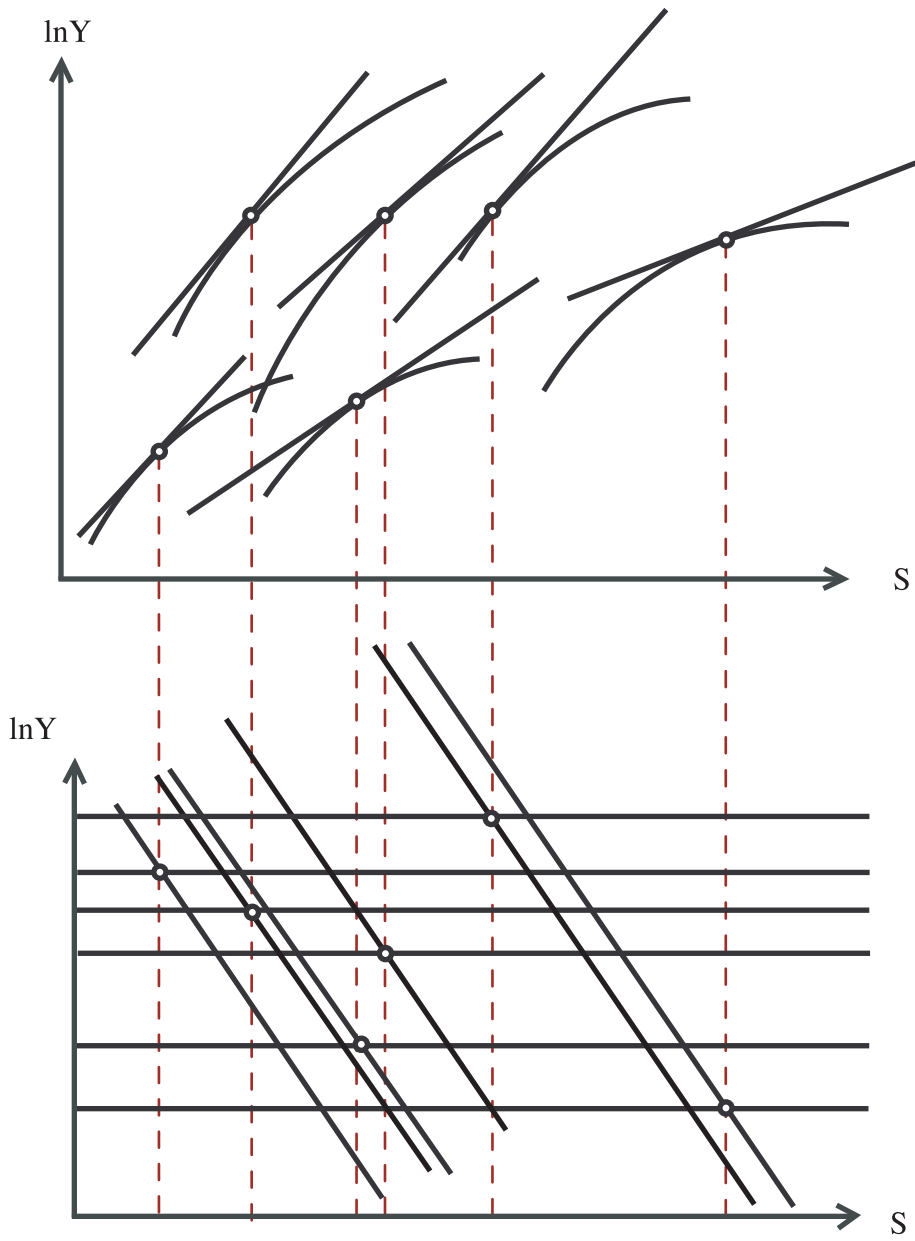
Figure 1

Figure 2

TABLE 1
DESCRIPTIVE STATISTICS FOR OVERALL SAMPLE

	Mean	Std.Dev.	Nobs.
AGE 1979	17.948	2.336	4393
AGE 1980	18.906	2.320	4393
NOMINAL WAGE 1985	12,717.90	10,721.20	3499
NOMINAL WAGE 1986	15,394.90	12,535.40	3370
NOMINAL WAGE 1987	17,553.20	13,584.00	3425
NOMINAL WAGE 1988	20,892.90	32,787.40	3405
NOMINAL WAGE 1989	21,992.50	18,249.70	3368
SCHOOLING 1985	12.525	2.354	3622
SCHOOLING 1986	12.636	2.455	3516
SCHOOLING 1987	12.747	2.510	3449
SCHOOLING 1988	12.812	2.594	3468
SCHOOLING 1989	12.848	2.621	3490
KWW	6.208	2.125	4393
AFQT	48.856	29.385	4087
EXPERIENCE 1985	4.073	3.071	4393
EXPERIENCE 1986	4.748	3.350	4393
EXPERIENCE 1987	5.443	3.638	4393
EXPERIENCE 1988	6.162	3.941	4393
EXPERIENCE 1989	6.908	4.256	4393
MOTHER'S SCHOOLING*	11.186	3.152	4139
FATHER'S SCHOOLING*	11.396	3.944	3978
FAMILY SIZE 1979	3.992	2.162	4393

*=as reported by the NLSY79

Source of data: NLSY79

TABLE 2
WORK EXPERIENCE COHORT FREQUENCY DISTRIBUTION: 1985-1989

FTE Years of Work Experience 1985-1989	Nobs.	% of sample
0	206	4.69%
1	483	10.99%
2	728	16.57%
3	997	22.70%
4	1157	26.34%
5	1221	27.79%
6	1186	27.00%
7	1060	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%

note: sample is based on those individuals whose wages > $500 and whom
are not enrolled in school currently or anytime after 1989

Source of data: NLSY79

TABLE 3
EQ 15 LOG EARNINGS FUNCTION: SELECTION CRITERION

ABILITY MEASURE:				AFQT		
FTE Work Experience Cohort	Nobs.	AIC	SC	PC	S.E.E.	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	0.290	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**

AIC=Akaike Information Criterion
SC=Schwarz Criterion
PC=Amemiya's Prediction Criterion

Bolded figures correspond to the minimum AIC, SC, PC, and Std. Error.
Bolded figures correspond to the largest R^2.

note: samples are based on those individuals whose wages >$500 and whom are not enrolled
in school currently or anytime after 1989

Source of data: NLSY79

TABLE 4
DESCRIPTIVE STATISTICS 13 YEAR WORK EXPERIENCE COHORT

	Mean	Std.Dev.	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**	9.633	4.095	215
FATHER'S SCHOOLING**	9.047	4.733	215
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

note: sample is based on those individuals whose wages >$500 and
whom are not enrolled in school currently or anytime after 1989

*=as reported by the NLSY79
**=adjusted for missing values

Source of data: NLSY79

TABLE 5
ESTIMATED SCHOOLING MODEL

MODEL/ESTIMATION STRATEGY:			UNRESTRICTED/OLS				
COHORT:			13 FTE YEARS				
EQ:	14	15	17	19		22	
DEP. VARIAB.:	ln(earnings)		estimated r	estimated i		years of school completed	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
CONSTANT	8.295	7.619	0.290	0.151	0.153	7.317	7.420
	(43.832)***	(14.426)***	(2.965)***	(14.764)***	(4.023)***	(13.449)***	(2.406)**
SCHOOLING	0.140	0.290	-1.871E-02	---	---	---	---
	(8.528)***	(2.965)***	(-2.056)**				
AFQT*SCHOOLING	---	4.008E-04	---	---	---	---	---
		(3.241)***					
AFQT	---	---	4.008E-04	---	---	3.089E-02	2.142E-02
			(3.241)***			(6.738)****	(2.659)**
SCHOOLING^2	---	-9.356E-03	---	---	---	---	---
		(-2.056)**					
FATHER'S SCHOOLING	---	---	---	-2.442E-03	-1.970E-03	0.106	0.131
				(-3.717)***	(-2.064)**	(2.862)***	(1.455)
MOTHER'S SCHOOLING	---	---	---	-3.092E-03	-2.760E-03	0.147	0.165
				(-3.494)***	(-1.919)*	(3.088)***	(1.580)
FATHER'S SCHOOLING DUMMY	---	---	---	-2.851E-02	-2.499E-02	1.335	1.524
				(-3.062)***	(-1.856)*	(2.659)***	(1.496)
MOTHER'S SCHOOLING DUMMY	---	---	---	-1.089E-02	-9.800E-03	0.523	0.582
				(-0.945)	(-0.878)	(0.855)	(0.772)
FAMILY SIZE	---	---	---	-3.895E-05	-5.500E-03	2.962E-02	2.080E-03
				(-0.035)	(-0.034)	(0.495)	(0.504)
R^2	0.255	0.299	---	0.235	---	0.446	---
adj R^2	0.251	0.289	---	0.216	---	0.430	---
S.E.E.	0.490	0.477	---	2.912E-02	2.506E-02	1.544	1.562
Nobs.	215	215	215	215	215	215	215
Estimated at sample mean r, i	0.140	0.095	0.095	0.095	0.103	---	---
Estimated at sample mean optimal years of schooling	---	---	---	---	---	11.377	11.377

(t-statistic)
*,**,***=significant at the 10, 5, and 1% level

note: sample is based on individuals whose wages >$500 and are not enrolled in school currently or anytime after 1989
(5), (7)= derived form

Source of data: NLSY79

TABLE 5 CONTINUED
ESTIMATED SCHOOLING MODEL

MODEL/ESTIMATION STRAGEGY:	RESTRICTED/NLSUR			
COHORT:	13 FTE YEARS			
EQ:	15	17	19	22
DEP. VARIAB.:	ln(earnings)	estimated r	estimated i	years of school completed
	(8)	(9)	(10)	(11)
CONSTANT	7.898	0.236	0.134	7.324
	(35.786)***	(6.049)***	(8.086)***	(14.652)***
SCHOOLING	0.236	-1.395E-02	---	---
	(6.049)***	(-2.927)***		
AFQT*SCHOOLING	4.229E-04	---	---	---
	(3.119)***			
AFQT	---	4.229E-04	---	3.089E-02
		(3.119)***		(6.004)***
SCHOOLING^2	-6.977E-03	---	---	---
	(-2.927)***			
FATHER'S SCHOOLING	---	---	-1.494E-03	0.107
			(-1.995)*	(2.840)***
MOTHER'S SCHOOLING	---	---	-2.071E-03	0.148
			(-2.161)**	(3.882)***
FATHER'S SCHOOLING DUMMY	---	---	-1.879E-02	1.347
			(-2.010)**	(2.904)***
MOTHER'S SCHOOLING DUMMY	---	---	-7.354E-03	0.527
			(-0.872)	(0.944)
FAMILY SIZE	---	---	-3.897E-04	2.793E-02
			(-0.481)	(0.472)
R^2	0.297	---	---	0.446
adj R^2	---	---	---	---
S.E.E.	---	---	---	---
Nobs.	215	215	215	215
Estimated at sample mean r, i	0.096	0.096	0.096	---
Estimated at sample mean optimal years of schooling	---	---	---	11.403

(t-statistic)
*,**,***=significant at the 10, 5, and 1% level

note: sample is based on individuals whose wages >$500 and are not enrolled in school currently or anytime after 1989
(7)=derived reduced form

Source of data: NLSY79