

NBER WORKING PAPER SERIES

EARNINGS, SCHOOLING,
AND ABILITY REVISITED

David Card

Working Paper No. 4832

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 1994

I am grateful to Janet Currie, John Ham, Robert Hutchens, and Kenneth McLaughlin for helpful comments on earlier drafts and to Alan Krueger and Cecilia Rouse for many useful discussions on this topic. I also thank Jin Heum Park for his assistance in preparing the paper. This paper is part of NBER's research program in Labor Studies. Any opinions expressed are those of the author and not those of the National Bureau of Economic Research.

© 1994 by David Card. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

NBER Working Paper #4832
August 1994

EARNINGS, SCHOOLING,
AND ABILITY REVISITED

ABSTRACT

This paper presents a survey and interpretation of recent research on the return to education. The empirical findings in a series of current papers suggest that the causal effect of education on earnings is understated by standard estimation methods. Using a simple model of optimal schooling developed by Gary Becker (1967), I derive an explicit formula for the conventional estimate of the return to schooling and for alternative instrumental variables and fixed-effects estimators. The analysis suggests that instrumental variables estimates based on "interventions" that affect the schooling choices of children from relatively disadvantaged family backgrounds will tend to exceed the corresponding OLS estimates.

David Card
Department of Economics
Princeton University
Princeton, NJ 08544
and NBER

One of the most important "facts" about the labor market is that individuals with more education earn higher wages. Hundreds of studies from many different countries confirm that average wages are higher for more highly-educated workers (see e.g. Psacharopoulos (1985)). Despite this evidence, most economists are reluctant to interpret the earnings gap between more and less educated workers as an estimate of the causal effect of schooling. Education is not randomly assigned across the population; rather, individuals make their own schooling choices. Depending on how these choices are made, measured earnings differences between workers with different levels of schooling may over-state or under-state the causal effect of education.

This paper begins with a review of eight recent studies that attempt to estimate the causal effect of schooling. Five of the studies use an arguably exogenous source of variation in education outcomes to form an instrumental variables estimate of the return to education. Three others attempt to control for unobserved attributes that may confound a causal analysis of education and earnings. A simple tabulation of the estimates from these diverse studies reveals an unexpected finding -- in all but one case the "corrected" estimate of the effect of schooling is noticeably larger than the corresponding "uncorrected" (ordinary least squares) estimate. Contrary to conventional wisdom, these studies suggest that the causal effect of education is actually understated by a simple comparison of wages between more and less educated workers.¹

¹The conventional wisdom is summarized in Ehrenberg and Smith (1991, pp. 320-322). In his widely-cited 1977 survey, Griliches concluded that OLS estimates tend to give unbiased or even negatively biased estimates of the causal effect of education. It is interesting to speculate as to why the conventional wisdom is apparently at odds with Griliches' carefully argued conclusion.

Of course it is possible to criticize the identification strategy used in any of these studies. In the absence of a randomized experiment, inferences about the causal effects of education are necessarily derived from non-experimental methods based on untestable identification assumptions. Furthermore, in each of the studies, the "corrected" estimate of the return to education is relatively imprecise. In my opinion, however, the similarity of the findings across the studies suggests the need for a careful rethinking of the forces that determine individual schooling attainment, and the potential biases that endogenous school choice creates in alternative estimators of the return to education.

The second section of the paper presents a model of endogenous schooling based on a supply-demand framework developed by Becker (1967). Using a set of simplifying assumptions on functional forms and the distributions of unobservables, I derive the formula for the conventional estimate of the return to education (based on a simple regression of log wages on observed education). I then compare this estimator to alternatives based on the kinds of estimation procedures used in recent studies of the return to education. This analysis suggests that instrumental variables estimates based on "exogenous" factors that affect the schooling choices of children who would normally obtain low levels of schooling will tend to exceed the corresponding ordinary least squares estimates of the return to education.

The reasoning behind this conclusion is fairly simple. In Becker's model, individuals invest in schooling until the marginal return to schooling is equated with their marginal discount rate. The population of less-educated workers therefore consists of a mixture of individuals with low returns to education (less able individuals) and individuals with high discount rates (individuals from poorer families, or with a stronger distaste for education). If

a certain intervention (such as a compulsory schooling law, or the introduction of a nearby college) induces people from the low-education group to increase their schooling, the associated "return" will reflect the marginal returns to schooling for the low-education group. This marginal return may well exceed the average return to schooling for the population as a whole if most of the people with low education have high discount rates rather than low ability.²

I. A Survey of Recent Estimates of the Return to Education

Table 1 presents a brief survey of eight recent papers that attempt to estimate the causal effect of schooling on earnings.³ The studies are divided into two groups based on methodological approach: the first group of studies use instrumental variables methods; the second group employ variants of a fixed effects estimator.

As a framework for the interpretation of these studies, consider a naive structural model of schooling and earnings:

$$(1a) \quad S_i = X_i\gamma + v_i$$

$$(1b) \quad \log y_i = X_i\beta + S_i\rho + u_i$$

Here S_i refers to the years of education of individual i , y_i refers to a measure of earnings, X_i represents a vector of control variables (including for example age, region, and race), and u_i and v_i represent a pair of residuals. The coefficient ρ in equation (1b) is the causal effect of

²Lang (1993) has labelled this phenomenon as "discount rate bias".

³Most of the pre-1986 literature is surveyed in Willis (1986), and much of the pre-1976 literature is surveyed in Griliches (1977) and Rosen (1977).

education: it gives the expected percentage gain in earnings if a randomly selected member of the population were to receive an additional year of schooling.⁴

It is well known that ordinary least squares (OLS) estimation of equation (1b) gives rise to a consistent estimate of ρ if and only if u_i and v_i are uncorrelated. There are a variety of reasons why the unobserved determinants of schooling may be positively or negatively correlated with the unobserved determinants of earnings, including the effects of unobserved ability and measurement error in schooling (see below). One strategy for dealing with this correlation is to identify a set of variables that affect schooling but not earnings (controlling for schooling). These variables can then be used to form instrumental variables (IV) estimates of the return to education. This is the procedure followed in the first 5 studies in Table 1.

Angrist and Krueger (1991a) use as an instrument for education the quarter of an individual's birth. There is a small but systematic quarterly pattern in completed schooling attainment for men born in the 1930s-1950s. Angrist and Krueger attribute this pattern to compulsory schooling laws. In most states, people born in the same calendar year start school at the same time. As a result, men born earlier in the year reach the minimum school-leaving age at a lower grade than men born later in the year. Their empirical analysis

⁴In many models of education (including the one developed in Section II, below) the effect of an additional year of education varies across individuals, and may vary within individuals depending on their years of schooling. This introduces a variety of issues of interpretation of "the" causal effect of education -- see Angrist and Imbens (1993) for a general discussion of this issue. It is also worth emphasizing that equations (1a) and (1b) constitute (at best) a partial equilibrium model: an intervention that affects the schooling levels of a significant fraction of the population might effect ρ . Finally, I make no distinction between a productivity-based interpretation of ρ and a signalling interpretation.

confirms that people born early in the year have relatively low levels of both schooling and earnings. Assuming that quarter of birth exerts no independent effect on earnings, it can thus serve as a legitimate instrument for education. As shown in the first row of Table 1, the IV procedure leads to 28% higher estimated return to education than the corresponding OLS procedure.

Angrist and Krueger (1991b) use a different IV procedure based on the lottery numbers assigned during the Viet Nam era draft. Since enrolled students could obtain draft exemptions, many analysts have argued that the draft lottery led to higher college enrollment rates, particularly for men whose lottery numbers implied the highest risk of induction.⁵ Angrist and Krueger use the lottery numbers (which were based on month and day of birth) to instrument years of education and veteran status for men in the age cohorts eligible for the draft. The resulting IV estimate of the return to education is some 10% above the corresponding OLS estimator.⁶

Butcher and Case (1993) analyze the effect of sibling composition on educational attainment. Since each child's sex is random, the sex composition of a family of given size is randomly determined. They offer a variety of evidence suggesting that women with one or more sisters have significantly less education than women from the same-sized families with only brothers. Butcher and Case then use an indicator variable for the presence of sisters as

⁵E.g. Bowen and Rudenstein (1992, chapter 3).

⁶In subsequent work Angrist and Krueger have found that the first stage equation in this specification shows only a marginally significant effect of lottery numbers on education. In light of this result the IV estimate should be interpreted with caution, since the IV estimator may be biased toward the OLS estimator when the instruments are "weak" (see Angrist and Krueger (1993)).

an instrument for education in a wage equation that also includes controls for family size. Although the resulting estimate of the return to education is relatively imprecise, it is substantially (100%) above the corresponding OLS estimate.

Kane and Rouse's (1993) study is primarily concerned with the relative labor market valuation of credits from regular (4-year) and junior (2-year) colleges. Their findings suggest that credits from the two types of colleges are interchangeable: given this conclusion they measure schooling in terms of total college credits. In analyzing the earnings effects of college credits, they compare OLS specifications against IV models that use the distance to the nearest 2-year and 4-year colleges and state-specific tuition rates as instruments. Their IV estimates of the return to college credits are 13-50% above the corresponding OLS specifications.

Finally, Card (1993) uses a simple indicator for the presence of a nearby college as an instrument for schooling. Men who grow up near a 4-year college have significantly higher education and earnings than other children. When college proximity is used as an instrument for schooling, the resulting IV estimator is 80% above the corresponding OLS estimator. As one might expect, most of the effect of college proximity is concentrated on children from poorer families. This pattern suggests an alternative specification that uses the interaction of college proximity with poor family background as an instrument for schooling. This instrument allows a test of the assumption (implicit in Kane and Rouse and in the simple IV results in Card) that college proximity has no direct effect on wages. The IV results from this alternative specification are 30% above the OLS results and give no indication that college proximity has a direct effect on wages, controlling for education.

An alternative to the use of instrument variables for educational attainment is to assume that the residuals components u_i and v_i in equations (1a) and (1b) have a restrictive covariance structure, and then use either repeated observations over time for the same individual or observations for different individuals from the same family to "difference out" the correlation between u_i and v_i . Variants of this methodology are pursued in the three studies summarized in panel II of Table 1.

Angrist and Newey (1991) assume that the period-specific error component of the wage equation for individual i in period t , u_{it} , and the corresponding period-specific component of the schooling equation, v_{it} , can be decomposed as:

$$u_{it} = \alpha_i + \eta_{it}$$

$$v_{it} = \theta_i + \xi_{it}$$

They further assume that any correlation between u_{it} and v_{it} arises through the time-invariant components (i.e. $E[\eta_{it}\xi_{it}] = 0$). Provided that some individuals acquire more schooling over time, these assumptions imply that ρ can be consistently estimated by a standard fixed effects estimator.⁷ An important limitation of this identification strategy is the assumption that earnings of individuals who have not yet completed schooling fully reveal the value of the fixed effect α_i . One might argue that individuals who are still in school take on part-time or

⁷As noted by Hausman and Taylor (1981), it may be possible to estimate ρ even if schooling is fixed over time if the means of certain time-varying covariates are orthogonal to the unobserved component of wages and also correlated with schooling. In their analysis they consider using means of indicators for the incidence of unemployment and bad health, as well as the mean of potential experience, as instruments for schooling. Their IV results show a large (over 100%) increase in the estimated return to education relative to an OLS specification.

"dead-end" jobs that do not reward the kinds of skills or attributes that affect completed schooling and post-graduate earnings.⁸

Angrist and Newey's results (in row 6 of Table 1) show a substantially higher return to education from the fixed effect model than the corresponding OLS model. This is even more remarkable in view of the fact that any measurement error in schooling is likely to be exacerbated in the fixed effects analysis, leading to a downward measurement error bias in the fixed effects specification relative to OLS (see below). Angrist and Newey's findings, however, suggest that other sources of bias in the OLS and fixed effects specifications dominate the relative attenuation effect of measurement error.

The assumptions underlying the analysis in Ashenfelter and Krueger (1994) are similar to the assumptions in Angrist and Newey, although Ashenfelter and Krueger use wage and schooling data for identical twins observed at the same point in time, rather than repeated observations for the same individual over time. Thus, Ashenfelter and Krueger assume that the error components in wages and schooling for twin j from family i can be decomposed as:

$$u_{ij} = \alpha_i + \eta_{ij}$$

$$v_{ij} = \theta_i + \xi_{ij}$$

They further assume that any correlation between u_{ij} and v_{ij} derives from a correlation between α_i and θ_i (i.e., the family effects). Provided that at least some pairs of twins have

⁸One could also argue that schooling choices are correlated with transitory wage disturbances. For example, individuals with negative transitory wage disturbances may decide to enroll in school. Something like this behavior is well-documented in the training literature (Lalonde (1993)).

different levels of schooling, ρ can be estimated by a "within-family" estimator analogous to the standard fixed effects estimator.

As in Angrist and Newey, Ashenfelter and Krueger find that the use of a fixed effects estimator leads to a rise in the estimated return to education relative to the OLS estimator. They also implement an instrumental variables/within-family estimator, based on the twins' responses about each others' levels of education. If the measurement errors in the twins' reports of education are uncorrelated, this estimator eliminates any downward bias associated with measurement error. The resulting within-twin instrumental variables estimator is substantially higher than either the OLS or unadjusted within-family estimator: on the order of the estimates reported by Butcher and Case (1993) and Card (1993). Ashenfelter and Krueger also implement a more complex measurement-error corrected estimator that allows for correlation in the twin's reports of their own and their sibling's education. This estimator gives a slightly lower estimate than the naive within-family IV estimator, but still 57% above the OLS estimate.

Finally, Ashenfelter and Zimmerman (1993) apply similar methods as Ashenfelter and Krueger (1994) to data on brothers from the original NLS Young Men data file. Contrary to the results of Ashenfelter and Krueger, Ashenfelter and Zimmerman find that the estimated return to schooling from a within-family fixed-effects model is lower than the corresponding OLS estimate. Recognizing that the fixed-effects estimator is highly sensitive to measurement error in schooling, they also report adjusted estimates that incorporate an exogenous estimate of the fraction of noise in measured schooling. Assuming that 10 percent of the variance of

schooling is attributable to measurement errors, the adjusted within-family estimate is just below the OLS estimate.⁹

In my opinion, the estimates in Table 1 point toward two conclusions. First, as argued by Griliches in his 1977 survey, OLS estimates of the return to schooling seem to be biased downward relative to alternative estimators that account for the correlation between the unobserved determinants of wages and schooling. This is particularly true when the alternative estimators are either robust to measurement error in schooling (as is the case for the instrumental variables estimators reviewed in Table 1) or build in some parametric adjustment for measurement error. Second, and more tentatively, the downward bias in the OLS estimators is sizeable: at least 10 percent and perhaps more like 30%. The median percentage gap between the IV and OLS estimates for the 8 studies in Table 1 is 28% (Angrist and Krueger (1991a)) or 33% (using the smaller IV estimate from Card (1993)).

II. A Simple Model of Endogenous Schooling

This section lays out a very simple theoretical model of schooling and earnings that can be used to interpret some of the recent econometric studies of the return to education. The motivation for building and analyzing such a model is twofold. First, many analysts seem to have a strong a priori belief that the return to education from a conventional OLS regression is an upward-biased estimate of the causal effect of schooling. Thus, they tend to

⁹Ashenfelter and Zimmerman report corrected estimates assuming that either 6.7% or 20% of the cross-sectional variance of schooling is attributable to measurement error. I have interpolated their estimates to obtain an estimate of the within-family coefficient assuming that the fraction of measurement error is 10%.

reject the IV and fixed effects estimates in Table 1 as "implausible". One objective of this section is therefore to restate the canonical model of Becker (1967) in an analytically tractable form and underscore its implications for the endogeneity biases in a conventional human capital earnings regression. Second, I believe that further research on the return to education would probably benefit from a more explicit theoretical framework than the naive econometric model of equations (1a) and (1b). The Becker model provides an obvious place to start.

Assume that an individual chooses schooling to maximize a utility function U defined over average earnings per year (y) and years of schooling (S). To keep things as simple as possible, I assume that

$$U(y, S) = \log y - \phi(S)$$

where ϕ is an increasing convex function.¹⁰ The individual's opportunities are summarized by a function $y=g(s)$, representing the level of earnings available at each level of education.

The first-order condition for optimal schooling is:

$$\frac{g'(S)}{g(S)} = \phi'(S).$$

Assuming that $g(S)$ is log-concave, the optimal level of schooling equates the marginal rate of return to schooling with the marginal cost, as in Figure 1.

¹⁰The simplest derivation of this utility function assumes that the individual maximizes the discounted present value of income, discounts the future at a constant rate r , and earns nothing while in school. In this case $U(y, S) = \log(y) - rS$: see Willis (1986, pp. 552-553).

To make this model operational I assume that the marginal return to schooling ($g'(S)/g(S)$) and the marginal rate of substitution ($\phi'(S)$) are linear functions with person-specific intercepts and homogenous slopes:

$$(2a) \quad \frac{g'(S)}{g(S)} \equiv \beta_i(S) = b_i - k_1 S \quad (k_1 \geq 0)$$

$$(2b) \quad \phi'(S) \equiv \delta_i(S) = r_i + k_2 S \quad (k_2 \geq 0).$$

Figure 2 illustrates the marginal efficiency and marginal cost schedules for several different individuals, as well as the optimal levels of schooling S_i^* defined by

$$(3) \quad S_i^* = (b_i - r_i) / k, \quad \text{where } k = k_1 + k_2.$$

Note that $\beta_i(S)$ and $\delta_i(S)$ are measured in units of percentage points per year. Becker interprets $\delta_i(S)$ as the rate of return required for funds used to finance the S th year of schooling. Although this interpretation is difficult to justify analytically (unless $\delta_i(S)$ is constant), it has considerable intuitive appeal and I shall refer to $\delta_i(S)$ as i 's "marginal discount rate".

In this model schooling choices vary across individuals for two reasons: because individuals have different returns to schooling (i.e. variation in b_i); and because individuals have higher or lower marginal rates of substitution between schooling and future earnings (i.e. variation in r_i). Loosely, variation in b_i corresponds to variation in "ability" whereas variation r_i corresponds to variation in "access to funds" (family wealth) or in "tastes for

schooling".¹¹ Importantly, the relevant variation in ability concerns the slope of the log earnings function, rather than its intercept. Indeed, as was pointed out by Griliches (1977) in the context of a slightly different version of Becker's model, individuals with higher earnings opportunities at each level of education (i.e. with higher intercepts in their log earnings functions) may well invest less in schooling, since they have a higher opportunity cost of attending school.

For simplicity I assume that b_i and r_i are symmetric random variables with an arbitrary correlation. To the extent that "ability" is either inherited or directly affected by family background, one might expect that b_i and r_i are negatively correlated. To see this, note that higher ability parents will tend to acquire more education and earn higher incomes. Assuming that higher-income families have lower discount rates (or stronger tastes for education), their children will have lower marginal disutilities of schooling. Thus, if ability is partially inherited, the children of higher-ability parents will have higher b_i 's and lower r_i 's, while the children of lower-ability parents will have lower b_i 's and higher r_i 's.

Integration of (2a) leads to an equation for the log earnings of individual i :

$$(4) \quad \log y_i = a + b_i S_i - 1/2 k_1 S_i^2,$$

where a is a constant that may in principle vary across individuals (see below). Equations (3) and (4) together determine the joint distribution of earnings and schooling. Two aspects of

¹¹Note that I am abstracting from differences in the quality of schooling available to different individuals. Variation in school quality can be parameterized in this model by shifts in the distribution of b_i . Consistent with this interpretation, Card and Krueger (1992) show that the slope of the relation between education and earnings is positively related to school quality variables. Furthermore, they find a significant effect of school quality on the average level of schooling.

this joint distribution are of special interest: (1) is the graph of the cross-sectional relation between $\log y_i$ and S_i approximately linear? (2) what is the (population) regression coefficient of $\log y_i$ on S_i ?

The approximate linearity of the cross-sectional relation between log earnings and schooling is an important "stylized fact". Heckman and Plocheck (1974), Hungerford and Solon (1987), and Card and Krueger (1992) all present evidence suggesting that earnings are nearly log-linear with respect to schooling. Figure 3 presents some additional evidence on this issue taken from a recent paper by Park (1994). The figure shows the mean and various percentiles of log wages for single years of education in a sample of white men from the 1991 Current Population Survey.¹² Apart from a small "dip" at the 15th year of schooling (which is the focus of Park's paper) the graph is remarkably linear throughout the entire distribution of wages.¹³

At first blush it may seem that equation (4) is inconsistent with a linear relation between log wages and schooling unless the quadratic term is 0 (i.e. $k_1=0$). Because of the endogeneity of schooling, however, S_i and b_i are positively correlated across the population. Thus, in the absence of concavity in the relation between wages and schooling for a given level of ability, wages will be a convex function of schooling in the population as a whole.

Figure 4 illustrates the structure of the cross-sectional relation between log earnings and schooling. Among individuals with the same value of b_i , those with lower marginal

¹²The sample includes individuals age 34-65 who completed their last year of school.

¹³Park (1994) presents figures similar to Figure 3 for each year from 1979 to 1991, for various age ranges and for men and women. In every case the graph is approximately linear, although the slope of the log earnings-education relation varies from year to year.

discount rates choose higher levels of schooling, tracing out a concave earnings-schooling relation. Across individuals with differing abilities, those with higher levels of ability choose higher levels of schooling, tracing out a convex earnings-schooling relation. The overall cross-sectional relation is determined by a combination of within-group and between-group effects, and will tend to be more concave, the smaller is the variance of ability (b_i) relative to the variance of discount rates (r_i).

Ignoring other covariates, the theoretical regression coefficient (ρ) of $\log y_i$ on S_i can be derived as follows. By definition:

$$\begin{aligned}\rho &= \text{Cov}(\log y_i, S_i) \div \text{Var}(S_i) \\ &= E[\log y_i \cdot (S_i - \bar{S})] \div \text{VAR}(S_i)\end{aligned}$$

where \bar{S} represents the mean of S_i . Using equations (3) and (4),

$$\begin{aligned}E[\log y_i \cdot (S_i - \bar{S})] &= E[b_i S_i (S_i - \bar{S}) - 1/2 k_1 S_i^2 \cdot (S_i - \bar{S})] \\ &= E\left\{ b_i \frac{b_i - r_i}{k} \frac{(b_i - \bar{b}) - (r_i - \bar{r})}{k} - \frac{1}{2} k_1 S_i^2 (S_i - \bar{S}) \right\}\end{aligned}$$

where \bar{b} and \bar{r} denote the expectations of b_i and r_i , respectively. The variance of schooling is

$$\text{Var}(S_i) = \frac{1}{k^2} (\sigma_b^2 + \sigma_r^2 - 2 \sigma_{br})$$

where σ_b^2 and σ_r^2 represent the variances of b_i and r_i , respectively, and σ_{br} is their covariance. Assuming that b_i and r_i are symmetrically distributed,¹⁴ it is easy to show that the population regression coefficient is a weighted average of \bar{b} and \bar{r} :

$$(5) \quad \rho = (1-\alpha) \bar{b} + \alpha \bar{r},$$

where

$$\alpha = \frac{k_1}{k} - \lambda,$$

and

$$\lambda = \frac{\sigma_b^2 - \sigma_{br}}{(\sigma_b^2 - \sigma_{br}) + (\sigma_r^2 - \sigma_{br})}$$

is (loosely) the fraction of the variance of schooling attributable to variation in ability as opposed to variation in discount rates.¹⁵

In the simplest model of endogenous schooling where individuals maximize the discounted present value of earnings at fixed individual-specific discount rates (i.e. $k_2=0$), $\alpha = 1 - \lambda$, and the conventionally estimated return to schooling is

$$\rho = \lambda \bar{b} + (1-\lambda) \bar{r}.$$

When the marginal discount rate is increasing with further years of schooling the relative weight α in equation (5) may be outside of the $[0,1]$ interval. A necessary condition for $\alpha > 0$ is

¹⁴Symmetry implies that the third central moments of both b_i and r_i are zero. If b_i or r_i is non-symmetrically distributed, the formula for ρ contains a term in the corresponding third central moment.

¹⁵This interpretation is obvious if $\sigma_{br} = 0$. Otherwise, λ will be bounded by 0 and 1 as long as $\sigma_{br} < \sigma_b^2$.

$$(6) \quad k_1 \geq \lambda (k_1 + k_2) .$$

This inequality has a simple interpretation in terms of the covariation between the marginal rate of return to schooling β_i and the level of schooling S_i . Let $P(\beta_i | S_i)$ denote the least squares prediction of β_i , given observed schooling. Using the facts that $S_i = (b_i - r_i)/k$ and $\beta_i = b_i - k_1 S_i$, the least squares projection formula implies

$$P(\beta_i | S_i) = \bar{\beta} + (\lambda k - k_1)(S_i - \bar{S}) .$$

The condition that the OLS estimate ρ lies between \bar{r} and \bar{b} is therefore equivalent to the condition that the predicted marginal return to schooling (given observed schooling) is decreasing in S_i : in other words, that individuals with higher levels of schooling have lower marginal returns to schooling, on average.

A model such as the one described here raises an important conceptual question for applied work: what is "the" causal effect of education? If different individuals have different returns to education at the same level of schooling, or if each individual's return to schooling is strictly decreasing, then there is no unique causal effect of schooling. In the conceptual experiment in which a randomly selected individual is allocated an extra year of schooling, the expected increase in earnings is $\bar{\beta} = E(\beta_i)$, the average marginal return to education. The causal effect of an additional year of education for a particular individual, however, may be above or below $\bar{\beta}$.

Despite this ambiguity it is useful to compare the OLS estimate of the return to schooling with the average marginal return to schooling. Using the fact that:

$$\bar{\beta} = \bar{b} - k_1 \bar{S} \quad ,$$

equation (5) can be re-written as

$$(7) \quad \rho = \bar{\beta} + \lambda(\bar{b} - \bar{r}) .$$

An OLS regression of log earnings on schooling yields an upward-biased estimate of the average marginal return to schooling, with a bigger bias the larger is σ_b^2 (the variance in ability) relative to σ_r^2 (the variance in discount rates). The term $\lambda(\bar{b} - \bar{r})$ is an endogeneity bias that arises because people with higher marginal returns to education choose higher levels of schooling.

Implications for Instrumental Variables Estimates

What are the implications of this theoretical framework for the kinds of instrumental variables estimators of the return to schooling described in Table 1? Is it possible for the estimated rate of return from an instrumental variables (IV) procedure to exceed the corresponding OLS estimate? Consider an IV procedure based on a discrete indicator representing an "intervention" that affects one subsample of individuals (the treatment group) with no effect on another otherwise identical subsample (the control group). For example, the instrument may represent the event of growing up in a state with a higher compulsory schooling age (Angrist and Krueger (1991a)); or growing up in close proximity to a 4-year college (Card (1993)). Let \bar{y}_t represent the expectation of log earnings in the treatment group, let \bar{y}_c represent the expectation of log earnings in the control group, and let \bar{S}_t and \bar{S}_c represent the corresponding expectations of education in the two groups. Consider a pooled regression of log earnings on schooling, using treatment group status as an instrument for schooling. The resulting IV estimate of the return to schooling has probability limit

$$\rho_{iv} = \frac{\bar{y}_t - \bar{y}_c}{\bar{S}_t - \bar{S}_c}.$$

If all individuals in the population have the same (constant) marginal return to education then $\bar{y}_t - \bar{y}_c = \bar{\beta}(\bar{S}_t - \bar{S}_c)$, and $\rho_{iv} = \bar{\beta}$. The instrumental variables estimate that results from the intervention is therefore a consistent estimate of the average marginal rate of return to schooling in the population.

More generally, divide the treatment and control populations into subgroups $g=1, \dots, G$ with the property that individuals in each subgroup have (approximately) the same marginal return to schooling. Consider a "small" intervention that raises average schooling in subgroup g of the treatment population by the increment $\Delta \bar{S}_g$. Then

$$\rho_{iv} = \frac{\sum_g \Delta \bar{S}_g \beta_g w_g}{\sum_g \Delta \bar{S}_g w_g},$$

where β_g is the marginal return to schooling of subgroup g and w_g is the fraction of the population in subgroup g .¹⁶ If only one subgroup is affected by the intervention, ρ_{iv} will equal the marginal rate of return to schooling in that subgroup. Clearly, the instrumental variables estimator of the return to education from a particular intervention can exceed the conventional cross-sectional estimator if the intervention affects a sub-population with a sufficiently high marginal return to schooling.

¹⁶This argument requires that the effect of the intervention works only through the change in education (i.e. the intervention has no effect on earnings for individuals who do not change their schooling).

As noted above, subgroups with lower education in the absence of an intervention will tend to have higher marginal returns to schooling if the relative variation in ability is small (i.e. provided that condition (6) is satisfied). An intervention that affects individuals with below-average levels of education can therefore lead to an instrumental variables estimate of the return to education that exceeds the OLS estimate, provided that the variation in ability is small relative to the variation in discount rates. This provides a simple interpretation of the IV estimates of the return to schooling estimated by Angrist and Krueger (1991a) using differences in education and earnings induced by compulsory schooling legislation, and also of the estimates in Angrist and Krueger (1991b) based on the changes in education and earnings induced by the Viet Nam draft lottery.

IV procedures based on factors that lower the costs of schooling will similarly tend to yield relatively high estimates of the marginal return to schooling if most of the individuals affected by a change in costs are from families with high marginal discount rates. Since individuals invest in schooling until the marginal return is equated to the marginal discount rate, individuals from such families will stop investing at a point where the marginal return to schooling is relatively high. A program that reduces the cost of schooling for children from poorer family backgrounds will therefore tend to have a higher marginal return than suggested by a conventional (OLS) estimate of the return to education. This provides a simple interpretation of the findings in Kane and Rouse (1993) and Card (1993) based on college proximity and tuition costs.¹⁷

¹⁷It is more difficult to use the model to interpret the findings of Butcher and Case (1993) since it is unclear how sibling composition affects either the marginal returns or marginal costs of schooling.

Implications for Fixed Effects Estimates

The simple theoretical model developed above can also be used to interpret the findings of recent studies that use a fixed-effects type estimation methodology. Consider first the "within-person" estimation strategy of Angrist and Newey (1991). Under the assumptions of their model (including the critical assumption that the transitory components of earnings and schooling are orthogonal) the fixed effects estimator recovers an estimate of the marginal return to education for the subset of individuals who are observed working every year and increase their schooling over the sample period.¹⁸ I suspect that this subset is more likely to include men from poorer family backgrounds who begin the sample with relatively low levels of education. A simple interpretation of their decision to continue working and attend school part-time is that they lack the funds to continue full-time schooling. Other things equal, these individuals would be expected to have higher marginal returns to schooling (at least for the lower levels of schooling that they hold at the beginning of the sample period). An alternative is that some of the individuals realize while on the job that they have high marginal returns to schooling, and decide to acquire more education. In either case, the fact that an individual acquires more schooling while working suggests that the individual may have above-average marginal returns to schooling.¹⁹ Thus one might expect the return to schooling from Angrist and Newey's fixed effects procedure to exceed the OLS estimate based on the their entire sample.

¹⁸Angrist and Newey (1991) exclude full time students from their sample.

¹⁹The simple Becker model implicitly assumes that schooling is completed before work begins. Thus there is some ambiguity in the mapping of the model to situations where individuals are observed in school while working.

Consider next the studies based on differences in education of siblings (twins in Ashenfelter and Krueger (1994), brothers in Ashenfelter and Zimmerman (1993)). Assume that b_{ij} and r_{ij} , the intercepts of the marginal efficiency and marginal discount functions for sibling j from family i , can be decomposed as:

$$b_{ij} = b_j + b'_{ij}$$

$$r_{ij} = r_j + r'_{ij},$$

where the sibling-specific components are symmetrically distributed and orthogonal to the family components. Then the formulas derived above for the OLS regression coefficient ρ remain valid for the within-family fixed effects estimator, provided that the variances and covariances of the ability and discount terms are interpreted as within-family variances and covariances: i.e. $\text{var}(b'_{i2}-b'_{i1})$, $\text{var}(r'_{i2}-r'_{i1})$, etc.

Recall from equation (7) that

$$\rho = \bar{\beta} + \lambda (\bar{b} - \bar{r}),$$

where $\bar{\beta}$ is the mean marginal return to education (in the relevant population), λ is the relative fraction of the variance of schooling attributable to ability (versus discount rates or preferences), and \bar{b} and \bar{r} are the means of b_i and r_i . A comparison of the cross-sectional and within-family fixed effects estimators therefore depends on the relative magnitude of λ in the overall population and within families. One possibility is that variation in b_i is eliminated (or greatly attenuated) within families. According to the model developed here, this hypothesis implies that the within-family estimator should be below the cross-sectional estimator. In my opinion, however, a more likely hypothesis is that the relative variation in

r_i (which reflects differences in accessibility of funds and tastes for education) is reduced within-families. If this is the case, then schooling choices are more highly correlated with ability within families than across the overall population, leading to an upward bias in the within-family fixed effects estimator relative to the corresponding OLS estimate of the return to education. This hypothesis is consistent with the pattern of estimates reported by Ashenfelter and Krueger (1994) and to a lesser extent with the results of Ashenfelter and Zimmerman (1993).

Other Biases in the Measured Return to Education

The preceding analysis ignores two important issues that typically arise in the estimation of the return to schooling: (1) unobserved heterogeneity in the level of earnings, and (2) measurement error in observed schooling. Most discussions of "ability bias" begin with a model for the level of earnings that contains an individual-specific component (e.g. Griliches (1977)). This can be incorporated in equation (4) by adding a person-specific intercept a_i :

$$\log y_i = a_i + b_i S_i - 1/2 k_i S_i^2 .$$

The formula for the population regression coefficient of log earnings on schooling then includes a component representing the projection of a_i on S_i .

A correlation between a_i and S_i may arise for two reasons: a_i may be correlated with b_i (i.e., the intercept and slope of the individual-specific earnings function are correlated); or a_i may be correlated with r_i (i.e., potential earnings are correlated with family background or taste factors that determine individual discount rates). One might conjecture that a_i and b_i are

positively correlated whereas a_i and r_i are negatively correlated.²⁰ In this case, the OLS regression coefficient of $\log y_i$ on S_i is larger (more positive) than suggested by equation (7). On the other hand, an instrumental variables procedure should provide estimates that are "purged" of the correlation between schooling and a_i . Thus unobserved heterogeneity in the level of earnings will tend to lead to an upward bias in the OLS estimate of the return to schooling relative to an IV estimate based on a legitimate instrument.

The implications of an individual heterogeneity component a_i in the level of earnings for the fixed-effects studies reviewed in Table 1 depends on the type of fixed effect (family or individual) and on the relative impact of differencing on the variance components attributable to ability and discount factors. A person-specific fixed effect will eliminate a_i altogether: thus the within-person estimator of Angrist and Newey (1991) should be robust to this type of heterogeneity. A family effect will presumably eliminate some fraction of the person-specific fixed effect. However, the remaining sibling-specific component may be more or less correlated with within-family schooling outcomes, leading to an ambiguous bias relative to an OLS estimate.

A second potentially important source of divergence between alternative estimates of the return to education is measurement error in schooling. Studies of the measurement error in reported schooling from conventional micro surveys suggest that 10 percent of the variance

²⁰A positive correlation between a_i and b_i will arise if a_i is correlated with a measure of cognitive ability and if schooling and cognitive ability are "complements" (Hause (1972)). On the other hand, the "comparative advantage" finding of Willis and Rosen (1979) suggests that a_i and b_i may be negatively correlated. A negative correlation between a_i and r_i will arise if children of wealthier families face lower discount rates (or have stronger tastes for education) and if these children also tend to earn more because of better "connections" in the labor market.

in observed schooling is attributable to measurement error.²¹ In a univariate regression, this measurement error will lead to a 10 percent attenuation bias in the OLS regression coefficient. In a multivariate regression the degree of attenuation may be higher if the other covariates are correlated with the true level of schooling. In particular, the expected degree of attenuation in a multivariate model is $\theta_1 = (\theta_0 - R^2)/(1-R^2)$, where θ_0 is the ratio of the true variance of schooling to the total variance of observed schooling and R^2 is the R-squared of a regression of observed schooling on the included covariates. Many studies of the return to education include family background, race, region, and age variables that explain up to one-third of the variance of observed schooling. In such models the expected attenuation of the education coefficient may be as high as 15%, purely as a consequence of measurement error in reported education.

Assuming that the measurement errors in reported schooling for the same person over time or for different people from the same family obey the "standard" assumptions, the attenuating effect of measurement errors will be further accentuated in a fixed effects estimation strategy.²² Differencing over time or within families eliminates much of the true "signal" in measured education, while raising the variance of the measurement error of the observed differences. For this reason, Ashenfelter and Krueger and Ashenfelter and

²¹See e.g. Siegel and Hodge's (1968) analysis of measurement error in the 1960 Census.

²²The "standard" assumption is that the measurement error in one variable is orthogonal to the true value of that variable and to the measurement errors in other variables. There is evidence (Bound and Krueger (1991)) that measurement errors in reported earnings of the same individual are positively correlated over time and negatively correlated with the true value of earnings. It is an open question whether similar patterns hold in survey measures of education. Results in Ashenfelter and Krueger (1994) suggest that the measurement error in one sibling's report of his own and his sibling's education are positively correlated.

Zimmerman implement "corrected" fixed effects procedures that explicitly account for measurement error.

In contrast to its attenuating effect on OLS and fixed effects estimators, measurement error should not affect the estimated return to education from a valid instrumental variables procedure. Thus, the implicit correction for measurement error built into the IV estimates in Table 1 can account for a 10-15 percent rise in the return to schooling in moving from the OLS to the IV estimates.

To summarize the discussion so far, consider three alternative estimates of the return to education for a given cross-section: an OLS estimator based on a potentially noisy measure of schooling; a within-family fixed effect estimator; and an IV estimator based on the changes in earnings and schooling induced by an "exogenous" intervention that affects a subgroup g of the population. Recognizing the effects of measurement error and allowing for an individual heterogeneity component a_i in the level of earnings, the probability limit of the OLS estimator is:

$$(8) \quad \theta_0(\bar{\beta} + \lambda (\bar{b} - \bar{r}) + k \cdot [\sigma_{ab} - \sigma_{ar}] / [\sigma_b^2 + \sigma_r^2 - 2\sigma_{br}]) ,$$

where σ_{ab} and σ_{ar} denote the covariances of (a_i, b_i) and (a_i, r_i) , respectively. There are 3 sources of bias in the OLS estimator relative to the average marginal return to schooling ($\bar{\beta}$): (i) the "endogeneity" bias component $\lambda (\bar{b} - \bar{r})$ isolated in equation (7); (ii) heterogeneity bias attributable to unobserved components in the level of earnings; and (iii) attenuation bias due to measurement error.

The probability limit of the within-family estimator has the same form as equation (8), with the re-interpretation of the attenuation coefficient as the signal-to-total-variance ratio in

"within-family" schooling outcomes, and the re-interpretation of σ_b^2 , σ_r^2 , σ_{br} , σ_{ab} , and σ_{ar} as "within family" moments. Finally, the IV estimator has probability limit β_g , equal to the marginal return to education in the subgroup affected by the intervention.

Is There Evidence of Individual Heterogeneity in Returns to Education?

Consideration of the biases generated by measurement error and unobserved heterogeneity in the level of earnings suggests a number of reasons why the OLS, fixed-effects, and IV estimates of the rate of return to education will vary, even in the absence of endogeneity bias introduced by individual heterogeneity in the return to schooling. Since the analysis and interpretation of earnings and schooling data is greatly simplified when there is no individual heterogeneity in the slope of the earnings function, it is worth asking whether there is any need to study more complex models.

Consider the very simplest model with $\sigma_b^2 = \sigma_{br} = \sigma_a^2 = 0$ (i.e. no ability variation in either the slope or intercept of the earnings function) and $k_1 = 0$ (i.e. linearity of the earnings function). In this basis case each individual has constant returns to additional years of schooling and the causal effect of education is simply $\bar{\beta}$, the common marginal return to education. Differences in observed schooling arise solely from differences in discount rates, with each individual investing in education until his or her marginal discount rate is equal to $\bar{\beta}$. Under these assumptions the cross-sectional relation between log earnings and observed education is exactly linear. The probability limits of the 3 estimators are

$$\rho_{ols} = \theta_0 \cdot \bar{\beta} ,$$

$$\rho_{fe} = \theta_w \cdot \bar{\beta} ,$$

$$\rho_{iv} = \bar{\beta} ,$$

where θ_w denotes the "within-family" measurement error attenuation coefficient. Assuming no variation in ability and constant returns to schooling the only source of bias in the OLS or fixed effects estimators is measurement error in education. An IV estimator, on the other hand, provides a consistent estimate of $\bar{\beta}$. Therefore, correcting for attenuation bias, the three estimates of the return to education should be equal.

A variant of this basis case assumes that $k_1 = 0$ and $\sigma_b^2 = \sigma_{br} = 0$ (i.e., no ability component in the return to education), but introduces an unobserved ability component in the level of wages that is correlated with the determinants of schooling: $\sigma_a^2 > 0$ and $\sigma_{ar} < 0$. This is the model typically assumed in the literature on "ability bias" in the return to education, and is consistent with the naive econometric model specified by equations (1a) and (1b). Again, the cross-sectional relation between log earnings and schooling is exactly linear, and the causal effect of education is simply $\bar{\beta}$. However, the presence of individual heterogeneity in the level of earnings leads to the following probability limits for the three alternative estimators:

$$\rho_{ols} = \theta_0 \{ \bar{\beta} - k \cdot \sigma_{ar} / \sigma_r^2 \} ,$$

$$\rho_{fe} = \theta_w \{ \bar{\beta} - k \cdot \sigma_{ar}(w) / \sigma_r^2(w) \} ,$$

$$\rho_{iv} = \bar{\beta} ,$$

where $\sigma_{ar}(w)$ and $\sigma_r^2(w)$ denote the within-family covariance of a_i and r_i and the within-family variance of r_i , respectively. The OLS and fixed effects estimators yield upward-biased estimates of $\bar{\beta}$ to the extent that the intercept in the earnings function (a_i) is negatively correlated with the discount rate, and downward-biased estimates to the extent that $\theta_w \leq \theta_0 < 1$. Correcting for attenuation bias, the OLS and fixed effects estimates of the return to education will equal or exceed the IV estimate obtained from a legitimate IV procedure.

A final variant of the "no ability variation" model introduces a strictly concave education production function (i.e. $k_1 > 0$). In the absence of heterogeneity in the marginal return to education, concavity in the education production function should reveal itself in a simple scatter plot of observed wages and schooling. As noted earlier, however, such plots are remarkably linear. Thus the joint hypotheses of concavity in the education production function and no individual heterogeneity in the return to education are inconsistent with the observed linearity of the cross-sectional earnings-education relationship.

Is there enough evidence from the studies in Table 1 to reject the assumption of no heterogeneity in returns to education (together with the assumption of linearity in the education production function)? Assuming that the attenuation bias in OLS estimates of the return to education is 10-15 percent, the IV estimates in Table 1 should exceed the corresponding OLS estimates by no more than 10-15%. Any unobserved heterogeneity in the level of earnings that is positively correlated with schooling will further reduce this range. In fact, all of the IV estimates are at least 10% above the corresponding OLS estimates, and the median gap is 30%, although none of the IV estimates is significantly bigger than the corresponding OLS estimate.

Under these same assumptions the relative ranking of an OLS estimator and a within-family fixed effects estimator depends on the relative magnitude of the heterogeneity biases introduced by unobserved determinants of the level of earnings, and on the relative magnitude of the measurement error biases. If $|\sigma_{ar}(w)/\sigma_r^2(w)| \leq |\sigma_{ar}/\sigma_r^2|$, the heterogeneity bias is smaller in the fixed effects estimator. Note however that within-family differencing will tend to reduce both the variance of r_i and the covariance of a_i and r_i . Thus heterogeneity bias may be reduced or increased by within-family differencing. On the other hand, the attenuation bias induced by measurement error is presumably larger in the fixed effects estimator. If the relative measurement error biases dominate the comparison, then an uncorrected fixed effects estimate will be below the OLS estimate. This is the case in Ashenfelter and Zimmerman (1993) but not in Ashenfelter and Krueger (1994).²³

In my opinion, the results of the studies in Table 1 are not quite strong enough to reject the "no heterogeneity" assumption, together with assumption of linearity in the education production function. Although the relatively high IV and fixed effects estimates in some of the studies suggest the presence of heterogeneity in the education slopes, it is not possible to rule out the most elementary "no heterogeneity" model of the joint determination of schooling and earnings.

²³Both studies also present measurement-error-corrected fixed-effects estimates. The difference between the measurement-error-corrected within-family estimate and a measurement error-corrected OLS estimate gives an estimate of the relative magnitudes of σ_{ar}/σ_r^2 and $\sigma_{ar}(w)/\sigma_r^2(w)$. Assuming that the OLS estimate is attenuated by 10 percent, the difference between the measurement-error-corrected fixed effects estimate and the corrected OLS estimate is 0.039 in Ashenfelter and Krueger and -0.011 in Ashenfelter and Zimmerman. These studies thus give different impressions of the relative heterogeneity biases in fixed effects and OLS estimators.

One type of evidence that would be useful in evaluating the "no heterogeneity" model is more detailed information on the nature of the measurement errors in reported education. The presumption that measurement error can explain a 10-15 percent gap between OLS and IV estimates of the return to education relies on the untested assumption that measurement errors in schooling are uncorrelated with true schooling. If the measurement errors in education are mean-regressive, as is apparently true for the measurement errors in earnings (Bound and Krueger (1991)), then the gap explainable by measurement error is smaller and the potential role for models with individual-specific education slopes is greater.

Conclusions

This paper is motivated by a series of recent studies of the causal effect of education in the labor market. A survey of these studies suggests that simple cross-sectional (OLS) estimates of the return to education are biased downward relative to more sophisticated estimates that attempt to control for the endogeneity of schooling. Although a similar conclusion was reached by Griliches in his 1977 survey paper, many social scientists believe the opposite: that the cross-sectional correlation between education and earnings overstates the true effect of education.

I outline a simple theoretical model of optimal schooling, based on the framework of Becker (1967), that fully characterizes the cross-sectional relation between education and earnings. The model identifies two sources of heterogeneity in the population: variation in the marginal return to education at each level of schooling (loosely, differences in "ability"); and variation in the marginal rate of substitution between higher earnings in the future and more schooling (loosely, differences in "tastes" or "access to funds"). Individuals invest in

schooling until the marginal return to schooling is equated to their marginal discount rate. Except under very restrictive assumptions, equilibrium implies a non-degenerate distribution of marginal returns to education across the population. This distribution introduces some ambiguity into the interpretation of "the" causal effect of education: in essence, each person has his or her own causal effect.

The model generates a relatively simple expression for the OLS estimate of the return to schooling that depends on the means, variances, and covariances of the ability and discount rate variables across the population. According to the model, the OLS estimator gives an upward-biased estimate of the average marginal return to education. The size of the bias is proportional to the relative fraction of the variation in schooling outcomes that is attributable to differences in ability rather than differences in tastes or access to funds.

The model also offers an interpretation of estimates of the return to schooling based on instrumental variables like compulsory schooling laws or differences in the cost of college. If most of the variation in education outcomes is attributable to differences in discount rates then individuals with low levels of schooling will tend to have higher marginal returns to education than the population as a whole. Instrumental variables procedures that rely on changes in schooling for low-education and/or poor family background groups will therefore tend to identify relatively high marginal returns to schooling.

The model is also useful for interpreting estimates of the return to schooling based on differences in schooling over time (for the same person) or between siblings. In the latter case, a key question is whether inter-family differencing results in a greater relative reduction in the variation of ability or the variation of discount rates.

Two further issues that arise in estimation of the return to education are measurement error in schooling and unobserved heterogeneity in the level of earnings. Any comparison of different estimators of the return to schooling is complicated by the differential effects of measurement error. Some fraction of the apparent downward bias of OLS estimates of the return to schooling in the recent literature is surely due to measurement error. Whether measurement error is the whole story is an important question for further research. The existing studies provide such imprecise estimates that it is difficult to reject a pure measurement error explanation.

In my opinion, further research on the role of schooling in the labor market could usefully benefit from a more explicit consideration of the issues raised by a well-posed theoretical model. Among these issues: What are the underlying sources of variation in observed school choices? What variation is used by alternative econometric estimators of the return to schooling? Is the labor force reasonably well-described by a constant return to education for all workers? Can individuals anticipate their own returns to education? What combination of forces lead to the remarkable linearity in the observed cross-sectional relation between education and earnings?

References

- Angrist, Joshua D. and Whitney K. Newey. "Over-Identification Tests in Earnings Functions with Fixed Effects." Journal of Business and Economic Statistics 9 (July 1991): 317-323.
- Angrist, Joshua D. and Guido W. Imbens. "Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity." Unpublished Discussion Paper, Harvard University Department of Economics, 1993.
- Angrist, Joshua D. and Alan B. Krueger. "Does Compulsory Schooling Affect Schooling and Earnings?" Quarterly Journal of Economics 106 (November 1991): 979-1014. (1991a)
- Angrist, Joshua D. and Alan B. Krueger. "Estimating the Payoff to Schooling Using the Vietnam-Era Draft Lottery." Princeton University Industrial Relations Section Working Paper #290, August 1991. (1991b)
- Angrist, Joshua D. and Alan B. Krueger. "Split Sample Instrumental Variables". Unpublished Discussion Paper, Princeton University Department of Economics, July 1993.
- Ashenfelter, Orley and Alan B. Krueger. "Estimates of the Economic Return to Schooling for a New Sample of Twins." Princeton University Industrial Relations Section Working Paper #304, July 1992, revised April 1994.
- Ashenfelter, Orley and David Zimmerman. "Estimates of the Return to Schooling from Sibling Data: Fathers, Sons, and Brothers." National Bureau of Economic Research Working Paper 4491, October 1993.

- Becker, Gary S. Human Capital and the Personal Distribution of Income. Ann Arbor: University of Michigan Press, 1967.
- Bound, John and Alan B. Krueger. "The Extent of Measurement Error in Longitudinal Earnings Data: Do Two Wrongs Make a Right?". Journal of Labor Economics 9 (January 1991): 1-24.
- Bowen, William G. and Neil Rudenstein. In Pursuit of the Ph.D.. Princeton NJ: Princeton University Press, 1992.
- Butcher, Kristin F. and Anne Case. "The Effect of Sibling Composition on Women's Education and Earnings." Unpublished Discussion Paper, Princeton University Department of Economics, June 1993.
- Card, David. "Using Geographic Variation in College Proximity to Estimate the Return to Schooling." National Bureau of Economic Research Working Paper 4483, October 1993.
- Card, David and Alan B. Krueger. "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States." Journal of Political Economy 100 (February 1992): 1-40.
- Ehrenberg, Ronald and Robert Smith. Modern Labor Economics (4th edition). New York: Harper Collins, 1991.
- Griliches, Zvi. "Estimating the Returns to Schooling: Some Econometric Problems." Econometrica 45 (January 1977): 1-22.

- Hause, John C. "Earnings Profile: Ability and Schooling." Journal of Political Economy 80 Part II (May/June 1972): S108-S138.
- Hausman, Jerry A. and William E. Taylor. "Panel Data and Unobservable Individual Effects." Econometrica 49 (November 1981): 1377-1398.
- Heckman James J. and Solomon Polachek. "Empirical Evidence on the Functional Form of the Earnings-Schooling Relationship." Journal of the American Statistical Association 69 (June 1974): 350-354.
- Hungerford, Thomas and Gary Solon. "Sheepskin Effects in the Return to Education." Review of Economics and Statistics 69 (February 1987): 175-177.
- Kane, Thomas J. and Cecilia E. Rouse. "Labor Market Returns to Two- and Four-Year Colleges: Is a Credit a Credit and Do Degrees Matter?" Princeton University Industrial Relations Section Working Paper #311, January 1993.
- Lalonde, Robert J. "The Earnings Impact of U.S. Employment and Training Programs." Unpublished Discussion Paper, University of Chicago, 1993.
- Lang, Kevin. "Ability Bias, Discount Rate Bias, and the Return to Education." Unpublished Discussion Paper, Boston University Department of Economics, May 1993.
- Park, Jin Heum. "'Returns to Schooling: A Peculiar Deviation from Linearity.'" Unpublished Discussion Paper, Princeton University Department of Economics, 1994.
- Psacharopoulos, George. "Returns to Education: A Further International Update and Implications." Journal of Human Resources 20 (Fall 1985): 583-604.

Rosen, Sherwin. "Human Capital: A Survey of Empirical Research." In Ronald Ehrenberg, editor, Research in Labor Economics volume 1. Greenwich Connecticut: JAI Press, 1977.

Siegel, Paul and Robert Hodge. "A Causal Approach to the Study of Measurement Error." In Hubert Blalock and Ann Blalock, editors, Methodology in Social Research. New York: McGraw Hill, 1968.

Willis, Robert J. "Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Functions." In Orley Ashenfelter and Richard Layard, editors, Handbook of Labour Economics. New York NY: North Holland, 1986.

Figure 1
Optimal Schooling Choice

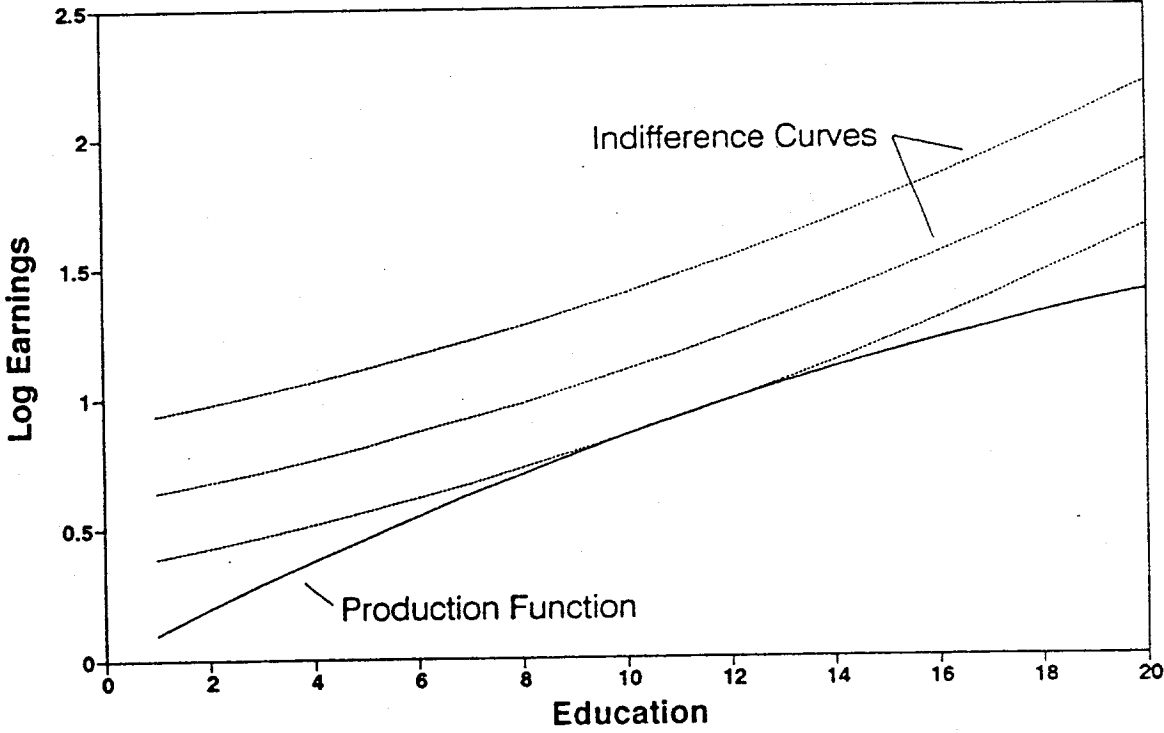


Figure 2
Marginal Efficiency and Discount Rate Schedules

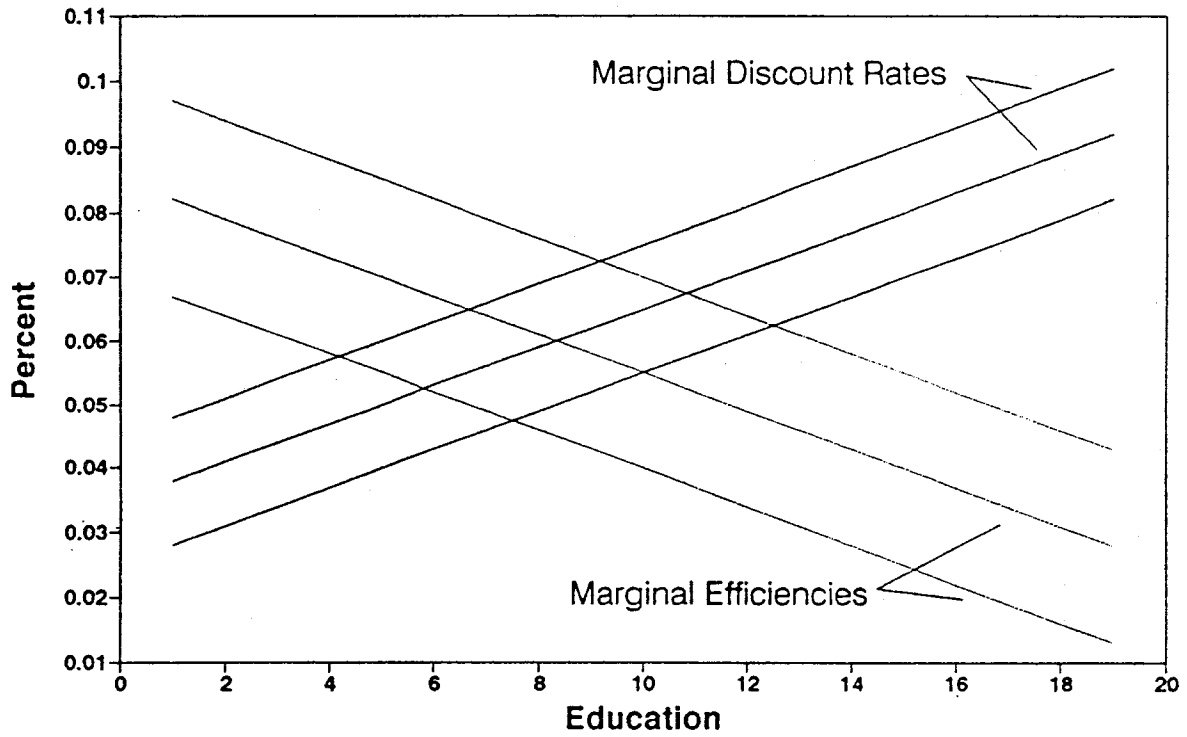


Figure 3
 Relation Between Wages and Education
 At Various Percentiles of Wages

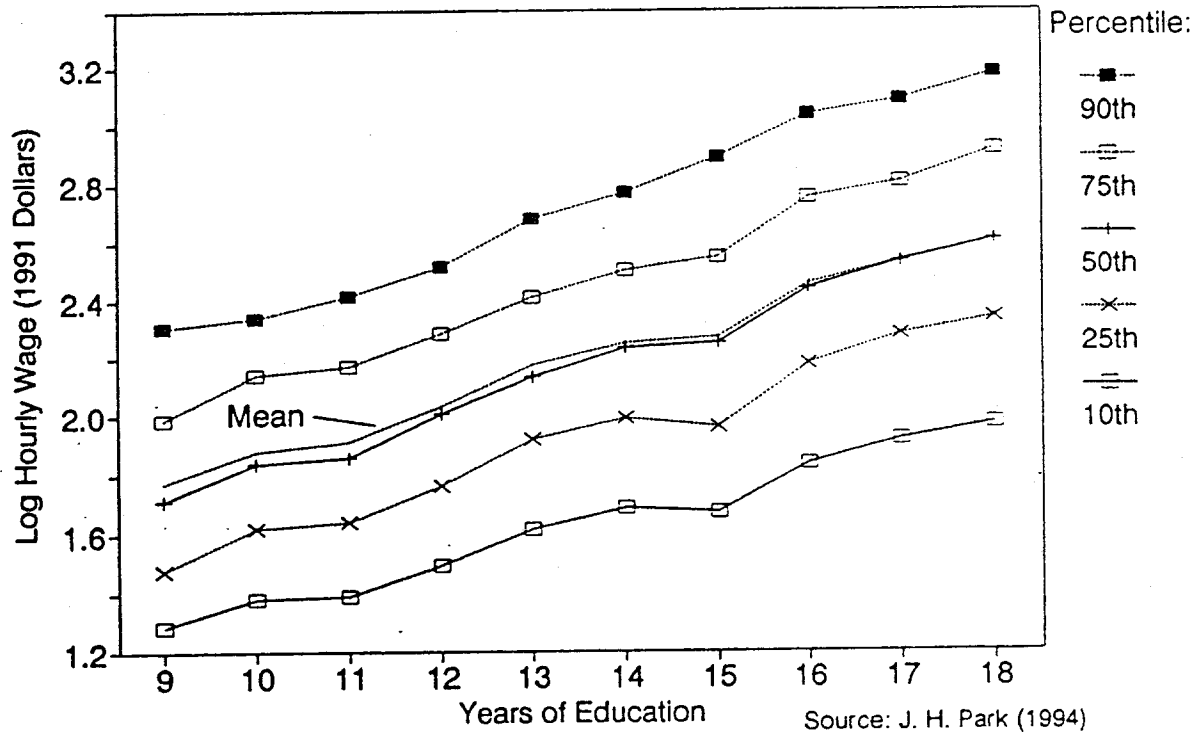


Figure 4
Cross-Sectional Earnings-Schooling Data

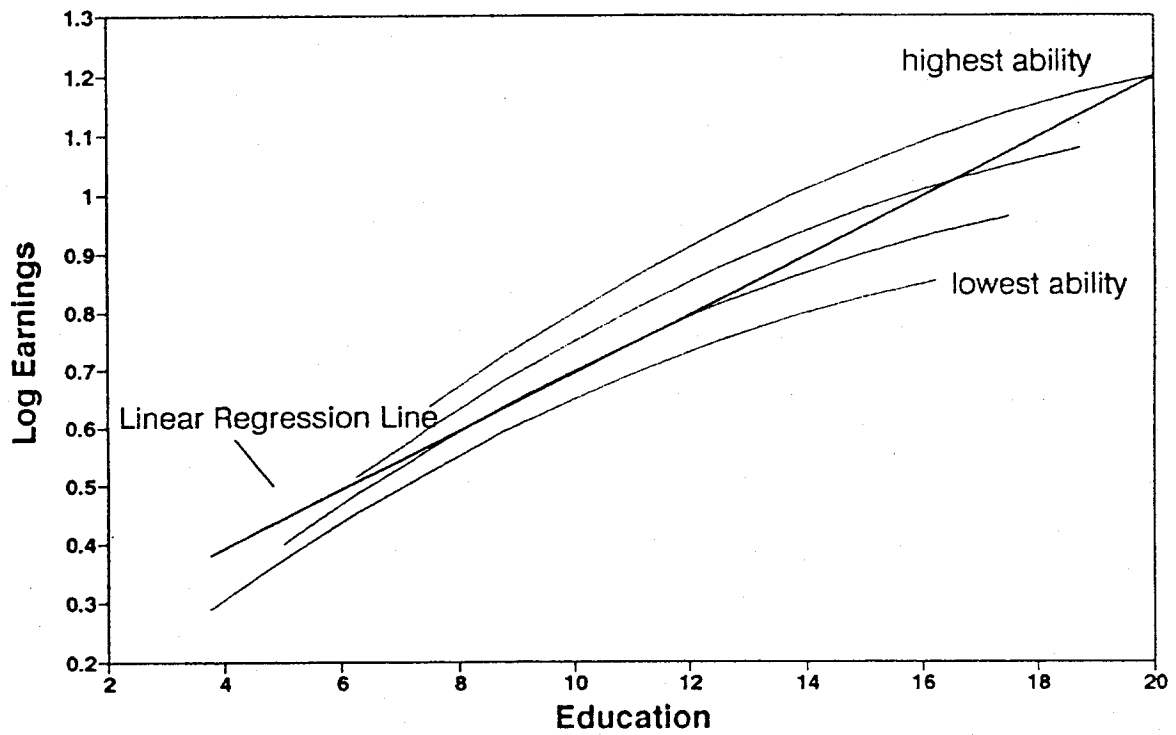


Table 1: Summary of Recent Estimates of the Return to Education

Study	Sample	Selected OLS and "Corrected" Education Coefficients	Comments
<u>I. Instrumental Variables Approaches</u>			
1. Angrist and Krueger (1991a)	1970 and 1980 Census. Men born 1920-20, 1930-39, 1940-49	Born 1930-39 in 1980 OLS: 0.063 (0.000) IV: 0.081 (0.011)	Instruments: year * quarter of birth; state * quarter of birth. Other covariates: race, central city, married, age, age-squared, state of residence, state of birth
2. Angrist and Krueger (1991b)	1979-85 CPS Men born 1944-53 (potentially drafted in Vietnam War)	OLS: 0.059 (0.001) IV: 0.066 (0.015)	Instruments: lottery number * year of birth. Other covariates: race, central city, married, year of birth region, veteran status (treated as endogenous)
3. Butcher and Case (1993)	1985 PSID White women Age 24+	OLS: 0.091 (0.008) IV: 0.185 (0.113)	Instrument: presence of any sisters. Other covariates: number of siblings, age, age squared, mother's & father's education, indicators for Catholic, oldest child, and family in poverty as child
4. Kane and Rouse (1993)	NLS Class of 1972 Men and Women, Valid Test Scores Working in 1986 Note: schooling measured by number of college credits divided by 30.	exclude test score, mother's & father's education OLS: 0.080 (0.005) IV: 0.091 (0.033) include test score, mother's & father's education OLS: 0.063 (0.005) IV: 0.094 (0.042)	Instruments: tuition at 2-yr and 4-yr colleges in state, distance to nearest colleges Other covariates: race, sex experience, number of years with child < 6, part-time status
5. Card (1993)	NLS Young Men Age 14-24 in 1966, working in 1976	OLS: 0.073 (0.004) IV: 0.132 (0.049) Use nearby college *parental education as instruments and include nearby college in wage equation IV: 0.097 (0.048)	Instruments: nearby college in county of residence 1966. Other covariates: region in 1966, SMSA in 1966, race, in South in 1976, family structure at age 14, father and mother education (main effects and interactions), quadratic in experience (treated as endogenous).

Note: estimated standard errors in parentheses. Table continues.

Table 1, continued

Study	Sample	Selected OLS and "Corrected" Education Coefficients	Comments
<u>II. Fixed Effects Approaches</u>			
6. Angrist and Newey (1991)	NLSY Men Age 18-28 in 1983, employed continuously 1983-87	OLS: 0.036 (0.003) Fixed Effect: 0.080 (0.015)	Other covariates: age, age squared
7. Ashenfelter and Krueger (1994)	1991 Twinsburg Survey data on identical twins	Cross-sectional OLS: 0.084 (0.014) Within twin pair OLS: 0.092 (0.024) Within twin pair IV: 0.167 (0.043) Within twin pair IV: 0.132 correlated (0.028) measurement error	Instruments: sibling's report of education Other covariates: age, age squared, race
8. Ashenfelter and Zimmerman (1993)	NLS Original Cohorts matched brothers sample, 1981 wages	OLS: 0.065 (0.011) Within family OLS: 0.047 (0.018) Within family OLS: 0.080 adjusted for (0.031) measurement error	Other covariates: age, age squared. Measurement error corrected within- family model assumes that 20% of variance of reported education is due to measure- ment error.

Note: estimated standard errors in parentheses.