



Munich Personal RePEc Archive

Ability Bias, Discount Rate Bias and the Return to Education

Lang, Kevin

Boston University

12 March 1993

Online at <https://mpra.ub.uni-muenchen.de/24651/>
MPRA Paper No. 24651, posted 29 Aug 2010 09:17 UTC

ABILITY BIAS, DISCOUNT RATE BIAS AND THE RETURN TO EDUCATION

Kevin Lang
Department of Economics
Boston University
270 Bay State Road
Boston, MA 02215

March 12, 1993

This study was supported in part by a Sloan Faculty Research Fellowship and NSF grant SES-9223349. I am grateful to Joshua Angrist, Paul Beaudry, David Card, Russell Cooper, Rena Eichler, Robert Gibbons, Shulamit Kahn, Lawrence Katz, Alan Krueger, Fabio Schiantarelli, Andrew Weiss and participants at workshops at Berkeley, Boston College, Laval, the NBER, the University of Pennsylvania, MIT and UCSD for helpful comments and discussions. Rena Eichler provided excellent research assistance. The usual caveat applies.

ABSTRACT

I reconsider various methods for correcting for bias in estimates of the returns to schooling. I argue that the literature on ability bias has ignored complications implicit in theoretical formulations of the choice of human capital. In particular, such models imply that adding ability to the wage equation may not be informative about the importance of bias and that variables correlated with the discount rate will generally not be suitable instruments for education. Indeed discount rate variation may generate downward bias. Estimation of a structural wage/schooling model suggests that OLS estimates of the return to schooling are biased downwards.

The return to education plays a key role in much of labor economics. Estimates of the return are central to discussions of the usefulness of education for development policy, for fighting poverty and for limiting race-related wage differentials. However, as economists began to estimate the impact of schooling on wages, they recognized that ability and schooling were likely to be correlated and that measures of ability were generally imperfect, if present at all in data sets. If ability affects earnings, the estimated return to education will therefore be biased upwards.

A considerable body of work has arisen to assess the significance of this ability bias. A fair assessment is that this vast body of research has been inconclusive. Some researchers find evidence of significant ability bias; many find evidence of little ability bias while yet others find that the return to education is biased downwards. Evidence of downwards bias has been particularly prevalent in recent clever instrumental variables estimates, leading many labor economists to conclude that ability bias is not a problem.

In this paper I argue that there is a second source of bias in OLS estimates of the return to education which economists have ignored. I call this "discount rate bias." It arises because individuals with different discount rates will have different equilibrium marginal returns to education. The wage equation may therefore be best understood using a random coefficients model in which there is correlation between schooling and the random part of the coefficient. While ability bias raises the estimated return to education, discount rate bias lowers it so that the net bias is indeterminate.

The paper is constructed as follows. Section I begins with a discussion of estimation bias. I argue that standard approaches to the estimation of the return to schooling do not provide consistent estimates. As part of the discussion I introduce the concept of discount rate bias. In section II, I develop and estimate a structural model of wages and schooling. The results

suggest that both ability and discount rate bias are important, but that the latter dominates so that OLS estimates of the return to schooling are biased substantially downwards.

I. Ability Bias: Theory

This section reviews the standard approaches to correcting the return to education for ability bias and argues that they are inconsistent. Moreover, both theorists and econometricians have ignored the possibility that variation in discount rates may also be an important source of bias. Since, as I show, discount rate variation biases the estimated return downwards, the two types of bias tend to offset each other. However, I argue below that their combined effect is not strictly additive.

I begin with a model in which the only source of variation in attained schooling is ability differences. I show that in such models, adding measures of ability to the wage equation need not be informative about the extent of ability bias.

Econometricians have generally started their investigation of ability bias by specifying a standard log wage equation of the form

$$(1) \quad \log w = XB + bs + ci + \epsilon$$

where s is schooling, i is innate ability and X is a vector of additional explanatory variables which for simplicity will be ignored in the following discussion. This equation can be interpreted as a specific functional form for the human capital production function. Ability bias arises in equation (1) if S and i are correlated, the econometrician does not measure i , and c is non-zero.

Although econometricians are inclined to favor linear functional forms and there is reasonable evidence that log wages are approximately linear in

schooling, equation (1) is in some ways an odd starting point for the investigation of ability bias. In the standard model of education with the usual simplifying assumptions¹ individuals remain in school until $d \ln w/ds$ equals their discount rate. In this case, equation (1) implies that schooling should be independent of ability since the return to education ($d \ln w/ds$) does not depend on ability.² Ability bias would therefore arise only if ability were correlated with tastes for education or the discount rate.

Early human capital theorists anticipated that bias would arise directly from the return to education being higher for more able workers (Becker 1975). While equation (1) cannot yield ability bias that arises in this way, it is easy to generate production functions which can. I provide an example below, but begin with a more general discussion.

Let us assume that

¹These assumptions are that workers maximize lifetime wealth, time spent in the labor force is independent of time spent in schooling (or equivalently lifetimes are infinite), there are no direct costs of education and the effect of experience on earnings is multiplicative. The assumption that tuition is free is probably most objectionable. Relaxing this assumption for college would mean that the return to schooling would exceed the discount rate for individuals getting exactly a high school degree and would be even higher for those getting some college education.

²Indeed with perfect capital markets, the only possible equilibrium is one in which b equals the interest rate or all workers get zero education. This point has certainly not been ignored in the literature. If one drops one or more of the assumptions in footnote 2, it is possible to treat (1) as the human capital production function. Thus Griliches (1977) discusses how different cost functions for schooling may generate a positive or negative correlation between ability and schooling. Lang and Ruud (1986) assume that the wage equation takes the form given by (1) but that the time it takes to complete a marginal year of completed school is increasing in schooling and decreasing in ability. This ensures that the second-order conditions are satisfied and that ability and schooling are positively correlated. On the other hand, Hause (1972) drops the assumption that the human capital production function is given by (1) and estimates both wage and log wage equations and allows ability to interact with schooling. See also the related work by Willis and Rosen (1979).

$$(2) \quad \log(w) = q(s, i) + e$$

where e is a stochastic term assumed orthogonal to i .³ I will refer to q as the human capital production function.

Schooling is determined by the first order condition $\partial q / \partial s = r$. If $\partial^2 q / \partial i \partial s$ is positive, in equilibrium, schooling and ability will be positively related so that we can write the demand for schooling as $s = s(i)$. The equation for schooling can be inverted to give innate ability as a function of schooling -- $i = i(s)$. The assumption that the cross-derivative is positive will be maintained throughout the remainder of this paper.

There are several important points to note about equation (2). First, there is no obvious reason based on (2) that s and e should be correlated. If the human capital production function is specified correctly and i can be measured, it may be estimated consistently by least squares.

Second, given the functional relation between schooling and ability, the human capital production function can be expressed as a function of s alone or as a function of i alone.⁴ We write the simplified wage equation with i removed as

$$(3) \quad \ln w = q(s, i(s)) + e = q^*(s) + e.$$

Equation (3) tells us that there is a functional form for the wage such that education provides all the information about its nonstochastic component.

³It may be helpful to think of i as ability which is productive in both school and work and e as ability which is only productive in work, but this interpretation is not necessary and will not be used in this paper.

⁴This is a standard point about the estimation of hedonic wage equations. The wage equation can be written as a function of worker characteristics alone, firm characteristics or the characteristics of the worker/job match. For a discussion of hedonics see Rosen (1974, 1986).

Including measures of i will not add to the equation's explanatory power and therefore will not change the estimated return to schooling. Nevertheless, an estimate of the effect of schooling on the wage which is based on this equation overestimates the effect of schooling on wages because it includes the effect of ability on wages and the relation between ability and schooling. In other words, the estimates will suffer from ability bias, but adding measures of ability to the equation has no effect on the estimated return.

If the functional form chosen for the wage equation is neither q nor q^* , adding ability to the equation can give perverse results. It may actually increase the estimated return to schooling. When the human capital production function is misspecified we can think of the error term as

$$(4) \quad \text{error} = q^*(s) - q'(s) + e$$

where q' is the misspecified production function. To see how including i in the equation will affect the estimated return, we can think of regressing the error term on q^* (or, more precisely, the derivatives of q^* with respect to its parameters) and on measures of i . There is no strong reason to expect the coefficient on i to be positive.

The following examples will illustrate this point.

Example 1: Suppose that

$$(5) \quad q = is - bs^2.$$

Then applying the first order condition gives the following equilibrium relation between ability and schooling:

$$(6) \quad i = r + 2bs.$$

Note that given the human capital production function (5), equilibrium implies that ability and schooling are linked by a linear relation. We can write the human capital production function as a function of s alone or of i alone or, equivalently, the wage equations as

$$(7a) \quad \ln w = rs + bs^2 + e$$

$$(7b) \quad = (i^2 - r^2)/4b + e.$$

Consider now what happens if we first estimate a wage equation with only schooling as an explanatory variable and then estimate the equation with measures of ability as well. If we have specified the log wage as quadratic in s , we can see from (7a) that the equation explains all the nonstochastic variation in wages. If we add measures of ability to (7a), they will not enter the equation significantly. Thus as argued more generally above, adding measures of ability to a properly specified q^* equation has no effect on the estimates of the return to schooling. Nevertheless, for all levels of schooling above zero, the estimated return to schooling ($r+2bs$) exceeds the true return to schooling (r).

What happens if we choose some other functional form for the wage equation? Suppose that instead we estimate a standard log wage equation with only a linear education term and intercept and try adding measures of ability. From equation (6), schooling and ability are perfectly correlated. Therefore it would not be possible to include both in an equation. Suppose, however, that our measure of ability was actually i^2 . Then from (7b), the best fit is

achieved using only a constant term and the ability measure. The coefficient on education will fall to zero. We might mistakenly conclude that education was completely unproductive.

On the other hand, consider the case where our measure of ability is $\log(i)$. In this case the coefficient on ability may well be negative since S^2 regressed on S and $\log(S)$ will give a negative coefficient and i is perfectly correlated with S . In this case, we might be tempted to "explain" the negative coefficient on ability by unmeasured work conditions which are correlated with ability.

While the first example demonstrates the general point made above, the fact that the wage equation given by (7a) is quadratic in schooling might lead us to doubt its empirical relevance. Many studies (although by no means all) find that the log wage equation is, in fact, well approximated by the use of a single linear schooling term and that no higher order terms are required.

The following example should put this concern to rest. I show that a family of production functions, which includes Cobb-Douglas as a special case, generates a log wage equation which is linear in schooling.

Example 2: Assume that the log wage is given by

$$(8) \quad \log(w) = f(i)s^b + XB + e \quad 0 < b < 1$$

where b is either a parameter or a second ability measure, X is a vector of additional explanatory variables.

Applying the first order condition to (8) gives the relation between innate ability and education

$$(9) \quad bf(i)s^{b-1} = r.$$

Here the source of variation in schooling is differences in the level of innate ability, i . Substituting (9) into (8) gives a standard log wage equation of the form

$$(10) \quad \log w = (r/b)S + XB + e.$$

Recall that individuals remain in school until the return to education equals the discount rate so that the true return to education is r . However, the OLS estimate of the return to education would be r/b . Thus estimation of equation (10) is subject to ability bias. The magnitude of this ability bias depends critically on the value of b . As b tends to 1, ability bias tends to zero. On the other hand, as b tends to zero, ability bias tends to infinity.

Note that this example generates the standard semi-log wage equation as a best-fitting functional form for wages. Note also that S and X explain all of the variation in the log wage except for the random component, e , which is orthogonal to all other variables. Adding innate ability, i , to the equation will have no effect whatsoever on the estimated coefficient. Moreover, the coefficient on ability will be zero.

In effect, the situation resembles that of multicollinearity. Each of the two explanatory variables, i and s , can be expressed as a perfect function of the other. In contrast with multicollinearity, the relation is nonlinear rather than linear. As a consequence, one variable can "pick up" much of the variation attributable to the other. In contrast with the case of multicollinearity, because there is not a perfect linear relation between the variables, the problem is not signalled by high standard errors.

So far the argument has proceeded under the assumption that ability is the sole source of variation in educational attainment. This assumption is

obviously unrealistic. Educational attainment may differ among equally able individuals because of differences in discount rates (or tastes which have much the same effect). Moreover, ability is not unidimensional. Consequently, the problem generated by a close functional relation does not arise in practice. However, provided ability is an important determinant of educational outcomes, the general argument will hold. In fact, it is trivial to generate examples with two dimensions of ability and discount rate variation in which adding ability to a standard log wage equation gives perverse results.⁵ Moreover, as will be discussed shortly, adding discount rate variation generates additional problems.

The concerns raised here about the usefulness of including measures of ability as controls for ability bias are justified by the variety of results in the empirical literature. Perverse signs for the ability coefficients are not unusual (e.g. Conlisk, 1971; Griliches, 1977; Lang and Ruud, 1986; Krueger, 1993); reductions in the return to education on the order of 20% are common (Chamberlain, 1978; Griliches, Hall, and Hausman, 1978) and large reductions occur in some data sets (Boissiere, Knight and Sabot, 1985) and when ability is treated as measured with error (Chamberlain, 1978).

⁵Assume that the log of output is given by is^b but that in contrast with the previous discussion both i and b vary so that there are two dimensions of ability. Further assume that r varies so that variation in education results in part from differences in discount rates which are independent of ability. More precisely r takes on the values (.0425, .0430, ..., .0475), i the values (.375, .380, ..., .425) and b the values (.430, .435, ..., .470) and each combination of i , b , and r values occurs exactly once (1089 observations). The true mean effect of education on productivity is .045, the mean of the discount rate. Regressing q on s gives an estimated return to schooling of .074. Adding i to the equation reduces the estimate only slightly to .067. Including b , which is also a measure of ability, raises the estimate to .075, and the sign of the coefficient on b is perverse. It is worth noting that there is nothing inherently bizarre in this example. The ability variable, b , is positively correlated with education and explains a substantial fraction but by no means all of the variation in education ($R^2 = .57$).

II. Discount Rate Bias: Theory

If adding measures of ability to the wage equation is not informative, a natural reaction is to attempt to develop instrumental variables estimators. In order to be able to use an IV estimator, we, of course, require some variable which affects educational attainment but which is not correlated with ability. The natural candidate is the discount rate. It is admittedly difficult to think of variables which are correlated with the discount rate but not with ability. What is more important, and perhaps surprising, is that even if we were able to isolate the discount rate to use as a first-stage regressor and if the discount rate were uncorrelated with ability, the resulting estimator would still generally be inconsistent.

To see this, note that if the discount rate varies, the schooling equation becomes $s = s(i,r)$. Again we can invert the schooling equation to get innate ability as a function of schooling and the discount rate — $i = i(s,r)$. Even if innate ability and the discount rate are uncorrelated, they are correlated once we condition on the level of schooling. For a given level of schooling, individuals with higher discount rates will have more innate ability.

When both ability and the discount rate vary, the human capital production function is written as

$$(11) \quad q(s, i(s,r)) = q^*(s,r).$$

Note that because the level of ability associated with a given level of schooling depends on the discount rate, when we eliminate ability from the reduced form, the discount rate remains in the wage equation.

The standard wage equation can be thought of as a linear approximation to the q^* equation. Linearizing q^* gives

$$(12) \quad q = q_0^* + q_s^*(s-s_0) + q_r^*(r-r_0) + \text{remainder term.}$$

If we regress the log wage on schooling alone, the error term contains the discount rate as one of its components. Since higher discount rates lower attained schooling, this component is correlated with education. Workers with high discount rates will have low levels of education.

Again this can be seen easily in the examples especially in equation (10) where (12) applies exactly. If r varies, then equation (10) can be thought of as a random coefficients model:

$$(13) \quad \log w = (\bar{r}/b)S + e + \nu S.$$

High values of ν (r) are associated with low values of S so that S is negatively correlated with the error term. Consequently, variation in the discount rate introduces a negative bias which may partially or completely offset the positive "ability" bias. Similarly, r is positively correlated with the error term so that an IV estimator based on r is inconsistent.

Figure 1 demonstrates this situation for a case where all individuals have one of two interest rates. The dashed lines represent the regression lines for each of the two sets of individuals. Note that the right dashed line involves individuals with lower discount rates and therefore generally higher levels of education. The solid line is the regression line. It is flatter than the either of the dashed lines and can easily be flatter than the mean discount rate.

These problems are reflected in experience with IV estimators. Some estimators produce estimated returns which exceed those obtained using OLS (Hausman and Taylor, 1981; Lang and Ruud, 1986) while others produce negative returns (Chamberlain, 1978; Ruud, 1981). A fair assessment of the literature

on ability bias is that "corrected" returns are centered on values only slightly below the OLS estimates, but there is considerable dispersion. Many researchers have drawn the reasonable conclusion that there is little evidence of significant ability bias.⁶

It is important to note that the derivation of both equations (12) and (13) assumes that innate ability varies systematically with education, holding r fixed. Thus the demonstration of the existence of this second sort of bias presupposes the existence of ability bias. It is also important to note that discount rate bias is different from standard specification bias. If we mistakenly used the linear form of equation (13) when the true equation linking wages to schooling (with ability eliminated from the equation) was quadratic, the coefficient on S might be greater than or less than r . However, this form of specification bias is not unique to the wage equation, and I do not consider it further.

A similar problem to discount rate bias arises in Angrist and Krueger (1991a) who use quarter of birth to instrument for education. Since quarter of birth works through the effect of compulsory school attendance laws, only the workers with the lowest level of education will be affected. These will tend to be workers with relatively high discount rates. The Angrist/Krueger estimator is consequently biased upwards.⁷

⁶The random coefficients specification in (13) also helps to explain why including b in the equation can have a perverse effect as described in footnote 5. If both b and r vary, the error term in the equation includes the deviation of r/b from its mean multiplied by S . It is not surprising that a regression of rS/b on b and S can generate a negative coefficient on b and therefore that the sign on b when added to the equation can be perverse.

⁷Angrist and Krueger are aware of this problem. One interpretation of their paper is that they are obtaining a consistent estimate of the return to schooling for individuals with low levels of education. They attempt to address the issue of the return to education at other levels of education by using draft lottery numbers as an instrument for education (Angrist and Krueger, 1991b) in the expectation that a high risk of being drafted would encourage young men to make use of the college deferment. Their findings are similar to those in their earlier paper. Unfortunately, lottery number is only weakly related to education leading to problems of interpretation.

Recognizing the existence of discount rate bias also casts doubt on the usefulness of studies of the return to education which rely on twins. The usual argument is that twins (especially identical twins) are likely to be similar with respect to innate ability. Thus, the argument continues, comparing twins removes ability bias and provides an unbiased estimate of the return to education.

However, although we may diminish ability as a source of variation in education, discount rate differences may remain as the source of variation. As just discussed, discount rate bias lowers the estimate of the return to education. Of course, it is not obvious whether comparing twins has a bigger effect on discount rate or ability variation. Compared with randomly chosen individuals from the general population, we would expect twins to be more similar with respect to both ability and discount rates. Since the two types of bias interact in a complex fashion, there is no strong reason to expect studies of twins to produce higher or lower estimates of the return to education than do standard cross-sectional studies.⁸

While thinking about bias in terms of the interaction of discount rate and ability bias appears to be a useful approach, because of the random coefficients structure to the model, there is no easy way of thinking about how variation in discount rates and ability interact. In the next section, I

⁸Ashenfelter and Krueger (1992) find significant downward bias. Early studies found significant upward bias (see Behrman et al, 1980; Taubman, 1977) but some critics have attributed this to greater measurement error in such studies (Chamberlain and Griliches, 1977). See also Taubman (1976) and Griliches (1979). The discussion of the relation between family and different types of bias in this paper is reminiscent of the discussion of the effect of family on schooling and ability in the latter paper. Griliches also argues that there is little reason for preferring studies of twins to studies of siblings. Studies of siblings have not produced a consistent set of results.

therefore estimate a structural model which allows us to look at these questions more directly.

III. A Structural Model of the Human Capital Production Function and Schooling

The argument in the previous section implies that it may be impossible to correct for ability bias by adding measures of ability to a standard semi-log wage equation. However, if we were able to specify the human capital production function correctly, neither the problem of a close functional relation between ability and schooling nor the discount rate bias which occurs when the production function is misspecified would arise. Consequently, we could get unbiased estimates of the return to education. I therefore experiment with estimating (8) and (9) directly to determine whether this suggests substantially different estimates of the return to schooling compared with standard OLS estimates.

I begin by specifying a human capital production function similar to the one used in the earlier discussion so that

$$(14) \quad \ln w = (Z\Gamma + ci)S^b + XB + e.$$

This human capital production function gives rise to the following education equation:

$$(15) \quad \ln S = (\ln r^* - \ln b - \ln(Z\Gamma + ci))/(b-1) + v.$$

where r^* is the geometric mean discount rate and v is proportional to the deviation of the log discount rate from this mean.

Assuming that discount rates and productivity are not correlated, equations (14) and (15) can be estimated as a pair of nonlinear seemingly unrelated equations with cross-equation constraints.

It is important to be clear about the objective of this estimation. I do not expect to provide compelling estimates of the "true" return to education and the underlying demand for education equation. That would require considerably more confidence in the appropriateness of the functional form chosen in (14) than can be justified for any functional form chosen a priori. Indeed, it is fairly obvious that the model cannot account for the concentration of schooling at high school and university graduation without some improbable distributions of ability and discount rates. Any model which did account for these concentrations would require even stronger functional form and distributional assumptions.

Instead, what I hope to demonstrate is that estimates of the return to education are sensitive to the assumption implicit in most empirical work that b equals 1. Moreover, the estimates suggest that there is considerable variance in both ability and discount rates so that both types of bias discussed in the previous section are likely to be important. In addition, (14) is attractive because it produces the standard log wage equation once $ZI+ci$ has been eliminated from the equation and satisfies the second order conditions for wealth maximization.

In assessing the usefulness of this structural approach, it is important to remember that the standard semi-log specification is not a structural but rather assumes a different, and in my opinion, even less plausible structure. I will provide evidence that the structural model provides more reasonable estimates on some key points.

The data for this section are drawn from the National Longitudinal Survey of Youth (NLSY). The sample was restricted to individuals in the nonmilitary sample, who were employed in 1988 and for whom data were available on the Armed Services Vocational Aptitude Battery and from the 1979 school survey.

Individuals with missing data on any of the other key variables were also dropped. The final sample contains 4253 observations. Means for the variables used in the analysis are included in Table 1.

Most of the variables are self-explanatory. The school quality variables are drawn from the supplementary survey of the last secondary school the individual attended. The ability variable is based on the results of the Armed Forces Qualifying Test (AFQT) administered as part of the NLSY in 1980. Although the AFQT is a respected aptitude test, like all such tests it is imperfect. One potentially serious problem is that the AFQT may measure what has been learned in school. In this case, the AFQT score might capture some of the productivity effect of education.

Fortunately, because a large fraction of respondents had not completed their schooling in 1980, it is possible to get around this problem. I regress the AFQT score on completed schooling as of 1980 (measured in the 1981 survey since the AFQT was administered after the end of the 1979-80 school year) and the three measures of school quality. Since more able students will tend to get through school faster and be less likely to drop out, OLS estimation of this equation would suggest that schooling has too great an effect on the AFQT. This problem is resolved by using age as an instrument for education in the equation.

The results (shown in Table 2) are consistent with the view that the AFQT is endogenous to education and that learning does take place in school. Each year of education is associated with a 3.8 point increase in AFQT. Lower pupil/teacher ratios and higher fractions of teachers with MAs or PhDs are associated with higher AFQT scores. Higher teacher salaries are associated with lower AFQT scores although this effect falls just short of significance at the .1 level.

Table 2 must be interpreted with caution. No attempt has been made to control for the endogeneity of school quality measures. It is possible that these variables are capturing neighborhood or family effects. Similarly, it is possible that the effect ascribed to education is simply an age effect. Nevertheless, these results at least suggest that school quality measures may affect learning.

The ability measure used in this paper is the residual from the equation shown in Table 2. It is the part of the AFQT which cannot be accounted for by schooling attainment at the time the test was administered.

The first column of Table 3 gives the results of estimating equation (14) simultaneously with (15). The first part gives the parameters (Γ, c) for the terms which are interacted with schooling. This is followed by the education exponent (b) and by the linear terms⁹ (B) and by the value of $\ln r^*$. For purposes of comparison, the second column gives standard OLS estimates of equation (14) including the full set of schooling interaction terms but with the exponent on schooling assumed to be 1. In the first column, the estimated exponent (b) is significantly less than 1. The results suggest that the marginal product of education diminishes as the level of education increases.¹⁰

⁹It would be natural to include the AFQT residual in the linear terms. Unfortunately, considerable efforts to converge the model with the AFQT residual in both the interaction and linear terms were unsuccessful.

¹⁰It is important to remember that if the model given by (14) and (15) is correct, equation (14) is only identified when estimated along with (15). The wage equation, itself, can be estimated equally well with a linear model or with the form given by (14). As a result the nonlinear model and the OLS results are not nested. The nonlinear model has an additional parameter not found in the OLS equation but imposes additional cross-equation constraints. Nevertheless, I also estimated equation (14) by itself. Not surprisingly, the standard errors of the parameter estimates are quite high. The estimated exponent is .935 with a standard error of .2. The remaining coefficients generally lie between the coefficients in the first two columns of Table 3.

Strikingly, however, this structural model suggests that the return to education is higher than the estimate obtained by OLS estimation. The mean return to education using this structural model is 13% compared with only 8.2% for the OLS estimates. Moreover, there is a large difference in the range of returns suggested by the two sets of estimates. While the highest estimated return to education is between 14% and 15% in both models, this structural model suggests that the lowest return to education is 13% compared with negative estimated returns for some individuals in the OLS estimates. Thus this structural model is able to remove some of the implausible results associated with the OLS estimates.

Further evidence of the advantages of this structural model comes from its ability to explain educational attainment. Under almost any economic model, anything which increases the return to education for a fixed level of education (in other words, shifts the demand for educated labor to the right) should increase the amount of education individuals choose. Increases in the value of the interaction terms, $Z\Gamma+ci$ represent such a shift. If we regress educational attainment on $Z\Gamma+ci$ using the OLS estimates, the R^2 is only .04. In contrast, the R^2 for a regression of educational attainment on the same interaction terms using the structural model is .29. We should not exaggerate the explanatory power of the structural model since the coefficients for the interaction terms are chosen, in part, for their ability to explain educational attainment. Nevertheless, the difference is striking.

It is also noteworthy that the OLS model gives rather poor results regarding the effect of school quality on the return to education. The pupil/teacher ratio has an insignificant effect on the return to education while teachers' educational credentials and salaries appear to have perverse effects. In contrast, when we use the structural model, higher pupil/teacher

ratios reduce the return to schooling by a statistically significant amount, and higher educational credentials for teachers increase the return. Again the effect is significant at the .05 level using a one-tailed test. The sign on teachers' salaries remains perverse, but the effect falls far short of statistical significance.

It is worth examining the conclusions we might draw from other approaches to correcting for bias. Column 3 of Table 3 gives OLS estimates when the AFQT residual is excluded from the model. For this model, the mean return to education is .094 compared with .082 from column 2. A researcher examining the OLS results might reasonably conclude that ability bias probably induces a small upwards bias in the estimated return to education.

In sum, there is both ability bias and discount rate bias in the estimates of the return to schooling. In this data set, the latter seems to dominate so that the OLS estimate of the return to schooling is biased downwards. Standard approaches to correcting for bias exacerbate the problem.

IV. Conclusions

The results in this paper suggest that there is significant bias in standard OLS estimates of the return to education. While neither the magnitude nor the direction of this bias can be determined a priori, the results from the NLSY suggest that there is significant downwards bias. The direction of the bias is consistent with other recent empirical papers on estimating the return to schooling.

They demonstrate that we cannot cavalierly dismiss the significance of estimation bias. These results therefore suggest the need for considerable caution when assessing the impact of various policies on the return to education since we cannot assume that estimation bias is invariant to policy.

REFERENCES

Angrist, Joshua D. and Krueger, Alan B., "Does Compulsory School Attendance Affect Schooling and Earnings?" Quarterly Journal of Economics, 106 (November 1991a): 979-1014.

Angrist, Joshua D. and Krueger, Alan B., "Estimating the Payoff to Schooling Using the Vietnam-era Draft Lottery," Princeton, unpublished, 1991b.

Ashenfelter, Orley, and Krueger, Alan, "Estimates of the Economic Return to Schooling from a New Sample of Twins," Industrial Relations, Princeton University Working Paper No. 304, June 1992.

Becker, Gary S., Human Capital, 2nd ed., Chicago: Chicago University Press, 1975.

Behrman, Jere R., Hrubec, Zdenek, Taubman, Paul, and Wales, Terence, Socioeconomic Success: A Study of the Effects of Genetic Endowments, Family Environment, and Schooling, Amsterdam: North-Holland, 1980.

Boissiere, M., Knight, J.B. and Sabot, R.H., "Earnings, Schooling, Ability, and Cognitive Skills," American Economic Review, 75 (Dec. 1985): 1016-29.

Chamberlain, Gary, "Omitted Variable Bias in Panel Data: Estimating the Returns to Schooling," Annales de l'insee, 30-31 (April-Sept., 1978): 49-82.

Chamberlain, Gary and Griliches, Zvi, "More on Brothers," in Taubman, Paul, ed., Kinometrics: The Determinants of Success Within and Between Families, Amsterdam: North-Holland, 1977.

Conlisk, John, "A Bit of Evidence on the Income-Education-Ability Interrelation," Journal of Human Resources, 6 (Summer 1971): 358-62.

Griliches, Zvi, "Estimating the Returns to Schooling: Some Econometric Problems," Econometrica, January 1977, 45, 1-22.

Griliches, Zvi, Sibling Models and Data in Economics - The Beginnings of a Survey," Journal of Political Economy, October 1979, 87, S37-64.

Griliches, Zvi, Hall, Bronwyn H., and Hausman, Jerry A., "Missing Data and Self-Selection in Large Panels," Annales de l'insee, 30-31 (April-Sept., 1978): 137-76.

Griliches, Zvi, and Mason, William M., "Education, Income, and Ability," Journal of Political Economy, 80 (May/June 1972): S74-S103.

Hause, John, "Earnings Profile: Ability and Schooling," Journal of Political Economy, 80 (1972): S108-38.

Hausman, Jerry A., and Taylor, William E., "Panel Data and Unobservable Individual Effects," Econometrica, 49 (Nov. 1981): 1377-98.

Krueger, Alan, "How Computers Have Changed the Wage Structure: Evidence from Microdata, 1984-89," Quarterly Journal of Economics, 108 (February 1993): 33-60.

Lang, Kevin and Ruud, Paul, "Returns to Schooling, Implicit Discount Rates and Black-White Wage Differentials," The Review of Economics and Statistics 68 (February 1986): 41-47.

Rosen, Sherwin, "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," Journal of Political Economy, 82 (January/February 1974): 34-55.

Rosen, Sherwin, "The Theory of Equalizing Differences," in Orley Ashenfelter and Richard Layard, eds., Handbook of Labor Economics, Amsterdam: North-Holland, 1976, 641-92.

Ruud, Paul A., "Misspecification in Limited Dependent Variable Models," unpublished doctoral dissertation, Massachusetts Institute of Technology, 1981.

Taubman, Paul, "Determinants of Earnings - Genetics, Family and Other Environments - A Study of White Male Twins," American Economic Review, December 1976, 66, 858-70.

Taubman, Paul, ed., Kinometrics, Amsterdam: North-Holland, 1977.

Willis, Robert J. and Rosen, Sherwin, "Education and Self-Selection," Journal of Political Economy, 87 (October 1979): S7-S36.

TABLE 1

Summary Statistics for Key Variables

	MEAN	STD DEV
In US at 14	0.98	0.10
Mother's Education	11.27	2.99
Father's Education	11.33	3.83
Number of Siblings	3.56	2.44
Pupil/Teacher Ratio	19.31	5.78
%Teachers with MA/PhD	47.47	23.43
Starting teachers' salary	10.74	1.15
Male	0.51	0.49
Black	0.19	0.39
Hispanic	0.13	0.33
Education	13.23	2.26
Experience	8.25	2.90
Bad Health	0.03	0.19
Urban	0.78	0.41
Unemployment Rate	2.57	0.87
South	0.36	0.48
Never Married	0.37	0.48
AFQT	70.11	20.51
Log (hourly wage)	6.67	0.50

TABLE 2

Determinants of 1980 AFQT Scores

Constant	28.65 (3.81)
Education Completed on 1981 Survey	3.77 (0.25)
%Teachers with MA or PhD	0.04 (0.01)
Pupil/Teacher Ratio	-0.16 (0.04)
Starting Salary for Teacher with BA/1000	-0.35 (0.22)

Standard errors are in parentheses.

Instrumental variables estimation. Exogenous variables are the three school quality variables and year of birth.

TABLE 3

Parameter Estimates of Structural Wage/Schooling Model

	<u>Nonlinear</u>	<u>OLS</u>	<u>OLS</u>
(Schooling Interactions) [Γ, c in (14) and (15)]			
Constant	.213 (.030)	.125 (.039)	.129 (.040)
AFQT residual	.113E-03 (.483E-04)	.436E-03 (.374E-04)	
In US at 14	-.153E-02 (.926E-03)	.040 (.027)	.036 (.028)
Mother's Educ.	.211E-03 (.945E-04)	-.301E-04 (.233E-03)	.189E-03 (.236E-03)
Father's Educ.	.239E-03 (.105E-03)	.279E-03 (.181E-03)	.490E-03 (.183E-03)
Siblings	-.115E-03 (.564E-04)	-.100E-03 (.230E-03)	-.291E-03 (.233E-03)
Pupil/Teacher Ratio	-.622E-04 (.291E-04)	-.172E-04 (.636E-03)	-.109E-03 (.646E-03)
%Teachers MA/PhD	.731E-05 (.436E-05)	-.367E-04 (.128E-03)	-.783E-04 (.130E-03)
Teacher's Salary	-.351E-07 (.603E-07)	-.659E-05 (.259E-05)	-.587E-05 (.263E-05)
Male	-.710E-03 (.335E-03)	-.031 (.594E-02)	-.027 (.603E-02)
Black	.253E-02 (.109E-02)	.019 (.841E-02)	.015 (.853E-02)
Hispanic	.169E-02 (.749E-03)	.969E-02 (.964E-02)	.012 (.980E-02)
Education exponent [b in (14)]	.862 (.041)	1 assumed	1 assumed

(continued next page)

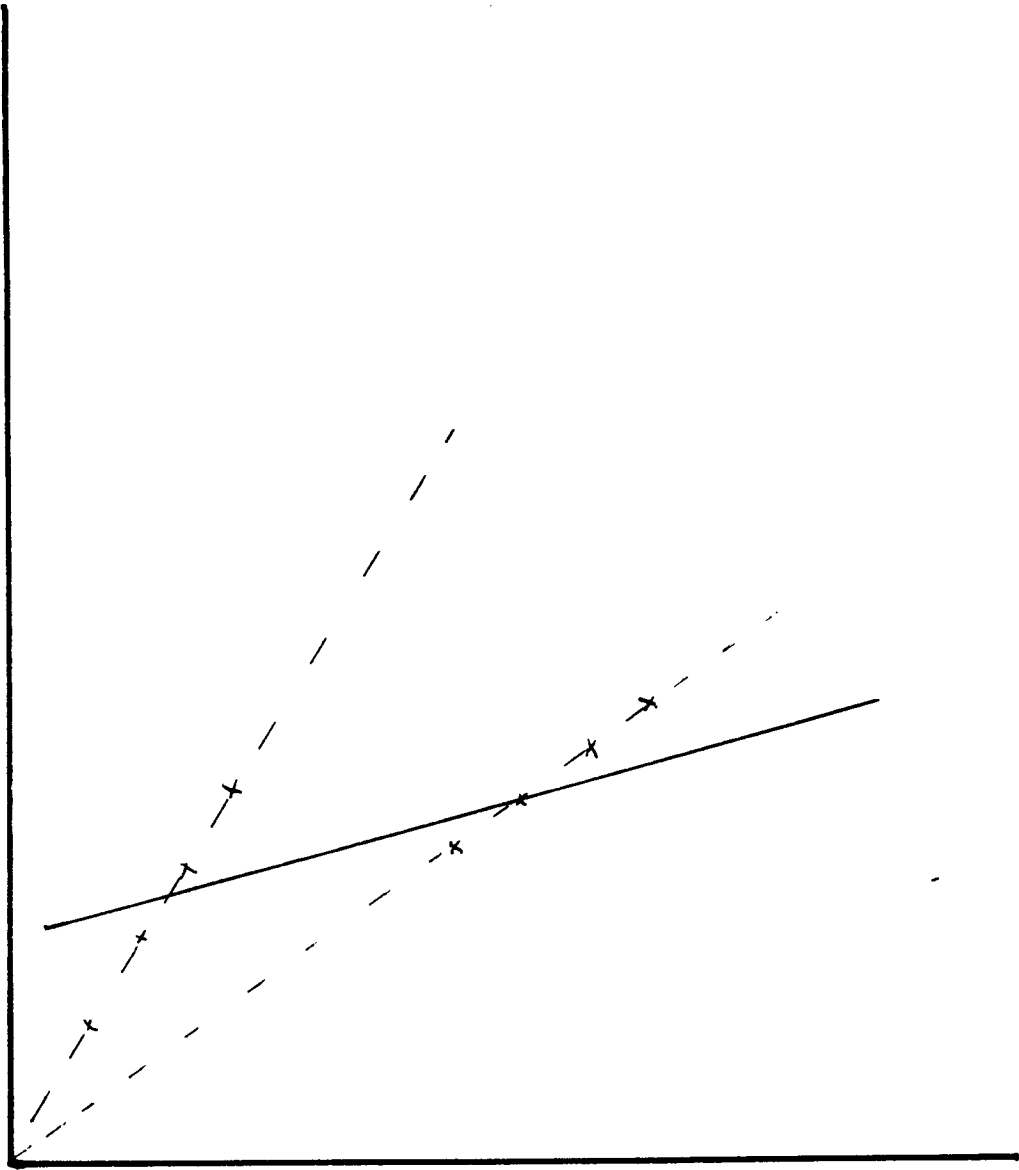
Table 3 (cont.)

(Linear terms)			
[B in (14)]			
Constant	4.222	4.597	4.484
	(.148)	(.541)	(.550)
Experience	.029	.030	.027
	(.319E-02)	(.319E-02)	(.322E-02)
Bad Health	-.093	-.084	-.096
	(.034)	(.034)	(.034)
Urban	.098	.098	.105
	(.018)	(.018)	(.018)
Unemployment Rate	-.080	-.084	-.084
	(.793E-02)	(.793E-02)	(.806E-02)
South	-.045	-.039	-.051
	(.015)	(.015)	(.015)
Never Married	-.081	-.084	-.089
	(.014)	(.014)	(.015)
In US at 14	.117	-.479	-.389
	(.068)	(.376)	(.382)
Pupil/Teacher Ratio	.264E-02	.846E-03	.293E-02
	(.122E-02)	(.816E-02)	(.829E-02)
%Teachers MA/PhD	.671E-03	.164E-02	.196E-02
	(.316E-03)	(.172E-02)	(.175E-02)
Teachers Salaries	.840E-05	.959E-04	.864E-04
	(.626E-05)	(.354E-04)	(.359E-04)
Male	.242	.636	.583
	(.014)	(.080)	(.081)
Black	-.128	-.280	-.318
	(.019)	(.111)	(.113)
Hispanic	.025	-.066	-.130
	(.022)	(.126)	(.128)
(Discount rate)			
[ln r* in (15)]			
Mean log r	-2.046	-	-
	(.033)		
Mean return to educ.	0.131	0.082	0.094
Max. return to educ.	0.149	0.144	0.144
Min. return to educ.	0.121	-0.012	0.020

Standard errors are in parentheses.

Figure 1

log
wage



schooling