TECHNICAL WORKING PAPER SERIES

BIASES IN TWIN ESTIMATES OF THE RETURN TO SCHOOLING: A NOTE ON RECENT RESEARCH

David Neumark

Technical Working Paper No. 158

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

Jere Behrman, Eli Berman, and McKinley Blackburn provided helpful comments. Orley Ashenfelter and Alan Krueger kindly supplied some additional estimates not provided in their 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.

NBER Technical Working Paper #158
June 1994

BIASES IN TWIN ESTIMATES OF THE RETURN TO SCHOOLING: A NOTE ON RECENT RESEARCH

ABSTRACT

Ashenfelter and Krueger's (1993) within-twin, measurement-error-corrected estimate of the return to schooling is about 13-16 percent. If their estimate is unbiased, then their results imply considerable downward measurement error bias in uncorrected within-twin estimates of the return to schooling, and considerable downward omitted ability bias in cross-section estimates.

This note points out that if there are ability differences among twins, then AK's IV estimator exacerbates the omitted ability bias in the within-twin estimate. Thus, upward omitted ability bias in within-twin estimates may provide an alternative explanation of the surprisingly high estimates of the return to schooling that AK obtain, and permit their results to be reconciled with upward, rather than downward omitted ability bias in cross-section estimates.

David Neumark
Department of Economics
University of Pennsylvania
Philadelphia, PA 19104
and NBER

I. Introduction

Within-twin estimates of the return to schooling hold out the promise of eliminating omitted ability bias. Earlier studies using identical twins (e.g., Behrman, et al., 1980) find that the within-twin estimate of the return to schooling is considerably lower than the cross-section estimate. However, Griliches (1979) points out that downward bias attributable to random measurement error in schooling may be exacerbated by differencing across twins, and suggests that this might explain the results of these earlier studies.

Ashenfelter and Krueger (1993, henceforth AK) propose correcting for this measurement error, using schooling as reported by each twin's sibling to construct an instrumental variable for self-reported schooling.[1] They report two main findings from data they collected on a new twin sample including self-reports and twin-reports of schooling. First, within-twin estimates of the return to schooling without correcting for measurement error (b_{WT}) are no lower, and if anything are somewhat higher, than OLS cross-section estimates (b_{LS}). Second, the measurement-error-corrected, within-twin estimate of the return to schooling (b_{IV}) is about 13-16 percent, as much as double b_{WT}. Their interpretation of the latter finding is that measurement error in schooling is seriously exacerbated in the within-twin estimate (b_{WT}).

Based on their findings, AK conclude that "the economic returns to schooling may have been badly underestimated in the past" (p. 1). But this underestimation cannot be attributed solely to measurement error in b_{WT}. If b_{IV} is unbiased, then the IV estimates of the return to schooling in the range

---

[1] An earlier version of this paper circulated as NBER Working Paper No. 4143.

of 13-16 percent imply that existing cross-section estimates are biased downward by omitted ability, perhaps by as much as 30-40 percent.

The purpose of this note is to offer an alternative interpretation of AK's IV estimates of the return to schooling, under which their estimates may be upward biased. Like AK's work, this interpretation builds on Griliches' (1979) argument that within-twin estimates do not necessarily eliminate all sources of bias in estimates of the return to schooling. AK's paper deals with the potential exacerbation of measurement error bias in within-twin estimates. But Griliches also pointed out that, as long as "ability" is not purely genetic, upward ability bias may persist, and even be exacerbated, in within-twin estimates.

The point that this note makes is that if there are ability differences among twins, then AK's IV estimator exacerbates the omitted ability bias in the within-twin estimate. Thus, upward omitted ability bias in within-twin estimates may provide an alternative explanation of the surprisingly high estimates of the return to schooling that AK obtain.

II. Ashenfelter and Krueger's Approach and Results

The starting point is a model relating the log wage to a linear function of schooling and unobserved ability,

(1) $\ln(w) = \beta S + \lambda A + \varepsilon$,

where $\text{plim}(S\cdot\varepsilon) = \text{plim}(A\cdot\varepsilon) = 0$. The equation can be interpreted as the regression of log wages on schooling and ability after partialling out other control variables. Assume for now that S is measured without error. Also assume that A is correlated with S through the equation

(2) $S = \gamma A + \eta$,

where $\text{plim}(\eta\cdot\varepsilon) = 0$. η can be interpreted as unmeasured "opportunities" for schooling, including factors such as access to financing, as in Becker's

Woytinsky Lecture schooling model (1975). Then when the estimated model excludes A, OLS estimates of β (b_{LS}) are biased, as

$$\operatorname{plim}(b_{LS}) = \beta + \lambda\sigma_{AS}/\sigma_S^2 = \beta + \lambda\gamma\sigma_A^2/\sigma_S^2 \quad . \tag{3}$$ [2]

If λ and γ are positive, as has been assumed in much of the literature, then b_{LS} is upward biased.

If ability is identical across twins, then b_{WT} is an unbiased estimate of the return to schooling, since A drops out of the within-twin equation

$$\Delta\ln(w) = \beta\Delta S + \Delta\varepsilon \quad . \tag{4}$$

However, measurement error in schooling may also cause b_{WT} to be less than b_{LS}. If S is now measured schooling, S^* true schooling, and $S = S^* + \nu$ with $\operatorname{plim}(S^* \cdot \nu) = 0$, then the differenced observed equation is

$$\Delta\ln(w) = \beta\Delta S - \beta\Delta\nu + \Delta\varepsilon \quad , \tag{5}$$

in which case

$$\operatorname{plim}(b_{WT}) = \beta\cdot[1 - \sigma_{\Delta\nu}^2/\sigma_{\Delta S}^2] = \beta\cdot[1 - \sigma_\nu^2/\{\sigma_S^2(1-\rho_S)\}] \quad , \tag{6}$$

where ρ_S is the correlation between reported schooling within twin pairs. As Griliches (1979) pointed out, as long as $\rho_S > 0$, b_{LS} is less downward biased than b_{WT} because of measurement error, since

$$\operatorname{plim}(b_{LS}) = \beta\cdot[1 - \sigma_\nu^2/\sigma_S^2] + \lambda\gamma\sigma_A^2/\sigma_S^2 \quad . \tag{7}$$ [3]

AK's innovation is to correct for measurement error in b_{WT} by instrumenting for ΔS with the difference between the schooling levels of each

---

[2] For expositional ease, this note sometimes uses the term bias to refer to asymptotic bias or inconsistency. Also, variance and covariance terms refer to population parameters, unless otherwise noted.

[3] In general, when the equation estimated includes other variables measured without error, the attenuation bias (i.e., $[1 - \sigma_\nu^2/\sigma_S^2]$) in equation (7) becomes $[1 - \sigma_\nu^2/\{\sigma_S^2(1-R^2)\}]$, where R^2 is from the regression of the error-ridden variable on the other variables (Griliches and Ringstad, 1971).

member of the twin pair, as reported by the other member. They assume that this difference, $\Delta S'$, also measures ΔS^* with error,

$$(8) \qquad \Delta S' = \Delta S^* + \Delta\nu' \quad ,$$

with $\text{plim}(\Delta S^* \cdot \Delta\nu') = 0$, but also assume that $\text{plim}(\Delta\nu \cdot \Delta\nu') = 0$. In this case, $\Delta S'$ is a valid instrument for ΔS in equation (5), and b_{IV}, the IV estimator using this instrument, is a consistent estimator of β.

In AK's Table 3, $b_{WT} = .092$ and $b_{IV} = .167$. On the other hand, for this specification

$$(9) \qquad \text{plim}(b_{WT}/b_{IV}) = [1 - \sigma_{\Delta\nu}^2/\sigma_{\Delta S}^2] = \sigma_{\Delta S',\Delta S}/\sigma_{\Delta S}^2 \quad .$$

AK's data on own- and twin-reports of schooling yield estimates of $\sigma_{\Delta S',\Delta S}/\sigma_{\Delta S}^2$ (the reliability ratio of schooling differences) of either .55 (2.158/3.902) or .58 (2.158/3.691), very close to the actual ratio b_{WT}/b_{IV} of .55. Based on these results, AK conclude that, conditional on the validity of their instrument, measurement error imparts a considerable downward bias to b_{WT}, and suggest that the return to schooling is about 16 percent.

Of course, AK's results imply that it is not just b_{WT} that is biased downward. Their results also imply that b_{LS}, the cross-section estimate of the return to schooling, is biased downward because of omitted ability. There are two ways to see this. First, b_{LS} in Table 3 is .084. Using AK's estimate of .88 of the reliability ratio in schooling levels, and the fact that the R^2 from the regression of schooling levels on the other regressors in Table 3 is .02, the attenuation bias in b_{LS} is .88. Thus, taking b_{IV} as a consistent estimate of β, the omitted ability bias in equation (7) is $.084 - .88 \cdot .167 = -.063$.[4] Second, if the only problem were measurement error that is

---

[4] The implied downward bias is slightly larger (.070) using their higher estimate of the reliability ratio of .92. Similar results are implied by other specifications and estimators reported by AK. In Table 5 they report

exacerbated in b_{WT}, then we would expect to find b_{WT}/b_{LS} equal to the ratio of the attenuation bias from measurement error in b_{WT} to that in b_{LS}, a ratio that is below one. In all of AK's estimates, though, the ratio b_{WT} to b_{LS} exceeds this predicted ratio, and in most of their estimates b_{WT}/b_{LS} exceeds one. For example, in Table 3 $b_{WT}/b_{LS} = 1.10$, whereas the predicted ratio based on the reliability ratios is .66. Again taking b_{IV} as a consistent estimate of β, this implies ability bias of -.059 in b_{LS}. Thus, considerable downward ability bias in b_{LS} is required to reconcile AK's alternative estimates of β, and their estimates of the reliability ratios of schooling levels and differences.[5]

III. Omitted Ability Bias in Within-Twin Estimates

While downward omitted ability bias in cross-section estimates of the

---

estimates adding to the wage equation additional variables that vary within twin pairs, in which case $b_{LS} = .094$ and $b_{IV} = .179$. In this case the lower estimate of the attenuation bias in b_{LS} is .85 (given that the R^2 from the regression of schooling levels on the levels of the regressors is .21), so that the implied ability bias is -.058.

In Table 6, AK use an alternative IV estimator that allows for correlated measurement error in a twin's report of his own and his sibling's schooling. In these estimates, the implied downward ability bias in b_{LS} is smaller, ranging from -.018 to -.030 depending on the specification (again using the lower estimate of the reliability ratio). Based on the discussion that follows, it is easy to show that any exacerbation of the bias in b_{WT} from using b_{IV} is lessened if there is positively correlated measurement error of this type and their alternative IV estimator is used.

[5] AK interpret the large difference between b_{IV} and b_{WT} as evidence that "measurement error may lead to considerable underestimation of the returns to schooling in studies based on siblings" (pp. 1-2). But they do not interpret the large difference between b_{IV} and b_{LS} as implying considerable downward ability bias in b_{LS}, instead concluding "we find some weak evidence that unobserved ability may be negatively related to schooling level" (p. 1).

return to schooling is conceivable (Griliches, 1977), the magnitudes (if not the sign) of this bias implied by AK's estimates may seem surprising. However, if the omitted ability bias is not completely eliminated in the within-twin difference, then there is an alternative potential explanation of AK's results that does not require downward ability bias in b_{LS}.

Griliches (1979) noted that if the "ability" that is rewarded in labor markets has more than a purely genetic component, then even MZ twins will not have identical ability. He showed that, in the absence of measurement error, within-twin or within-sibling estimates are more upward biased from omitted ability if the ratio of the variance of within-family ability differences to the variance of within-family schooling differences is greater than the ratio of their cross-section counterparts. To see this, let A have an individual (A') and family (A'') component, and similarly for η in equation (2). Then ignoring measurement error,

$$(10) \qquad \mathrm{plim}(b_{LS}) = \beta + \lambda\gamma\sigma_{A}^{2}/\sigma_{S}^{2} = \beta + \lambda\gamma\sigma_{A}^{2}/\{\gamma^{2}\sigma_{A}^{2} + \sigma_{\eta}^{2}\}$$

$$(11) \qquad \mathrm{plim}(b_{WT}) = \beta + \lambda\gamma\sigma_{\Delta A'}^{2}/\sigma_{\Delta S'}^{2} = \beta + \lambda\gamma\sigma_{\Delta A'}^{2}/\{\gamma^{2}\sigma_{\Delta A'}^{2} + \sigma_{\Delta\eta'}^{2}\} \quad .$$

While $\sigma_{\Delta A'}^{2}$ is undoubtedly smaller than σ_{A}^{2}, the omitted ability bias in b_{WT} can be greater than that in b_{LS} if σ_{η}^{2} is sufficiently greater than $\sigma_{\Delta\eta'}^{2}$. This may occur if parents treat their twin offspring similarly in financing schooling, partly ignoring differences in A, or perhaps even offsetting them (in which case $\sigma_{\Delta\eta'\Delta A'} \neq 0$).[6] Thus, upward ability bias in b_{WT} can explain

---

[6] Existing research points to some evidence of differences in ability or endowment within twin pairs. Jencks and Brown (1977) report correlations of an academic aptitude test of .86 among MZ twins in the Project Talent study, and Behrman, et al. (1980) report correlations of .76 among MZ twins on the General Classification Test, for the Navy subsample of the NRC twin sample. However, these correlations may fall short of one because of measurement error, rather than because of true ability differences.

In addition, Behrman, et al. (1993) report that for 69 percent of the why b_{WT} exceeds b_{LS}, despite the greater attenuation bias in b_{WT}, without resorting to the conclusion that b_{LS} is biased downward.

Moreover, upward ability bias in the within-twin estimator may have implications for b_{IV}. If there are individual differences in ability within twin pairs, then upward omitted ability bias in b_{WT} is exacerbated in b_{IV}. The plims of the alternative estimators of β when there is both omitted ability and measurement error are given by:

$$\text{(12)} \qquad \text{plim}(b_{LS}) = \beta\cdot[1 - \sigma_{\nu}^2/\sigma_S^2] + \lambda\gamma\sigma_A^2/\sigma_S^2$$

$$\text{(13)} \qquad \text{plim}(b_{WT}) = \beta\cdot[1 - \sigma_{\Delta\nu}^2/\sigma_{\Delta S}^2] + \lambda\gamma\sigma_{\Delta A'}^2/\sigma_{\Delta S}^2$$

$$\text{(14)} \qquad \text{plim}(b_{IV}) = \beta + \lambda\gamma\sigma_{\Delta A'}^2/\sigma_{\Delta S}^{*2} \quad .$$

If the omitted ability bias is positive, then equation (14) implies that b_{IV} is upward biased. Furthermore, because $\sigma_{\Delta S}^2 > \sigma_{\Delta S}^{*2}$, the upward ability bias in b_{IV} exceeds that in b_{WT}. Given AK's estimate of $\sigma_{\Delta S}^2/\sigma_{\Delta S}^{*2}$ of 1/.58 = 1.72, the implied upward ability bias in b_{IV} is 72 percent larger than that in b_{WT}.[7] Finally, b_{WT} may provide a less biased estimate than b_{IV} because the

---

identical twin pairs in their sample, birthweights differ by at least four ounces, and that for 48 percent birthweights differ by at least eight ounces. They also report that a four-ounce birthweight advantage is associated with a schooling advantage of almost one-half year.

Finally, there is some evidence that within-twin variation in ability has a smaller effect on schooling than does cross-section variation in ability, suggesting that bias from omitted ability could be exacerbated in within-twin differences. Griliches (1979) reports that there is a significantly lower effect of IQ on schooling within families than across families, and that this attenuation is more severe for identical twins. However, Behrman, et al. (1993) argue that these results may stem from unobserved differences in endowments that are smaller within twin pairs than across individuals.

[7] For the within-twin estimator allowing for correlated measurement error, the estimate of $\sigma_{\Delta S}^2/\sigma_{\Delta S}^{*2}$ = 1/.84 = 1.19, so the omitted ability bias in b_{IV} is 19 percent larger than that in b_{WT}. But the contrast between b_{IV} and b_{WT} for this estimator is considerably smaller to begin with, as the ratio b_{IV}/b_{WT} = 1.21 (.129/.107), vs. 1.82 (.167/.092) for the estimator assuming classical

measurement error and omitted ability bias in b_{WT} are offsetting, whereas b_{IV} has only upward ability bias.

It is also worth noting that there is no inconsistency between the fact that the ratio b_{WT}/b_{IV} "matches" the noise-to-signal ratio in schooling differences--as is the case in AK's estimates--and the claim that b_{IV} may be upward biased because of omitted ability differences between twins. To see this, rewrite the plim of b_{WT}

$$\text{(15)} \qquad \text{plim}(b_{WT}) = \beta\cdot[1 - \sigma_{\Delta\nu}{}^2/\sigma_{\Delta S}{}^2] + \{\lambda\gamma\sigma_{\Delta A'}{}^2/\sigma_{\Delta S^*}{}^2\}\cdot\{\sigma_{\Delta S^*}{}^2/\sigma_{\Delta S}{}^2\}$$
$$= \{\beta + \lambda\gamma\sigma_{\Delta A'}{}^2/\sigma_{\Delta S^*}{}^2\}\cdot[1 - \sigma_{\Delta\nu}{}^2/\sigma_{\Delta S}{}^2] \quad .$$

Equations (14) and (15) imply that, just as in the pure measurement error case (equation (9)), $\text{plim}(b_{WT}/b_{IV}) = [1 - \sigma_{\Delta\nu}{}^2/\sigma_{\Delta S}{}^2]$. Thus, given AK's estimate of $\sigma_{\Delta\nu}{}^2/\sigma_{\Delta S}{}^2$, the ratio of b_{WT} to b_{IV} should be the same as that implied by pure measurement error, even when there is upward ability bias in both estimators.

Equations (12), (14), and (15) cannot be solved for the three unknowns β, $\lambda\gamma\sigma_A{}^2/\sigma_S{}^2$, and $\lambda\gamma\sigma_{\Delta A'}{}^2/\sigma_{\Delta S^*}{}^2$, because equations (14) and (15) are linearly dependent. Nonetheless, these equations can be used to suggest some bounds on β. For example, for the estimates in Table 3, assuming that omitted ability bias in the cross-section estimate is positive, equation (12) suggests that β satisfies $b_{LS} - \beta\cdot[1 - \hat{\sigma}_\nu{}^2/\hat{\sigma}_S{}^2] > 0$, or $\beta < .098$. This upper bound for β is below the IV estimate of the return to schooling reported by AK, while suggesting upward ability bias of .035 in b_{WT}.

Finally, note that endogeneity bias that persists in within-twin estimates will also be exacerbated by AK's IV estimator.[8] Dropping the

---

measurement error.

[8] Behrman, et al. (1993) is perhaps the only paper that studies endogeneity bias in twin differences. Using birthweight differences as an instrument for

assumption that $\mathrm{plim}(\eta\cdot\varepsilon) = 0$, in the presence of measurement error,

(16) $$\mathrm{plim}(b_{LS}) = \beta\cdot[1 - \sigma_{\nu}^{2}/\sigma_{S}^{2}] + \sigma_{\varepsilon\eta}/\sigma_{S}^{2}$$

(17) $$\mathrm{plim}(b_{WT}) = \beta\cdot[1 - \sigma_{\Delta\nu}^{2}/\sigma_{\Delta S}^{2}] + \sigma_{\Delta\varepsilon\Delta\eta}/\sigma_{\Delta S}^{2}$$

(18) $$\mathrm{plim}(b_{IV}) = \beta + \sigma_{\Delta\varepsilon\Delta\eta}/\sigma_{\Delta S}^{*2} \quad .$$

If $\sigma_{\varepsilon\eta} < 0$, as Griliches (1977) suggested, then b_{LS} may be biased downward more than b_{WT}, which again can explain AK's finding that b_{LS} is less than b_{WT}. Furthermore, equation (18) shows that any endogeneity bias in b_{WT} will be exacerbated in b_{IV}. The net effect, though, is unclear. If $\sigma_{\Delta\varepsilon\Delta\eta} < 0$ (which seems likely, if $\sigma_{\varepsilon\eta} < 0$), then in contrast to the omitted ability case, b_{IV} is downward biased, in which case the true return to schooling may exceed AK's IV estimates. If instead $\sigma_{\Delta\varepsilon\Delta\eta} > 0$, then b_{WT} and b_{IV} are upward biased, as in the omitted ability case, and b_{WT} may be less biased than b_{IV}, since b_{WT} has offsetting downward and upward biases, whereas b_{IV} is only biased upwards.[9]

IV. Conclusion

The rationale for within-twin estimation of the return to schooling is the presumption that identical twins have equal ability, which drops out of

---

schooling differences, they find no evidence of endogeneity bias in within-twin estimates of the return to schooling, as the estimated return is unchanged after instrumenting. However, as the authors note, there is no direct evidence that birthweight does not also affect wages directly, independently of its effect on schooling, in which case it is not a valid instrument. Furthermore, that study does not attempt to estimate the endogeneity bias in cross-section estimates, with which to compare the bias in the within-twin estimates. This dearth of evidence reflects the difficulty of thinking of instruments for schooling that vary across twins, and hence could be used to correct within-twin estimates for endogeneity bias.

[9] In either case, as in the omitted ability case, the same ratio of b_{WT} to b_{IV} should be obtained as that implied by pure measurement error, since it is still true that $\mathrm{plim}(b_{WT}/b_{IV}) = [1 - \sigma_{\Delta\nu}^{2}/\sigma_{\Delta S}^{2}]$.

the within-twin difference. However, this does not explain the source of schooling differences within twin pairs. The notion that within-twin estimates provide a "natural experiment" for estimating the return to schooling is based on the assumption that schooling differences within twin pairs represent random (true) variation. However, once alternative reasons for schooling differences among twins are considered, the conditions for this experiment to be valid may be violated, and may, in some circumstances, imply that the bias in within-twin estimates is greater than that in cross-section estimates. That was the message of Griliches' (1979) review of sibling and twin studies.

Ashenfelter and Krueger's approach to correcting for measurement error in schooling in within-twin estimates of the return to schooling is potentially valuable because it can tell us whether lower within-twin than cross-section estimates are attributable to the removal of ability bias, or the exacerbation of measurement error bias. However, if biases other than measurement error remain in the within-twin estimates, these biases may be exacerbated by AK's within-twin IV estimator, and this estimator may be upward biased.

References

Ashenfelter, Orley, and Alan Krueger. 1993. "Estimates of the Economic Return to Schooling from a New Sample of Twins." Mimeograph, Princeton University.

Becker, Gary S. 1975. Human Capital (Chicago: University of Chicago Press).

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

Behrman, Jere R., Mark R. Rosenzweig, and Paul Taubman. 1992. "Endowments and the Allocation of Schooling in the Family and in the Marriage Market: The Twin Experiment." Mimeograph.

Griliches, Zvi. 1979. "Sibling Models and Data in Economics: Beginnings of a Survey." *Journal of Political Economy* 87(5, Pt. 2): S37-64.

________. 1977. "Estimating the Returns to Schooling: Some Econometric Problems." *Econometrica* 45(1): 1-22.

Jencks, Christopher, and Marsha Brown. 1977. "Genes and Social Stratification: A Methodological Exploration with Illustrative Data." In Paul Taubman, Ed., Kinometrics: Determinants of Socioeconomic Success Within and Between Families (Amsterdam: North-Holland Publishing Company), pp. 169-233.