Biases in twin estimates of the return to schooling

David Neumark *

Department of Economics, Michigan State University, East Lansing, MI 48824-1038, U.S.A.

Received 21 March 1996; accepted 12 February 1997

Abstract

Recent within-twin estimates of the return to schooling that use instrumental variables to correct for measurement error are considerably higher than existing estimates. This paper extends Griliches’ (Griliches (1979) Sibling models and data in economics: beginnings of a survey, Journal of Political Economy 87, S37–S64) analysis of sibling or twin estimates of the return to schooling to show that small ability differences among twins can lead to more upward omitted ability bias—and more upward bias overall—in the instrumental variables estimate that corrects for measurement error than in the standard within-twin estimate. This point is relevant for any application in which differencing may not fully remove the omitted variable, and may explain high within-twin instrumental variable estimates of the return to schooling. [JELI21] © 1999 Elsevier Science Ltd. All rights reserved.

1. Introduction

Within-twin estimates of the return to schooling hold out the promise of eliminating omitted ability bias. Earlier studies using identical twins find within-twin estimates that are considerably lower than cross-section estimates (e.g. Behrman et al., 1980). However, such findings may be attributable to exacerbation of bias from measurement error in schooling that is caused by differencing across twins (Griliches, 1979).

Ashenfelter and Krueger (1994) (henceforth AK) collect data on twin pairs including schooling of each individual reported by that person and their twin, to obtain an instrumental variable (IV) with which to correct within-twin differences for measurement error. They find that the within-twin IV estimates of the return to schooling are in the range of 13–18%, as much as double the standard within-twin estimate. AK interpret this evidence as implying that measurement error in schooling is seriously exacerbated in the within-twin estimate, and that “the economic returns to schooling may have been underestimated in the past” (p. 1157).

However, AK also report that within-twin estimates of the return to schooling without correcting for measurement error are high relative to OLS cross-section estimates, compared with what is predicted based on the extent of measurement error in schooling levels and differences, and upward omitted ability bias in the cross-section estimates. This may be attributable to downward, rather than upward omitted ability bias in the cross-section estimates of the return to schooling. However, the high within-twin estimates may also be attributable to remaining ability differences between twins that, although smaller, lead to greater upward bias in standard within-twin estimates than in the cross-section estimates (Griliches, 1979).

Building on the possibility of upward omitted ability bias in standard within-twin estimates, this paper extends the analysis in Griliches (1979), and shows that the within-twin IV estimator amplifies the bias from any omitted ability differences between twins, relative to the standard within-twin estimator. Moreover, it shows that if omitted ability biases cross-section estimates of the return to schooling upward, and is not fully removed by differencing within twin pairs, then the within-twin IV estimator is upward biased (possibly substantially) relative to the standard within-twin estimator, and possibly...
also relative to the cross-section estimator. This point is relevant for any application in which instrumental variables estimation is used for differenced data, when the differencing may not fully eliminate the omitted variable. In addition, it may explain the unusually high estimates of the return to schooling that AK obtain, and reconcile their results with upward, rather than downward omitted ability bias in cross-section estimates of the return to schooling.

2. Within-twin estimation of the return to schooling

The starting point is a model relating the log wage to a linear function of schooling \((S)\) and unobserved ability \((A)\),

\[
\ln(w) = \beta S + \lambda A + \epsilon, \tag{1}
\]

where \(\text{plim}(A_\epsilon) = \text{plim}(A_\epsilon) = 0\). “Ability” is a catch-all for omitted variables that affect wages and may be correlated with schooling. Eq. (1) 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

\[
S = \gamma A + \eta, \tag{2}
\]

where \(\text{plim}(\eta_\epsilon) = 0\). \(\eta\) can be interpreted as unmeasured “opportunities” for schooling, including factors such as access to financing (Becker, 1975, Chapter IV). When the estimated model excludes \(A\), OLS estimates of \(\beta\) \((b_{LS})\) are biased (asymptotically), as

\[
\text{plim}(b_{LS}) = \beta + \lambda \frac{\sigma_{s,\epsilon}}{\sigma_s^2} = \beta + \lambda \frac{\sigma^2_{\epsilon}}{\sigma_s^2}. \tag{3}
\]

If \(\lambda\) and \(\gamma\) are positive, as has been assumed in much of the literature, then \(b_{LS}\) is upward biased.\(^1\)

If \(A\) is identical across twins, then it drops out of the within-twin equation

\[
\Delta \ln(w) = \beta \Delta S + \Delta \epsilon, \tag{4}
\]

and the standard within-twin estimate \((b_{WT})\), which is the OLS estimate of \(\beta\) in Eq. (4), is unbiased. However, measurement error in schooling may also cause \(b_{WT}\) to be less than \(b_{LS}\). If \(S^*\) is true schooling, and measured schooling is \(S = S^* + \nu\) with \(\text{plim}(S^*\nu) = 0\), then the differenced observed equation is

\[
\Delta \ln(w) = \beta \Delta S - \beta \Delta \nu + \Delta \epsilon, \tag{5}
\]

and

\[
\text{plim}(b_{WT}) = \beta [1 - \frac{\sigma^2_{s,\epsilon}}{\sigma^2_s}] = \beta [1 - \frac{\sigma^2_{\epsilon}}{\sigma_s^2}(1 - \rho_s)]. \tag{6}
\]

\(\rho_s\) is the correlation between reported schooling within twin pairs. As Griliches (1979) pointed out, as long as \(\rho_s > 0\), the downward measurement error bias in \(b_{LS}\) is less than that in \(b_{WT}\), since in the presence of measurement error\(^2\)

\[
\text{plim}(b_{LS}) = \beta [1 - \frac{\sigma^2_{\epsilon}}{\sigma_s^2}] + \lambda \frac{\sigma^2_{\epsilon}}{\sigma_s^2}. \tag{7}
\]

3. The within-twin IV estimator

AK correct for measurement error in \(b_{WT}\) by instrumenting for \(\Delta S\) with the difference between the schooling levels of each member of the twin pair, as reported by the other member. They assume that this difference, \(\Delta S^*\), also measures \(\Delta S^*\) with error,

\[
\Delta S' = \Delta S^* + \Delta \nu', \tag{8}
\]

with \(\text{plim}(\Delta S^* \times \Delta \nu') = 0\), and assume that \(\text{plim}(\Delta \nu' \times \Delta \nu') = 0\). In this case, \(\Delta S'\) is a valid instrument for \(\Delta S\) in Eq. (5), and, maintaining our earlier assumptions, the IV estimator \((b_{IV})\) is a consistent estimator of \(\beta\). This approach 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. Following AK’s lead, Behrman et al. (1994) use a similar measurement error correction, using children’s reports of twin fathers’ schooling to construct an instrument, and Miller et al. (1995) carry out a replication of AK’s analysis using Australian Twin Register Data.

In AK’s Table 3, \(b_{WT} = 0.092\) and \(b_{IV} = 0.167\). On the other hand, for this specification

\[
\text{plim}(b_{WT}/b_{IV}) = [1 - \frac{\sigma^2_{s,\epsilon}}{\sigma^2_s}] = \frac{\sigma^2_{s,\Delta S}/\sigma^2_{\Delta S}}{s^2}. \tag{9}
\]

AK’s data on own and twin reports of schooling yield estimates of \(\sigma^2_{s,\Delta S}/\sigma^2_{\Delta S}\) (the reliability ratio of schooling differences) of either 0.55 (2.158/3.902) or 0.58 (2.158/3.691), very close to the actual ratio \(b_{WT}/b_{IV}\) of 0.55. 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 closer to 16%.

\(^1\)Throughout, bias is used to refer to asymptotic bias or inconsistency. Also, variance and covariance terms refer to population parameters, unless otherwise noted.

\(^2\)When the equation estimated includes other variables measured without error, the noise-to-signal ratios \((\sigma^2_{s,\epsilon}/\sigma^2_s)\) in Eq. (7) and \(\sigma^2_{s,\epsilon}/\sigma^2_s\) in Eq. (6) are divided by \((1 - R^2)\), where \(R^2\) is from the regression of the error-ridden variable on the other included variables (Griliches and Ringstad, 1971). This adjustment is used below when AK’s estimates are used to compute the implied noise-to-signal ratios.
4. Omitted ability bias in within-twin estimates

Although measurement error in schooling also leads to some downward bias in the cross-section estimate $b_{LS}$, if $b_{IV}$ is consistent then AK’s results also imply that omitted ability biases $b_{LS}$ downward, in contrast to the more common presumption. To see this, note that if the only problem is measurement error in schooling that is exacerbated in $b_{WT}$, then we would expect $b_{WT}/b_{LS}$ to equal the ratio of the attenuation bias from measurement error in $b_{WT}$ to that in $b_{LS}$, a ratio that is below one. We would expect $b_{WT}/b_{LS}$ to be less than this ratio if there is upward omitted ability bias in $b_{LS}$. In all of AK’s estimates, however, $b_{WT}/b_{LS}$ exceeds this predicted ratio, and in some of their estimates it exceeds one. For example, in Table 3 $b_{WT}/b_{LS} = 1.10$, whereas the predicted ratio based on the reliability ratios is 0.66. If $b_{IV}$ is a consistent estimate of $\beta$, then the omitted ability bias in $b_{LS}$ is $-0.059$. 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 twin’s schooling. In these estimates, the implied downward ability bias in $b_{LS}$ is required to reconcile AK’s alternative estimates of $\beta$ with their estimates of the reliability ratios of schooling levels and differences.

While substantial downward omitted ability bias in cross-section estimates of the return to schooling is conceivable (Griliches, 1977), $b_{WT}$ may be high relative to $b_{LS}$ for a different reason. Although differing within twin pairs surely eliminates much of the omitted ability, Griliches (1979) notes that if the “ability” that is rewarded in labor markets has more than a purely genetic component, then even among monozygotic (MZ) twins ability differences will remain. Griliches shows that $b_{WT}$ may even be more upward biased from omitted ability than is $b_{LS}$, when within-family ability differences are an important determinant of within-family schooling differences. In a particular example he develops, $b_{WT}$ is more upward biased if the ratio of the variance of within-family ability differences to the overall variance of ability is greater than the ratio of the variance of within-family schooling differences to the overall variance of schooling, or if ability is more “familial” than schooling. The intuition is straightforward. To take an extreme example, when the within-family ability variance is zero, none of the within-family variance in schooling is attributable to ability differences; this is the case in which twins provide an ideal “natural experiment.” In contrast, when the within-family ability variance is relatively high, within-family schooling differences are more likely to reflect within-family ability differences, in which case within-twin variation may not provide a better experiment than cross-sectional variation. A similar problem to that noted by Griliches can arise if the partial regression coefficient of ability on schooling is weaker within twin pairs, but if the ability–earnings link (captured in $\lambda$ in Eq. (1)) is stronger within twin pairs than in a cross-section, perhaps because there is stronger specialization—based on ability—in earnings enhancing activities among twins than in a cross-section.

The existing research, in fact, finds some evidence of differences in ability (or endowments) within twin pairs, although it has not explored direct effects of these differences on within-twin wage differences. Jencks and Brown (1977) report correlations of an academic aptitude test of 0.86 among MZ twins in the Project Talent study, and Behrman et al. (1980) report correlations of 0.76 among MZ twins on the General Classification Test, for the Navy subsample of the NRC twin sample. Of course, these correlations may fall short of one because of measurement error, rather than because of true ability differences. Behrman et al. (1994) report more convincing evidence. For 69% of the identical twin pairs in their Minnesota Twins Registry sample, birthweights differ by at least four ounces, and for 48% birthweights differ by at least eight ounces. Furthermore, birthweight appears to be a relevant endowment, as a four ounce birthweight advantage is associated with a schooling advantage of almost one-half year (although the authors assume that birthweight has no independent effect on wages). Of course, we can only speculate regarding within-twin differences in unobservables.

Thus, upward omitted ability bias in $b_{WT}$, rather than downward omitted ability bias in $b_{LS}$, can potentially explain why in AK’s data $b_{WT}/b_{LS}$ exceeds the ratio predicted on the basis of measurement error bias in $b_{LS}$ and $b_{WT}$ upward ability bias in $b_{LS}$. AK acknowledge that unobserved differences may persist among identical twins, biasing within-twin estimates (p. 1171). However, they do not explore the implications of such unobserved differences for their IV estimator.

5. Implications of within-twin ability differences for the IV estimator

Although AK’s IV estimator eliminates measurement error bias in the within-twin estimate, it amplifies the omitted variable bias from any ability differences within twin pairs, possibly substantially. This may explain the...
surprisingly high estimates of the return to schooling that they obtain from the within-twin IV estimator.

Letting \( A \) have individual (\( A' \)) and family (\( A'' \)) components, the plims of \( b_{\text{WT}} \) and \( b_{\text{IV}} \), when there is both omitted ability and measurement error, are:

\[
\text{plim}(b_{\text{WT}}) = \beta [1 - \sigma_{\Delta A}^2/\sigma_{\Delta S}^2] + \lambda \sigma_{\Delta A}^2/\sigma_{\Delta S}^2 \tag{10}
\]

\[
= \beta [1 - \sigma_{\Delta A}^2/\sigma_{\Delta S}^2] + \lambda \gamma \sigma_{\Delta A}^2/\sigma_{\Delta S}^2 \tag{11}
\]

\[
\text{plim}(b_{\text{IV}}) = \beta + \lambda \sigma_{\Delta A}^2/\sigma_{\Delta S}^2 = \beta \tag{11}
\]

\[+ \lambda \gamma \sigma_{\Delta A}^2/\sigma_{\Delta S}^2.\]

If the omitted ability bias is positive, then Eq. (11) implies that \( b_{\text{IV}} \) is upward biased. Furthermore, because \( \sigma_{\Delta S}^2 < \sigma_{\Delta A}^2 \) (i.e. the variance of true schooling is less than the variance of reported schooling), the upward ability bias in \( b_{\text{IV}} \) exceeds that in \( b_{\text{WT}} \). The intuition parallels the earlier result from Griliches (1979): given that measurement error contributes relatively more to the within-family variance of schooling than to the overall variance of schooling, ability is more “familial” relative to true schooling than to measured schooling. Moreover, the amplification of the omitted ability bias in \( b_{\text{IV}} \) can be substantial. Given AK’s estimate of \( \sigma_{\Delta A}^2/\sigma_{\Delta S}^2 \) of 1/0.58 = 1.72, the implied upward ability bias in \( b_{\text{IV}} \) is 72% larger than that in \( b_{\text{WT}} \).

It is easy to show that any amplification of the omitted ability bias in \( b_{\text{WT}} \) from instrumenting is lessened but nonetheless persists using AK’s alternative IV estimator for the case of correlated measurement error. For the within-twin estimator allowing for correlated measurement error, the estimate of \( \sigma_{\Delta A}^2/\sigma_{\Delta S}^2 \) of 1/0.84 = 1.19, so the omitted ability bias in \( b_{\text{IV}} \) is 19% larger than that in \( b_{\text{WT}} \). But the contrast between \( b_{\text{IV}} \) and \( b_{\text{WT}} \) for this estimator is considerably smaller to begin with, as the ratio \( b_{\text{IV}}/b_{\text{WT}} = 1.21 \times 0.129/0.107, \) vs 1.82 \times 0.167/0.092 \) for the estimator assuming classical measurement error.

The amplification of upward omitted ability bias in within-twin estimates by the IV estimator would not necessarily be troubling if, overall, \( b_{\text{IV}} \) nonetheless provided a less biased estimate of \( \beta \). However, Eqs. (10) and (11) suggest that this is unlikely to be the case. The measurement error and omitted ability biases in \( b_{\text{WT}} \) would be in opposite directions, and, as suggested by at least some recent research using alternative approaches, are of similar absolute magnitudes (Ashenfelter and Zimmerman, 1993; Angrist and Krueger, 1991). In contrast, \( b_{\text{IV}} \) would have only upward omitted ability bias.

It is also worth noting that there is no inconsistency between the ratio \( b_{\text{IV}}/b_{\text{WT}} \) closely matching the reliability ratio in schooling differences—as predicted if the measurement error bias in \( b_{\text{WT}} \) is eliminated in \( b_{\text{IV}} \), and differencing within twin pairs eliminates \( A \) (see Eq. (9))—and the claim that the difference between \( b_{\text{IV}} \) and \( b_{\text{WT}} \) may reflect bias from omitted ability differences between twins that is amplified by AK’s IV estimator. To see this, rewrite the plim of \( b_{\text{WT}} \) as

\[
\text{plim}(b_{\text{WT}}) = \beta [1 - \sigma_{\Delta A}^2/\sigma_{\Delta S}^2] + [\lambda \gamma \sigma_{\Delta A}^2/\sigma_{\Delta S}^2] \cdot [1 - \sigma_{\Delta A}^2/\sigma_{\Delta S}^2] \tag{12}
\]

Eqs. (11) and (12) imply that, just as in the pure measurement error case, \( \text{plim}(b_{\text{IV}}/b_{\text{WT}}) = [1 - \sigma_{\Delta A}^2/\sigma_{\Delta S}^2] \). Thus, given AK’s estimate of \( \sigma_{\Delta A}^2/\sigma_{\Delta S}^2 \), the ratio of \( b_{\text{IV}} \) to \( b_{\text{WT}} \) should be the same as that implied by pure measurement error, even when there is upward ability bias in both estimators.

Eqs. (7), (11) and (12) cannot be solved for the three unknowns \( \beta, \lambda \gamma \sigma_{\Delta A}^2/\sigma_{\Delta S}^2 \), and \( \lambda \gamma \sigma_{\Delta A}^2/\sigma_{\Delta S}^2 \), because the latter two equations are linearly dependent. Nonetheless, they can be used to suggest some bounds. For example, for the estimates in AK’s Table 3, assuming that omitted ability bias in the cross-section estimate is positive, Eq. (7) suggests that \( \beta \) satisfies \( b_{\text{IV}} - \beta [1 - \sigma_{\Delta A}^2/\sigma_{\Delta S}^2] > 0 \), or \( \beta < 0.098 \). This upper bound for \( \beta \) is below the IV estimate of the return to schooling reported by AK, and consistent with upward ability bias of 0.035 in \( b_{\text{WT}} \).

This section focuses on the implications of omitted ability bias that persists in the within-twin estimates for AK’s IV estimator of the return to schooling. It is also the case, however, that endogeneity bias that persists in within-twin estimates will be amplified by AK’s IV estimator.\(^3\) Dropping the assumption that \( \text{plim}(\eta | e) = 0 \), in the presence of measurement error,

\[
\text{plim}(b_{\text{WT}}) = \beta [1 - \sigma_{\Delta A}^2/\sigma_{\Delta S}^2] + \sigma_{\Delta \Delta \eta}/\sigma_{\Delta S}^2 \tag{13}
\]

\[
\text{plim}(b_{\text{IV}}) = \beta + \sigma_{\Delta \Delta \eta}/\sigma_{\Delta S}^2. \tag{14}
\]

Eq. (14) shows that any endogeneity bias in \( b_{\text{WT}} \) will be amplified in \( b_{\text{IV}} \). Whether \( b_{\text{IV}} \) is more or less biased than \( b_{\text{WT}} \) depends, however, on the sign of \( \sigma_{\Delta \Delta \eta} \).

6. A partial evaluation of the competing explanations

The previous section shows that the exacerbation of upward omitted ability bias in the within-twin estimate is a possible explanation of the high returns to schooling that AK estimate, but it does not prove that this source of bias explains their results. From Eqs. (7) and (10), we know that

---

\(^3\) Behrman et al. (1994) use birthweight differences as an instrument for schooling differences, and find no evidence of endogeneity bias in within-twin estimates of the return to schooling. They acknowledge, however, that birthweight may also affect wages directly, in which case it is not a valid instrument.
\[ \text{plim}(b_{WT}/b_{S})) = \frac{\beta + \lambda \sigma_{A^T,SA^T}/\sigma_{A^S}^2}{\beta + \lambda \sigma_{A,SA}/\sigma_{A}^2} \cdot \frac{RR_{A^T}}{RR_{A}}. \]  

where \( RR_{A^T} \) and \( RR_A \) are the reliability ratios of schooling differences and levels, respectively. Because, as noted in Section 4, \( b_{WT}/b_{S} \) exceeds \( RR_{A^T}/RR_{A} \), the explanation based on exacerbation of upward omitted ability bias in the within-twin estimate implies that \( \sigma_{A^T,SA^T}/\sigma_{A^S}^2 \) (assuming that \( \lambda > 0 \)). Alternatively, if differencing eliminates \( A \), so that \( \sigma_{A^T,SA^T} = 0 \), then AK’s results imply that either \( \lambda < 0 \), or \( \sigma_{A,SA}/\sigma_{A}^2 \) (with the other > 0). They prefer the latter explanation, concluding that “we find some weak evidence that unobserved ability may be negatively related to schooling level” (p. 1157). Obviously, in the absence of data on the variables that underlie \( A \)—which is the reason twin data are used in the first place—we cannot determine which of these scenarios is more likely to hold. Moreover, even a data set on twins that has information on some component of \( A \), such as a test score, will not be decisive because this component will almost surely fail to capture all of the unobservables that underlie \( A \). Nonetheless, such data can provide suggestive evidence, and can help to answer the more general question of whether there is likely to be omitted ability bias that persists in within-twin estimates, whether or not it is exacerbated relative to the cross-section estimates.

Evidence on test scores indicates quite clearly that insofar as these scores capture the relevant components of \( A \), \( \lambda \) is greater than zero; that is, test scores are positively associated with wages (see Blackburn and Neumark, 1995, and the literature reviewed therein). AK appear to agree with this conclusion, as indicated by the above quote.

We can obtain some information on \( \sigma_{A^T,SA^T}/\sigma_{A^S}^2 \) and \( \sigma_{A,SA}/\sigma_{A}^2 \) from Table 2 in Jencks and Brown (1977), for a sample of MZ twins from the Project Talent data. Computing these two ratios, correcting for measurement error in schooling using the reliability ratios for schooling estimated by AK (since measurement error in schooling affects these estimates), we find that \( \sigma_{A^T,SA^T}/\sigma_{A^S}^2 = 1.37 \), while \( \sigma_{A,SA}/\sigma_{A}^2 = 3.34 \). (Note that we have used \( T \) instead of \( A \) to indicate a test score that most likely does not capture all of \( A \).) These estimates lead to three tentative conclusions. First, the evidence from these particular test scores does not indicate that omitted ability bias is exacerbated within twin pairs in the Project Talent data, as the first ratio is smaller than the second. Second, at the same time, this evidence indicates that omitted ability bias still plagues within-twin estimates, since \( \sigma_{A^T,SA^T}/\sigma_{A^S}^2 \) exceeds zero and is sizable. Third, the positive sign of \( \sigma_{A,SA}/\sigma_{A}^2 \) is not consistent with AK’s conclusion that ability and schooling are negatively correlated. This third result implies that the evidence from the Project Talent data that omitted ability bias is not exacerbated in within-twin estimates cannot be used to bolster AK’s interpretation of their results, since they require an unobservable that is negatively associated with schooling. Thus, the evidence is not completely consistent with either AK’s explanation of their findings, or the alternative explanation offered in this paper. However, this evidence is suggestive at best, because there is no reason to believe that the test score in the Project Talent data captures all of the relevant unobservables.

Aside from the question of discriminating among the competing explanations of AK’s results, the Project Talent data provide evidence on the more general point that this paper makes, with the result that \( \sigma_{A^T,SA^T}/\sigma_{A^S}^2 \) and \( \sigma_{A,SA}/\sigma_{A}^2 \) are positive suggesting that within-twin estimates of the return to schooling are likely to be plagued by upward omitted ability bias. As this paper shows, such bias—even if it is reduced relative to cross-section estimates—is exacerbated by IV procedures to correct for measurement error such as those used by AK, as well as by Behrman et al. (1994) and Miller et al. (1995). As Eqs. (10) and (11) show, when upward omitted ability bias persists in within-twin differences, correcting for measurement error in within-twin schooling differences does not necessarily lead to less biased estimates of the return to schooling, and does lead to estimates that overall are upward biased.

### 7. 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 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 (and if twins are identical, we must wonder about the source of schooling differences between them), 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-sectional estimates. That—and not just the role of measurement error in biasing within-twin or within-sibling estimates of the return to schooling downward—was the message of Griliches’ (Griliches, 1979) review of sibling and twin studies.

Ashenfelter and Krueger’s IV approach is intended to correct within-twin estimates of the return to schooling for measurement error bias, and hence to tell us whether lower within-twin than cross-sectional estimates are attributable to the removal of ability bias, or the exacerbation of measurement error bias. However, while the IV estimator eliminates the measurement error bias in...
within-twin estimates, it amplifies any omitted variable bias. If there is upward omitted ability bias in within-twin estimates, the within-twin IV estimator is more upward biased than the standard within-twin estimator of the return to schooling, and may also be more upward biased than the cross-sectional estimator. Given that identical twins are unlikely to be identical in every respect, the within-twin IV estimator may be another example in which the IV “solution” actually worsens the econometric problem (see e.g. Bound et al., 1995).

Acknowledgements

Jere Behrman, Eli Berman, McKinley Blackburn, Elchanan Cohn, and an anonymous referee provided helpful comments. Orley Ashenfelter and Alan Krueger graciously supplied some additional estimates not provided in their paper.

References