Upward Bias in the Estimated Returns to Education: Evidence from South Africa

Article *in* American Economic Review · February 2003

DOI: 10.1257/000282803769206331 · Source: RePEc

1 author:

Thomas Hertz
United States Department of Agriculture

Upward Bias in the Estimated Returns to Education: Evidence from South Africa

By THOMAS HERTZ*

Ordinary least-squares (OLS) estimates of the proportionate increase in wages due to an extra year of education in the United States (the Mincerian rate of return) are believed to be reasonably consistent. It appears that upward bias due to omitted variables is roughly offset by attenuation bias due to errors in the measurement of schooling. Orley Ashenfelter and Cecilia Rouse (1998) find a net upward bias on the order of just 10 percent of the magnitude of the OLS estimate. David Card's (2001) survey of instrumental variables-based estimates reaches a similar conclusion, as do Ashenfelter et al. (1999).

This result need not obtain in all countries at all times.[1] Some (e.g., David Lam and Robert F. Schoeni, 1993) have suggested that omitted variables bias might be larger in less developed economies, where liquidity constraints and family background are likely to be important determinants of both education and earnings. To date, however, there are few estimates of the returns to schooling in developing countries that take account of both omitted variables and measurement error; an exception is Esther Duflo (2001) who finds no net upward bias for Indonesia.[2] In this paper I use the familiar within-family (fixed-effects) approach, as well as variants of the family-effects models described by Ashenfelter and David J. Zimmerman (1997) and Card (1999), to minimize omitted variables bias in a South African data set from 1993. I exploit the fact that about 13 percent of respondents were resurveyed (in 1998) to derive estimates of the reliability of measured schooling, which turns out to be rather low (on the order of 0.77). This suggests that the within-family fixed-effects estimates should be biased downwards to a considerable extent (Zvi Griliches, 1979). However, I also show that errors in the schooling variable are strongly correlated within the family, and that this reduces the degree of attenuation bias in the fixed-effects model. After correcting for these correlated, nonclassical measurement errors, I arrive at schooling coefficients for Africans that are less than half as large as the ordinary least-squares results. My preferred specification yields results on the order of 5 to 6 percent, whereas the initial OLS figures are 11 to 13 percent.

I. Models

A. *Omitted Variables Bias with No Measurement Error*

I first assume that log earnings are linearly related to perfectly measured years of schooling, with random, person-specific slopes and intercepts. Letting y_i stand for log monthly earnings of the ith employed person; α_i for the

* Department of Economics, American University, 4400 Massachusetts Avenue NW, Washington, DC 20016 (e-mail: hertz@american.edu). Many thanks to Orley Ashenfelter, Dale Ballou, Sam Bowles, John Buonaccorsi, Michael Carter, Anne Case, Angus Deaton, Darren Lubotsky, Julian May, Peter Moll, Christina Paxson, and two anonymous but meticulous referees for comments and advice; to the South African Labour and Development Research Unit, Data Research Africa, and the World Bank for the survey data; and to the MacArthur Foundation's Research Network on the Effects of Inequality on Economic Performance, and the Center for Health and Wellbeing at the Woodrow Wilson School of Public and International Affairs, for financial support.

[1] Jere R. Behrman and Mark R. Rosenzweig (1999, p. 167) observe: " 'Ability' bias is not a fundamental underlying parameter and may differ importantly depending on the nature of markets and policies for schooling and for labor." The amount of measurement error may also vary across studies depending on the care with which the questions are written, asked, answered, and transcribed.

[2] Duflo uses an instrumental variables estimator that exploits large policy-induced changes in the availability of education in Indonesia to estimate the return to schooling. Lam and Schoeni (1993) illustrate the effects of measurement errors of various magnitudes but their data from Brazil do not permit them to estimate the error variance. I have found no other studies from developing countries that incorporate corrections for measurement error based on empirical estimates of the reliability of the data or that are able to instrument for measured schooling.

person-specific intercept (which subsumes the usual error term); S_i for the true number of years of schooling completed; and β_i for the person-specific slope, we have:

$$y_i = \alpha_i + S_i \beta_i. \tag{1}$$

If log earnings do in fact depend linearly on schooling alone (up to random error), if schooling is perfectly measured, and if α and β are independent of S, then ordinary least squares will yield an unbiased estimate of $\bar{\beta} \equiv E(\beta_i)$, the mean marginal effect of schooling.[3]

We may then introduce a correlation between α_i, β_i, and S_i to capture the way in which unobserved family background, school quality, individual ability, and other factors will simultaneously influence educational attainment and income at any given level of education. Following Card (1999), this is formalized by taking the linear projections of $\alpha_i - \bar{\alpha}$ and $\beta_i - \bar{\beta}$ onto S_i:

$$\alpha_i = \bar{\alpha} + (S_i - \bar{S})\lambda + u_i \quad \text{and} \tag{2}$$

$$\beta_i = \bar{\beta} + (S_i - \bar{S})\varphi + v_i, \tag{3}$$

where the overbars indicate expectations across all i, and $E(u|S) = E(v|S) = 0$ by construction. Assuming that β and S are jointly symmetrically distributed, Card demonstrates that the asymptotic value of the ordinary least-squares estimate of β from equation (1) is:

$$\text{plim}(\beta^{\text{OLS}}) = \bar{\beta} + \lambda + \bar{S}\varphi. \tag{4}$$

I next consider households with two employed people. Husband-wife, parent-child, and sibling pairs are analyzed separately and these are later pooled with all other multiearner households. Log incomes are given by y_j for $j = 1, 2$, with husbands, parents, and older siblings assigned to the first position. Dropping the household subscript i for clarity, earnings are now given by:

$$y_j = \alpha_j + S_j \beta_j, \text{ with} \tag{5}$$

$$\alpha_j = \bar{\alpha}_j + (S_1 - \bar{S}_1)\lambda_{j1} + (S_2 - \bar{S}_2)\lambda_{j2} + u_j, \text{ and} \tag{6}$$

$$\beta_j = \bar{\beta}_j + (S_1 - \bar{S}_1)\varphi_{j1} + (S_2 - \bar{S}_2)\varphi_{j2} + v_j, \tag{7}$$

where $\bar{\beta}_j$ and $\bar{S}_j$ are expected values across households for all jth members. Combining (5)–(7) and taking the linear projection of the result onto S_1 and S_2 (gathering all constant terms into c_1 and c_2) yields the following:

$$y_1 = c_1 + \overbrace{(\bar{\beta}_1 + \lambda_{11} + \varphi_{11}\bar{S}_1)}^{\tau_{11}} S_1 + \overbrace{(\lambda_{12} + \varphi_{12}\bar{S}_1)}^{\tau_{12}} S_2 + e_1, \tag{8a}$$

$$y_2 = c_2 + \overbrace{(\lambda_{21} + \varphi_{21}\bar{S}_2)}^{\tau_{21}} S_1 + \overbrace{(\bar{\beta}_2 + \lambda_{22} + \varphi_{22}\bar{S}_2)}^{\tau_{22}} S_2 + e_2. \tag{8b}$$

These will be estimated jointly as seemingly unrelated regression equations (S.U.R.E.). Using this model I will first test for the presence of the family effect, by testing the hypothesis that τ_{12} and τ_{21} are zero. This test, and all that follow, will later be repeated in a measurement-error-adjusted model, as described in the next section.

The second question to be addressed is whether returns to education, including any family effects, are symmetric for husbands versus wives, older versus younger siblings, and parents versus children. A necessary and sufficient condition for symmetry is that all of the following are true: $\bar{\beta}_1 = \bar{\beta}_2$, $\bar{S}_1 = \bar{S}_2$, $\lambda_{11} = \lambda_{22}$, $\lambda_{12} = \lambda_{21}$ and likewise for φ. These assumptions are also sufficient, but not necessary, to imply $\tau_{11} = \tau_{22}$ and $\tau_{12} = \tau_{21}$. This latter pair of equalities may be tested empirically; if it is rejected we should be wary of ignoring the information contained in the subscripts—we should not treat fathers and sons, husbands and wives, and older and younger

[3] I interpret the return to education not as an *ex ante* attribute of the individual but as a realized outcome, one that is determined by personal and familial decisions and constraints, by schooling policies and practices, by macroeconomic and labor market conditions and institutions, and by chance.

siblings as interchangeable.[4] However, if the hypothesis is not rejected we cannot be sure that symmetry in fact holds.

There are two sets of identifying restrictions that allow us to generate separate estimates of $\bar{\beta}_1$ and $\bar{\beta}_2$ and to test for their equality. First, we may impose a restriction which I will call vertical uniformity, namely, that $\lambda_{11} = \lambda_{21}$ and $\lambda_{12} = \lambda_{22}$, and similarly for φ. If we also observe that $\bar{S}_1 = \bar{S}_2$ then we may retrieve estimates of both slope coefficients: $\bar{\beta}_1 = \tau_{11} - \tau_{21}$ and $\bar{\beta}_2 = \tau_{22} - \tau_{12}$. We may then test whether these are equal, which amounts to a conditional specification test of the family fixed-effects model, since if $\bar{\beta}_1 = \bar{\beta}_2 = \bar{\beta}$ (and assuming vertical uniformity) we may subtract (8b) from (8a), to arrive at the within-family equation:

$$(9)\quad (y_1 - y_2) = (c_1 - c_2) + \bar{\beta}(S_1 - S_2) + e.$$

This model may be estimated using OLS in the data set of within-family differences,[5] and under these assumptions we have $\text{plim}(\beta_\Delta^{\text{OLS}}) = \bar{\beta}$. Note that if the fixed-effects model is correct, then it again makes no difference whether we assign a given person to the first or second position. However, while we can test for the equality of $\bar{\beta}_1$ and $\bar{\beta}_2$ by assuming vertical uniformity, we cannot test the validity of that identifying restriction in its own right.[6]

If we assume symmetry but not vertical uniformity the fixed-effects estimator converges to:

$$(10)\qquad \text{plim}(\beta_\Delta^{\text{OLS}}) = \bar{\beta} + (\lambda_{11} - \lambda_{12}) + (\varphi_{11} - \varphi_{12})\bar{S}.$$

Although we cannot now identify $\bar{\beta}$, Card (1999) argues that we may bound it from above. If we assume that the correlation between a family member's β (or α) and his own education will be greater than the correlation between his β (or α) and the *other* family member's education, and if we observe that $\text{Var}(S_1) = \text{Var}(S_2)$, then we can show that $\lambda_{11} \geq \lambda_{12}$ and $\varphi_{11} \geq \varphi_{12}$.[7] Equation (10) then implies that $\beta_\Delta^{\text{OLS}}$ is also an upwardly biased estimator of $\bar{\beta}$.[8]

The second identifying restriction is to assume horizontal uniformity (a stronger assumption than Card's horizontal inequality), i.e., that $\lambda_{11} = \lambda_{12}$, $\lambda_{21} = \lambda_{22}$, and likewise for φ. This approach has been used by Ashenfelter and Zimmerman (1997) in the analysis of parent-child pairs.[9] As before, this assumption allows us to estimate $\bar{\beta}_1$ and $\bar{\beta}_2$ and provides an inde-

[4] Strictly speaking, interchangeability does not require this symmetry, since interchangeability also occurs under a different set of assumptions that generates the fixed-effects model; see below.

[5] Subtracting each member from the household mean yields identical parameter estimates to those obtained by subtracting one member's data from the other's, but will understate their standard errors unless one adjusts the degrees of freedom to reflect the inclusion of the household means (the implied fixed-effect terms). The correction factor that will later be applied is $\sqrt{(N - r)/(N - r - f)}$, where f is the number of families (households), r the number of regressors, and N is the number of observations.

[6] Furthermore, if we observe that $\bar{S}_1 \neq \bar{S}_2$, and if ϕ_{11} and ϕ_{22} are nontrivial, then we cannot isolate $\bar{\beta}_1$ and $\bar{\beta}_2$. In particular, we might have $\bar{\beta}_1 \neq \bar{\beta}_2$ even if $\tau_{11} - \tau_{21} = \tau_{22} - \tau_{12}$.

[7] This assumption does not preclude the possibility that parents may compensate by investing more in the schooling of the sibling with lower β (or α), so that they attain more years of schooling than they otherwise might. It does however imply that, on average across households, such compensatory behavior will not be strong enough to cause the sibling with the lower β (or α) to attain more years of schooling than the one with the higher β (or α).

[8] Card's argument can also be applied in the absence of symmetry. In the expression for the probability limit of the fixed-effects estimator, the term corresponding to $(\lambda_{11} - \lambda_{12})$ becomes:

$$\frac{(\lambda_{11} - \lambda_{21})\text{Var}(S_1) + (\lambda_{22} - \lambda_{12})\text{Var}(S_2) - (\lambda_{11} - \lambda_{21} + \lambda_{22} - \lambda_{12})\text{Cov}(S_1, S_2)}{\text{Var}(S_1 - S_2)}.$$

The numerator will be nonnegative if $(\lambda_{11} - \lambda_{21}) \geq 0$ and $(\lambda_{22} - \lambda_{12}) \geq 0$, and if the correlation between S_1 and S_2 does not exceed $\sqrt{\text{Var}(S_1)/\text{Var}(S_2)}$, where $\text{Var}(S_1)$ is assumed to be the smaller of the two variances. The latter requirement may be verified empirically: it is true in each of the subsets I analyze. The former assumption is the vertical equivalent of the horizontal inequality assumed by Card, and seems no less plausible. It again amounts to the claim that own effects are larger than cross effects, and it is a weaker assumption than vertical uniformity (which is $\lambda_{1j} - \lambda_{2j} = 0$.) Under this assumption we may again conclude that the fixed-effects estimator is upwardly biased by the presence of heterogeneity in α.

[9] Ashenfelter and Zimmerman first relax the assumption of vertical uniformity by allowing for proportionality of the following form: $\lambda_{21} = k\lambda_{11}$ and $\lambda_{22} = k\lambda_{12}$ for some constant k. This increases the number of parameters in the model and identification is then restored by the assumption of horizontal uniformity. Note that once horizontal uniformity is assumed any deviation from vertical uniformity is necessarily of this proportional type: thus the only operative restriction is that of horizontal uniformity.

pendent (albeit conditional) test of their equality.

To summarize: I first test for the presence of a family effect, without which none of this would be necessary. I then test for its symmetry, and for the validity of the assumptions underlying the fixed-effects equation (conditional on vertical uniformity). If neither of these tests is rejected then we have no evidence that is inconsistent with the proposition that the mean return to education may be identified by the standard fixed-effects equation. However, if we reject the assumption of vertical uniformity on a priori grounds, we may prefer the weaker claim: that the fixed-effects estimates serve as an upper bound on the true returns to schooling. This claim holds if symmetry holds, and if we accept Card's argument that the own-effect λ_{jj} is larger than the cross effect λ_{jk}; and likewise for φ. Finally, Ashenfelter and Zimmerman's assumption of horizontal uniformity provides a second set of estimates of $\bar{\beta}_1$ and $\bar{\beta}_2$, and again allows us to test for their equality. Because I cannot be sure which identifying restrictions are correct, I adopt each in turn to see what implications they have for our estimates of the returns to education, and the biases therein.

B. *Mean-Reverting Correlated Measurement Errors*

Thus far I have assumed that schooling is perfectly measured, but in fact it is not; moreover, the errors in this variable appear to be nonclassical: they are negatively correlated with the true level of schooling (mean-reverting), and positively correlated within the household. Mean-reversion is to be expected since people with very low levels of schooling cannot underreport by very much (they cannot report less than zero) but can overreport by a large amount. The survey instrument also has a maximum educational category: individuals near this ceiling cannot overreport by much but can underreport. Nor can large positive errors be introduced by interviewers or data-entry personnel when the true value is large; mutatis mutandis when the true value is small. This is most likely to be a problem when errors are both frequent and large in relation to the range of admissible responses, as they appear to be here, and when many people are at an extreme point, as occurs when many have no schooling.[10]

The independence of measurement errors within the family is compromised by the fact that in this survey, as in many others, a single respondent provides information on the educational attainment of all family members: a respondent who is prone to exaggeration might well exaggerate consistently. A given enumerator might also have an idiosyncratic interpretation of the coding scheme, and a given data-entry person might make the same kind of mistakes repeatedly. However, because enumerators and data-entry personnel were not the same for the 1993 survey as for the 1998 resurvey, the latter two channels are unlikely also to have created an *intertemporal* correlation in errors. For about 40 percent of the panel data set, the respondent was the same in both years; yet the interviews were five years apart, and the respondent's frame of mind need not have been constant. I will therefore assume that there is no intertemporal correlation in measurement error, beyond that generated by the correlation between error and level.[11]

Let observed schooling for person $j = 1, 2$ in each survey year be denoted by S_j^{93} and S_j^{98} (the household subscript i is again implicit). True schooling in 1993, the variable of interest, is still simply S_j. Although the 1993 sample is restricted to those who were at least 16 years old and not enrolled in school at the time, some may have returned to school between 1993 and 1998, so true schooling in 1998 will be $S_j + \delta_j$ for some $\delta_j \geq 0$. Those who return to school are sure to have been younger than average in 1993, since virtually no Africans are enrolled beyond age 26. Because national educational attainment has risen over time, they are likely to be better educated than average as well; thus it is probable that S_j and δ_j are positively correlated. In order to work around the complications created by this correlation, I will base my estimates of

[10] Thomas J. Kane et al. (1999) find evidence of mean-reverting errors even in relatively high-quality U.S. survey data. John Bound and Gary Solon (1999) note that instrumental variables estimators that are consistent under classical assumptions are inconsistent under mean-reverting errors.

[11] My results are not affected by allowing for an intertemporal correlation in errors, provided they are relatively small. In the notation developed below, the necessary assumption is that $\mathrm{Cov}(\varepsilon_j^{93}, \varepsilon_j^{98}) - \mathrm{Cov}(\varepsilon_j^{93}, \varepsilon_k^{98}) \cong 0$.

the measurement-error moments on that subset of the panel data set who were not known to be enrolled in school in either year,[12] and who were 27 or older in 1993. In the full labor force, enrollment rates were less than 1 percent for those in this age category, suggesting that reenrollment is rare and hence that $\delta \cong 0$.

Observed schooling in each year will thus be modeled as a mean-reverting linear function of true schooling in 1993. I do not constrain the slopes, intercepts, or error variances to be the same in the two periods. Thus, for $j = 1, 2$, the equations are:

$$S_j^{93} = c^{93} + \gamma S_j + \varepsilon_j^{93} \quad \text{and} \tag{11}$$

$$S_j^{98} = c^{98} + \theta S_j + \varepsilon_j^{98},$$

with $0 < \gamma, \theta < 1$. It will be assumed that $E(\varepsilon_j^t|S_j) = E(\varepsilon_j^t|S_k) = E(\varepsilon_j^{93}\varepsilon_k^{98}) = 0$ for all j, k, and t. I also impose the symmetry requirement that $\text{Cov}(\varepsilon_1^{93}, \varepsilon_2^{93}) = \text{Cov}(\varepsilon_1^{98}, \varepsilon_2^{98})$; later it will be shown that these covariances are positive. Finally, it is observed that the variances of schooling for different family members are comparable; thus I assume:

$$\text{Var}(S_1) = \text{Var}(S_2) \equiv \text{Var}(S),$$

$$\text{Var}(\varepsilon_1^{93}) = \text{Var}(\varepsilon_2^{93}) \equiv \text{Var}(\varepsilon^{93}), \text{ and}$$

$$\text{Var}(\varepsilon_1^{98}) = \text{Var}(\varepsilon_2^{98}) \equiv \text{Var}(\varepsilon^{98}).$$

The assumption that $0 < \gamma, \theta < 1$ implies that true and observed schooling are positively correlated and that the measurement errors are mean-reverting (negatively correlated with S_j).[13]

Now consider the ordinary least-squares regression of log earnings in 1993 on measured schooling in the same year, pooling family members (and so dropping the j subscripts). This yields:

$$\text{plim}(\beta^{\text{OLS}}) = \frac{R}{\gamma}(\bar{\beta} + \lambda + \bar{S}\varphi), \tag{12}$$

$$\text{where} \qquad R \equiv \frac{\gamma^2 \text{Var}(S)}{\gamma^2 \text{Var}(S) + \text{Var}(\varepsilon^{93})} \leq 1.$$

R is the reliability of measured schooling in 1993. If R/γ is less than one, mean-reverting measurement error imparts a downward bias to the OLS estimate of β; otherwise measurement error will bias β^{OLS} upwards, as do the omitted variables.[14]

A problem arises in that estimates of the measurement-error structure that are based on comparing the 1993 and 1998 data will depend on θ as well as γ, whereas the actual measurement errors in 1993, with which we are concerned, depend only on γ. This problem is addressed by comparing two instrumental variables estimates of the returns to schooling: the first uses the 1993 schooling data to instrument for the 1998 results, and the second uses the 1998 data as the instrument. The ratio of these two estimates is a measure of γ/θ, which will be needed below:

$$\beta_{93\to 98}^{IV} = \frac{\text{Cov}(S^{93}, y)}{\text{Cov}(S^{93}, S^{98})} \qquad \text{and}$$

$$\beta_{98\to 93}^{IV} = \frac{\text{Cov}(S^{98}, y)}{\text{Cov}(S^{93}, S^{98})}$$

$$\text{plim}\left(\frac{\beta_{93\to 98}^{IV}}{\beta_{98\to 93}^{IV}}\right) = \text{plim}\,\frac{\text{Cov}(S^{93}, y)}{\text{Cov}(S^{98}, y)} \tag{13}$$

$$= \frac{\gamma\,\text{Var}(S)(\bar{\beta} + \lambda + \bar{S}\varphi)}{\theta\,\text{Var}(S)(\bar{\beta} + \lambda + \bar{S}\varphi)}$$

$$= \frac{\gamma}{\theta}.$$

[12] Enrollment status was not asked of those over 24 years of age in 1993. In 1998 it was asked of all, as part of the process of defining labor force participation.

[13] Note that the measurement errors are $S_j^{93} - S_j = c^{93} + (\gamma - 1)S_j + \varepsilon_j^{93}$, and $S_j^{98} - S_j = c^{98} + (\theta - 1)S_j + \varepsilon_j^{98}$; these are negatively correlated with S_j, as desired, although their purely stochastic components (ε_j^t) are not.

[14] This latter outcome is most likely if the variance of the stochastic error-component is small in relation to the variance of true schooling. In particular:

$$\frac{R}{\gamma} < 1 \qquad \text{iff}$$

$$\frac{1}{2} - \sqrt{\frac{1}{4} - \frac{\text{Var}(\varepsilon^{93})}{\text{Var}(S)}} < \gamma < \frac{1}{2} + \sqrt{\frac{1}{4} - \frac{\text{Var}(\varepsilon^{93})}{\text{Var}(S)}}.$$

As an example, suppose that the ratio of $\text{Var}(\varepsilon^{93})$ to $\text{Var}(S)$ is 1 to 20. Then any value of γ between roughly 0.05 and 0.95 will satisfy this condition.

To introduce measurement error into the S.U.R.E. model I now let $\mathbf{S^{93}}$ represent the matrix $[S_1^{93}, S_2^{93}]$ and $\mathbf{Y}$ the matrix $[y_1, y_2]$. If the covariance structure of the stochastic measurement errors were known, then we could form a measurement-error-corrected estimator as follows:

$$\text{(14)} \quad \boldsymbol{\tau}^{\mathbf{ME}} \equiv \begin{bmatrix} \hat{\tau}_{11} & \hat{\tau}_{12} \\ \hat{\tau}_{21} & \hat{\tau}_{22} \end{bmatrix}$$

$$= \left((\mathbf{S^{93^T}S^{93}}) - N \begin{bmatrix} \text{Var}(\varepsilon^{93}) & \text{Cov}(\varepsilon_1^{93}, \varepsilon_2^{93}) \\ \text{Cov}(\varepsilon_2^{93}, \varepsilon_1^{93}) & \text{Var}(\varepsilon^{93}) \end{bmatrix} \right)^{-1} \times (\mathbf{S^{93^T}Y}), \quad \text{with}$$

$$\text{(15)} \qquad \text{plim}(\boldsymbol{\tau}^{\mathbf{ME}}) = \frac{1}{\gamma} \begin{bmatrix} \tau_{11} & \tau_{12} \\ \tau_{21} & \tau_{22} \end{bmatrix}.$$

The measurement-error variance and covariance terms may be written as follows:

$$\text{(16)} \quad \text{Var}(\varepsilon^{93}) = (1 - R)\text{Var}(S^{93}), \text{ and}$$

$$\text{(17)} \quad \text{Cov}(\varepsilon_1^{93}, \varepsilon_2^{93}) = (1 - \Omega)\text{Cov}(S_1^{93}, S_2^{93}),$$

$$\text{with} \qquad \Omega \equiv \frac{\gamma^2\text{Cov}(S_1, S_2)}{\gamma^2\text{Cov}(S_1, S_2) + \text{Cov}(\varepsilon_1^{93}, \varepsilon_2^{93})}.$$

A method for estimating R and Ω is outlined below; these will then be substituted into equations (16) and (17) to enable us to perform the specification tests with the S.U.R.E. model of equations (14) and (15).

The 1993 reliability R may be estimated by the coefficient from a regression of S^{98} against S^{93}, provided we correct for the ratio γ/θ:

$$\text{(18)} \qquad \hat{R} \equiv \frac{\gamma}{\theta} \left(\frac{\text{Cov}(S^{98}, S^{93})}{\text{Var}(S^{93})} \right)$$

$$= \frac{\gamma}{\theta} \left(\frac{\theta\gamma \text{ Var}(S)}{\gamma^2\text{Var}(S) + \text{Var}(\varepsilon^{93})} \right)$$

$$= \frac{\gamma^2\text{Var}(S)}{\gamma^2\text{Var}(S) + \text{Var}(\varepsilon^{93})}.$$

The corresponding regression in within-family differences, also corrected for γ/θ, yields the following:

$$\text{(19)} \quad \hat{R}_\Delta \equiv \frac{\gamma}{\theta} \left(\frac{\text{Cov}(\Delta S^{93}, \Delta S^{98})}{\text{Var}(\Delta S^{93})} \right) = \frac{R - \Omega\rho_s}{1 - \rho_s}$$

where ρ_s stands for $\text{Corr}(S_1^{93}, S_2^{93})$, the contemporaneous intrafamily correlation in observed schooling.[15] Thus, given estimates of γ/θ, ρ_S, $\hat{R}$, and $\hat{R}_\Delta$ we may use equation (19) to calculate the implied value of Ω.

The term $\hat{R}_\Delta$ also appears in the expression for the attenuation of the within-family fixed-effects estimator. Assuming symmetry, it converges to:

$$\text{(20)} \quad \text{plim}(\beta_\Delta^{\text{OLS}}) = \frac{R_\Delta}{\gamma} [\bar{\beta} + (\lambda_{11} - \lambda_{12}) + (\varphi_{11} - \varphi_{12})\bar{S}],$$

$$\text{with} \qquad R_\Delta \equiv \frac{R - \Omega\rho_s}{(1 - \rho_s)}.$$

Under the classical assumption of independent measurement errors, we have $\text{Cov}(\varepsilon_1^{93}, \varepsilon_2^{93}) = 0$. In this case $\Omega = 1$ and (R_Δ/γ) will be smaller than (R/γ), meaning that the fixed-effects estimator is more attenuated by measurement error than is the cross-sectional result, as emphasized by Griliches (1979). In general, however, this result only holds if $\Omega > R$. The intuition is that if the contemporaneous, within-family covariance in the stochastic component of the measurement errors is large (so that Ω is small) then within-family differencing will actually lessen the influence of measurement errors. Although this extreme case may be rare, it remains true that the attenuation of the within-family estimator is smaller when measurement errors are contemporaneously correlated than when they are independent.

We may test whether errors are in fact correlated in this way by forming the average of the 1993 and 1998 observations for each family member, which serves as a proxy for their true level of schooling. The deviations from this

[15] In estimating the full-sample fixed-effects equation, the pairwise correlation ρ_s will be replaced with Leslie Kish's (1965) intragroup correlation coefficient to allow for households of more than two employed people.

average in a given year (or half the difference between S_j^{93} and S_j^{98}) then contain information about $\mathrm{Cov}(\varepsilon_1^{93}, \varepsilon_2^{93})$. In particular:

$$\text{(21)} \qquad \mathrm{Corr}\left(\frac{S_1^{93} - S_1^{98}}{2}, \frac{S_2^{93} - S_2^{98}}{2}\right) = \frac{2\ \mathrm{Cov}(\varepsilon_1^{93}, \varepsilon_2^{93})}{\mathrm{Var}(\varepsilon^{93}) + \mathrm{Var}(\varepsilon^{98})}.$$

If the observed correlation on the left-hand side of this equation is positive then $\mathrm{Cov}(\varepsilon_1^{93}, \varepsilon_2^{93})$ must be positive as well, implying that $\Omega < 1$.

The measurement-error-corrected fixed-effects estimates for the 1993 subsamples are generated by using the estimates of γ/θ, $\hat{R}$, and Ω obtained from the panel data set, in conjunction with the observed subsample-specific 1993 values of ρ_S, to generate subsample-specific estimates of $\hat{R}_\Delta$ for 1993. These are then incorporated into the estimator as follows:

$$\text{(22)} \qquad \boldsymbol{\beta}_\Delta^{\mathbf{ME}} = (\mathbf{X}_\Delta^T\mathbf{X}_\Delta - \boldsymbol{\Sigma})^{-1}\mathbf{X}_\Delta^T\mathbf{y}_\Delta$$

where $(\mathbf{X}_\Delta, \mathbf{y}_\Delta)$ is the data set of differences from the within-household mean, and $\boldsymbol{\Sigma}$ is the cross-product matrix of measurement errors (also in differences). These cross products are assumed zero for all variables other than schooling (namely age, its square, and gender). For schooling we have:

$$\text{(23)} \qquad \Sigma_{\Delta S} = N(1 - \hat{R}_\Delta)\mathrm{Var}(\Delta S),$$

and (assuming symmetry)

$$\text{(24)} \qquad \mathrm{plim}(\beta_\Delta^{ME}) = \frac{1}{\gamma}[\bar{\beta} + (\lambda_{11} - \lambda_{12}) + (\varphi_{11} - \varphi_{12})\bar{S}].$$

These estimates should be upper bounds on the returns to schooling because the correction factor $\hat{R}_\Delta$ overstates the actual attenuation, which is R_Δ/γ, although Gunnar Isacsson (1999) presents evidence that the degree of overstatement may be small when the data are in differences.[16,17]

Finally, note that to use the coefficient from a regression of within-household schooling differences in one year against the other as an estimate of the reliability of the differenced data is computationally equivalent to treating schooling differences in one year as an instrument for schooling differences in the other.[18] This within-family IV estimator is similar to that used by Ashenfelter and Rouse (1998), except that the second measurement of schooling comes from a follow-up survey, not the other family member. My method extends this approach to the nationally representative data set for which the IV model cannot be estimated because the 1998 data are lacking, under the assumption that the smaller panel data set gives us reliable estimates of the magnitude of measurement error, and allowing for differences in the measurement error across years.

II. Data

The primary data are drawn from the Project for Statistics on Living Standards and Development (PSLSD), a survey of 8,848 households containing 43,974 individuals. The survey was part of the World Bank's series of Living Standards Measurement Surveys and was conducted during the second half of 1993 under the auspices of the Southern Africa Labour and Development Research Unit (SALDRU). The 1998

[16] Note that although the term $\gamma^2\mathrm{Var}(S)$ can be identified, γ and $\mathrm{Var}(S)$ cannot be separated.

[17] Letting $\mathbf{A}_\Delta = (\mathbf{X}_\Delta^T\mathbf{X}_\Delta - \boldsymbol{\Sigma})$ the variance-covariance matrix of the estimator is given by:

$$\mathbf{Cov}(\boldsymbol{\beta}_\Delta^{\mathbf{ME}}) = \left(\frac{\mathbf{y}_\Delta^T\mathbf{y}_\Delta - \boldsymbol{\beta}_\Delta^{\mathbf{ME}}\mathbf{A}_\Delta\boldsymbol{\beta}_\Delta^{\mathbf{ME}T}}{N - f - r}\right)\mathbf{A}_\Delta^{-1}\mathbf{X}_\Delta^T\mathbf{X}_\Delta\mathbf{A}_\Delta^{-1}\mathbf{y}_\Delta,$$

where N, f, and r, are the sample size, the number of families, and the number of regressors. Standard errors for this equation, as well as for the OLS and S.U.R.E. equations, should arguably be adjusted upwards to reflect the correlation of wage-equation residuals among members of the same household, as well as the effects of the clustered sampling strategy. For simplicity, conventional standard errors are reported for all regressions.

[18] More precisely, the within-household IV estimator that uses the 1993 data as the instrument is identical to the measurement-error-corrected within-household model that uses the regression coefficient of schooling differences in 1998 against schooling differences in 1993 as its estimate of R_Δ. The IV estimator that uses the 1998 data as the instrument is replicated by basing R_Δ on the regression of schooling in 1993 against schooling 1998, again in differences. If $\gamma = \theta$ and $\mathrm{Var}(\varepsilon^{93}) = \mathrm{Var}(\varepsilon^{98})$ then these two sets of estimates would also be identical.

TABLE 1—DESCRIPTIVE STATISTICS

Surveyed in:	1993	1993	1993	1993	1993	1993, 1998	1993, 1998	1993, 1998
	Africans 16–65 w/wages in 1993 (1)	Africans 16–65 2+ people w/wages in 1993 (2)	Husbands and wives (3)	Parents and children (4)	Siblings (5)	Africans 16–65 (6)	Africans 27–65 (7)	Africans 27–65 w/wages in 1993 (8)
Sample size	4,773	2,051	878	400	430	2,987	2,052	554
Families	3,642	920	439	200	215	938	887	431
Observations/family	1.3	2.2	2.0	2.0	2.0	3.2	2.3	1.3
Age	37.5	36.6	39.7	39.5	30.5	34.8	40.5	40.2
	(10.7)	(10.8)	(8.9)	(14.4)	(7.5)	(12.3)	(10.7)	(9.5)
Female	0.40	0.47	0.50	0.39	0.41	0.55	0.56	0.45
Years schooling in 1993	6.81	7.07	6.50	6.41	8.03	6.46	5.92	6.71
	(4.13)	(4.23)	(4.40)	(4.37)	(3.77)	(4.09)	(4.09)	(4.14)
S_1	na	na	6.60	4.80	7.70	na	na	na
			(4.38)	(4.11)	(3.90)			
S_2	na	na	6.41	8.02	8.37	na	na	na
			(4.43)	(4.02)	(3.63)			
p-value testing $\text{Var}(S_1) = \text{Var}(S_2)$	na	na	0.83	0.76	0.30	na	na	na
Years schooling in 1998	na	na	na	na	na	6.82	6.12	6.98
						(4.11)	(4.10)	(3.97)
Log wages in 1993	6.36	6.27	6.30	6.19	6.28	na	na	6.34
	(0.93)	(1.00)	(1.08)	(0.94)	(0.92)			(0.94)

Notes: Standard deviations are in parentheses. Log wages are logs of 1993 rands per month; at that time the rand was worth approximately $0.31. Wages include those from regular and casual employment, using only actual (as opposed to imputed) values, net of taxes but including all benefits.

follow-up, known as the KwaZulu-Natal Income Dynamics Study (KIDS), was limited to one province, covering 1,183 households who were PSLSD respondents in 1993 (13 percent of the total). The 1993 survey instrument and the sampling strategy are described in SALDRU (1994); details for the KIDS appear in John Maluccio et al. (1999).

Table 1 describes the primary analysis data set and its subsets, as well as the portion that forms the two-period panel. The first column is the maximal data set for which the relation between wage earnings and schooling in 1993 can be estimated for Africans of working age ($n = 4{,}773$).[19] The second column is limited to those households with two or more employed people with valid wage data ($n = 2{,}051$). The three smaller subsets consist of pairs: husbands and wives, heads of household and their oldest employed child who is living at home, and oldest vs. youngest siblings.[20] For each subset I

[19] The full 1993 sample ($n = 4{,}773$) includes all employed Africans in the PSLSD data set between the ages of 16 and 65 provided they had nonmissing education and gender data, were not known to be enrolled in school in 1993, had reported monthly earnings from regular or casual employment (counting only actual and not imputed earnings values), were not coded as "household help," "lodgers," or "other nonfamily," and did not come from survey clusters 217 and 218, the data for which were discovered to have been at least partially fabricated. Regular wage earnings are net of taxes but include benefits; these were added to reported casual wage earnings.

[20] Siblings were identified as those family members listed as "sons or daughters of the household head" (151 pairs) as well as 64 pairs consisting of a household head and their sister or brother.

TABLE 2—ESTIMATING THE RELIABILITY OF MEASURED SCHOOLING

Surveyed in:	1993, 1998	1993, 1998	1993	1993	1993	1993
	Africans 27–65 (1)	Africans 27–65 wages in 1993 (2)	Africans, 16–65 2+ people w/ wages in 1993 (3)	Husbands and wives (4)	Parents and children (5)	Siblings (6)
(1) $\text{Cov}(S^{98}, S^{93})/\text{Var}(S^{93})$	0.80 (0.013)	0.72 (0.027)	na	na	na	na
(2) $\text{Cov}(\Delta S^{98}, \Delta S^{93})/\text{Var}(\Delta S^{93})$	0.77 (0.015)	0.61 (0.038)	na	na	na	na
(3) Ratio γ/θ	na	1.066 (0.062)†	na	na	na	na
(4) $\hat{R} \equiv \frac{\gamma}{\theta}\text{Cov}(S^{98}, S^{93})/\text{Var}(S^{93})$	na	0.77 (0.046)†	***0.77***	***0.77***	***0.77***	***0.77***
(5) $\hat{R}_{\Delta} \equiv \frac{\gamma}{\theta}\text{Cov}(\Delta S^{98}, \Delta S^{93})/\text{Var}(\Delta S^{93})$	na	0.65 (0.150)†	**0.69**	**0.63**	**0.71**	**0.68**
(6) $\rho_S \equiv \text{Corr}(S_1^{93}, S_2^{93})$	0.42 (0.025)††	0.62 (0.051)††	0.52	0.65	0.43	0.55
(7) Imputed value of Ω	na	**0.85** (0.049)†	***0.85***	***0.85***	***0.85***	***0.85***
(8) $\text{Corr}\left(\frac{S_1^{93}-S_1^{98}}{2}, \frac{S_2^{93}-S_2^{98}}{2}\right)$	0.30 (0.027)††	0.24 (0.091)††	na	na	na	na
Sample size	2,052	554	2,051	878	400	430

Notes: Values in **bold** are imputed using equation (19); values in ***bold italics*** are carried over from the 1993–1998 calculation to the 1993 columns. All other figures are based on estimates described in the text. The final row reports the intrafamily correlation coefficient for deviations of reported schooling from its two-period average. Conventional standard errors appear in parentheses, except:
† Bootstrapped standard error
†† Asymptotic standard error of intragroup correlation coefficient.

test whether the variances of schooling are the same for the different family members; in no case is this hypothesis rejected.

The largest group for which schooling data are available in both years numbers 2,987 [column (6)]. Mean schooling rose by 0.36 years between the two surveys; limiting the sample to those 27 and older, or to those with wages in 1993, reduces this increment somewhat, but it remains significant. Given the low rates of reenrollment in this age group, this gap may signal a difference in the level of measurement error across years, which we have allowed for in our estimating strategy. The final column is the basis for my measurement-error estimates; this subsample is reasonably similar to the primary analysis sample in column (2) in terms of 1993 years of schooling (6.7 versus 7.1), age (40 versus 37), proportion female (0.45 versus 0.47), and 1993 log wages (6.34 versus 6.27).

III. Results

The first two columns of Table 2 present estimates of the reliability of measured schooling for the two age-restricted panel data subsets; the smaller of the two has complete 1993 wage data. The first row reports the coefficient from the regression of schooling in 1998 against schooling in 1993, with no correction yet for the ratio γ/θ. This yields a reliability that is lower than the U.S. survey rule of thumb of 0.90, especially in the sample of employed people (who more nearly resemble the 1993 wage earners we wish to analyze) where it falls to 0.72. In

TABLE 3—OLS, FIXED EFFECTS, AND SIMULTANEOUS EQUATION ESTIMATES OF THE RETURNS TO SCHOOLING WITH AND WITHOUT MEASUREMENT-ERROR CORRECTION
SAMPLE OF 439 HUSBAND-WIFE PAIRS

Model	Statistic estimated	Measurement-error correction: No	Measurement-error correction: Yes
OLS: Pooled	β	0.144 (0.006)	0.193 (0.008)
OLS: Husbands	β_1	0.143 (0.009)	0.188 (0.010)
OLS: Wives	β_2	0.145 (0.009)	0.191 (0.011)
Fixed effects: Pooled	β	0.017 (0.009)	0.037 (0.016)
S.U.R.E.	τ_{11}	0.100 (0.011)	0.132 (0.014)
	τ_{21}	0.090 (0.012)	0.113 (0.016)
	τ_{12}	0.067 (0.010)	0.079 (0.014)
	τ_{22}	0.089 (0.011)	0.113 (0.016)
	z for H_0: $\tau_{21} = 0$	7.78	7.70
	Prob $> z$	0.00	0.00
	z for H_0: $\tau_{12} = 0$	6.39	5.48
	Prob $> z$	0.00	0.00
Symmetry	χ^2 for H_0: $\tau_{11} = \tau_{22}$	0.35	0.55
	Prob $> \chi^2$	0.55	0.46
	χ^2 for H_0: $\tau_{12} = \tau_{21}$	1.62	1.87
	Prob $> \chi^2$	0.20	0.17
	χ^2 for joint test	3.08	3.33
	Prob $> \chi^2$	0.21	0.19
Vertical uniformity	$\beta_1 = \tau_{11} - \tau_{21}$	0.010 (0.010)	0.019 (0.015)
	$\beta_2 = \tau_{22} - \tau_{12}$	0.023 (0.010)	0.034 (0.015)
	χ^2 for H_0: $\beta_1 = \beta_2$	2.16	2.17
	Prob $> \chi^2$	0.14	0.14
Horizontal uniformity	$\beta_1 = \tau_{11} - \tau_{12}$	0.033 (0.019)	0.053 (0.027)
	$\beta_2 = \tau_{22} - \tau_{21}$	−0.001 (0.021)	0.000 (0.030)
	χ^2 for H_0: $\beta_1 = \beta_2$	0.92	1.17
	Prob $> \chi^2$	0.34	0.28

Notes: OLS and fixed-effects regressions predict log monthly earnings as function of own education, age, age squared, and gender. S.U.R.E. system predicts husband's earnings in one equation and wife's in the other, both as functions of both partners' educations, ages, and ages squared. Conventional standard errors appear in parentheses. Measurement-error corrections as described in text using sample-specific reliability estimates taken from Table 2: $\hat{R} = 0.77$, $\Omega = 0.85$, and $\hat{R}_\Delta = 0.63$.

the next row, we see that the basic reliability of the differenced data is lower still, as anticipated, falling to 0.61 in the employed subset.

The next row reports the ratio γ/θ, based on the ratio of IV estimates of the returns to schooling, as described in equation (13), for which wage data are required; the bootstrapped standard error is reported as well. This ratio is slightly above one, although we cannot reject the hypothesis of equality. Multiplying by this ratio raises the effective reliability of the 1993 schooling data in levels from 0.72 to 0.77 (row 4), and in differences from 0.61 to 0.65 (row 5). I will also report the final fixed-effects results on the assumption that γ/θ was one.

Row 6 reports sample-specific estimates of the within-group correlation coefficient for schooling; this varies from 0.42 to 0.65, and this difference has a considerable impact on the degree of attenuation of the within-family estimator. Row 7 of column (2) shows the value of Ω that satisfies equation (19), namely 0.85. Two tests demonstrate that Ω is significantly below one: first, the 95-percent confidence interval based on its bootstrapped standard error extends from 0.75 to 0.94; and second, in the final row we see that the within-family correlation in schooling as a deviation from its two-period average is positive and significant. As noted above, this positive correlation implies $\Omega < 1$.

TABLE 4—OLS, FIXED EFFECTS, AND SIMULTANEOUS EQUATION ESTIMATES OF THE RETURNS TO SCHOOLING WITH AND WITHOUT MEASUREMENT-ERROR CORRECTION SAMPLE OF 200 PARENT-CHILD PAIRS

		Measurement-error correction	
Model	Statistic estimated	No	Yes
OLS: Pooled	β	0.113 (0.010)	0.153 (0.012)
OLS: Parents	β_1	0.115 (0.014)	0.152 (0.017)
OLS: Children	β_2	0.112 (0.015)	0.149 (0.019)
Fixed effects: Pooled	β	0.026 (0.012)	0.045 (0.021)
S.U.R.E.	τ_{11}	0.091 (0.014)	0.122 (0.018)
	τ_{21}	0.071 (0.015)	0.092 (0.019)
	τ_{12}	0.048 (0.015)	0.063 (0.019)
	τ_{22}	0.087 (0.015)	0.117 (0.020)
	z for H_0: $\tau_{12} = 0$	3.30	3.27
	Prob $> z$	0.00	0.00
	z for H_0: $\tau_{21} = 0$	4.74	4.81
	Prob $> z$	0.00	0.00
Symmetry	χ^2 for H_0: $\tau_{11} = \tau_{22}$	0.03	0.03
	Prob $> \chi^2$	0.85	0.87
	χ^2 for H_0: $\tau_{12} = \tau_{21}$	0.95	0.99
	Prob $> \chi^2$	0.33	0.32
	χ^2 for joint test	1.56	1.63
	Prob $> \chi^2$	0.46	0.44
Vertical uniformity	$\beta_1 = \tau_{11} - \tau_{21}$	0.021 (0.014)	0.030 (0.020)
	$\beta_2 = \tau_{22} - \tau_{12}$	0.039 (0.015)	0.054 (0.021)
	χ^2 for H_0: $\beta_1 = \beta_2$	1.17	1.23
	Prob $> \chi^2$	0.28	0.27
Horizontal uniformity	$\beta_1 = \tau_{11} - \tau_{12}$	0.043 (0.024)	0.059 (0.032)
	$\beta_2 = \tau_{22} - \tau_{21}$	0.016 (0.025)	0.025 (0.033)
	χ^2 for H_0: $\beta_1 = \beta_2$	0.39	0.39
	Prob $> \chi^2$	0.53	0.53

Notes: OLS and fixed-effects regressions predict log monthly earnings as function of own education, age, age squared, and gender. S.U.R.E. system predicts parent's earnings in one equation and adult child's earnings in the other, both as functions of both members' educations, ages, ages squared, and genders. Conventional standard errors appear in parentheses. Measurement-error corrections as described in text using sample-specific reliability estimates taken from Table 2: $\hat{R} = 0.77$, $\Omega = 0.85$, and $\hat{R}_\Delta = 0.71$.

We also see that $\Omega > \hat{R}$, which is the criterion for the within-household estimator to be more attenuated than the OLS estimator.

The imputed value of Ω is then extended to the 1993 data sets (for which 1998 schooling data are not available) allowing us to derive a value of $\hat{R}_\Delta$ for each; these range from 0.63 to 0.71. It is worthy of notice that if Ω were assumed to be equal to unity (that is if we ignored the intrahousehold correlation in measurement errors) the estimates of $\hat{R}_\Delta$ would be much lower, ranging from 0.22 to 0.51 (not shown).

Tables 3, 4, and 5 present the OLS, fixed-effects, and S.U.R.E. models, with and without measurement-error corrections, for the set of husband-wife, parent-child, and sibling pairs. For husbands and wives (Table 3) both the pooled and the sex-specific OLS estimates of the schooling coefficient are on the order of 0.14. These rise to about 0.19 after taking account of errors in reported schooling, whose reliability (in levels) is assumed to be 0.77, based on the results in Table 2. The application of family fixed effects reduces the coefficient to 0.017; after correcting for measurement error (with $\hat{R}_\Delta = 0.63$) the estimate rises to 0.037. This figure is roughly one-quarter of the original pooled OLS estimate.

The next panel reports the results of the si-

TABLE 5—OLS, FIXED EFFECTS, AND SIMULTANEOUS EQUATION ESTIMATES OF THE RETURNS TO SCHOOLING WITH AND WITHOUT MEASUREMENT-ERROR CORRECTION
SAMPLE OF 215 SIBLING PAIRS

		Measurement-error correction	
Model	Statistic estimated	No	Yes
OLS: Pooled	β	0.102 (0.010)	0.133 (0.013)
OLS: Oldest siblings	β_1	0.111 (0.013)	0.144 (0.016)
OLS: Youngest siblings	β_2	0.090 (0.016)	0.118 (0.021)
Fixed effects: Pooled	β	0.032 (0.015)	0.049 (0.023)
S.U.R.E.	τ_{11}	0.087 (0.016)	0.119 (0.022)
	τ_{21}	0.051 (0.018)	0.065 (0.026)
	τ_{12}	0.033 (0.016)	0.034 (0.023)
	τ_{22}	0.059 (0.019)	0.075 (0.027)
	z for H_0: $\tau_{21} = 0$	2.84	2.51
	Prob $> z$	0.01	0.01
	z for H_0: $\tau_{12} = 0$	2.01	1.48
	Prob $> z$	0.04	0.14
Symmetry	χ^2 for H_0: $\tau_{11} = \tau_{22}$	1.02	1.19
	Prob $> \chi^2$	0.31	0.28
	χ^2 for H_0: $\tau_{12} = \tau_{21}$	0.41	0.60
	Prob $> \chi^2$	0.52	0.44
	χ^2 for joint test	1.15	1.32
	Prob $> \chi^2$	0.56	0.52
Vertical uniformity	$\beta_1 = \tau_{11} - \tau_{21}$	0.036 (0.016)	0.054 (0.024)
	$\beta_2 = \tau_{22} - \tau_{12}$	0.026 (0.017)	0.041 (0.025)
	χ^2 for H_0: $\beta_1 = \beta_2$	0.42	0.41
	Prob $> \chi^2$	0.52	0.52
Horizontal uniformity	$\beta_1 = \tau_{11} - \tau_{12}$	0.054 (0.028)	0.085 (0.041)
	$\beta_2 = \tau_{22} - \tau_{21}$	0.008 (0.032)	0.010 (0.048)
	χ^2 for H_0: $\beta_1 = \beta_2$	0.74	0.93
	Prob $> \chi^2$	0.39	0.33

Notes: OLS and fixed-effects regressions predict log monthly earnings as function of own education, age, age squared, and gender. S.U.R.E. system predicts oldest siblings' earnings in one equation and youngest siblings' earnings in the other, both as functions of both siblings' educations, ages, ages squared, and genders. Conventional standard errors appear in parentheses. Measurement-error corrections as described in text using sample-specific reliability estimates taken from Table 2: $\hat{R} = 0.77$, $\Omega = 0.85$, and $\hat{R}_\Delta = 0.68$.

multaneous equation model, first with no measurement-error correction, and then with $\hat{R} = 0.77$ and $\Omega = 0.85$. The estimates of the cross effects $\hat{\tau}_{12}$ and $\hat{\tau}_{21}$ are large (between 0.07 and 0.11) and significantly different from zero, confirming that family effects are nontrivial. Below these is the test of the hypothesis of symmetry in own and cross effects, which is not rejected at customary confidence levels. The next panel assumes vertical uniformity, resulting in an estimated schooling coefficient for men of 0.010, rising to 0.019 if measurement error is taken into account; for women the figures are 0.023 and 0.034. The difference between the male and female coefficients is not clearly significant ($p = 0.14$).

Under the assumption of horizontal uniformity, however, the men's coefficient rises to 0.033 (0.053 when corrected for measurement error) while the wives' fall to zero. Note, however, that the standard errors are roughly twice as large as in the previous panel, and thus the hypothesis of equal returns for men and women again cannot be rejected. This is a pattern that will repeat itself: the horizontal estimates are less precise and lead to an implausible divergence in the results for first and second family members.

TABLE 6—OLS, FIXED EFFECTS, AND MEASUREMENT-ERROR-CORRECTED RESULTS IN FULL SAMPLE AND SUBSET OF FAMILIES CONTRIBUTING TWO OR MORE OBSERVATIONS

		Two-worker families			
	Full sample OLS (1)	OLS (2)	Within-household differences (3)	Differenced and measurement-error-corrected $\gamma/\theta = 1.066$ $\hat{R}_\Delta = 0.69$ (4)	Differenced and measurement-error-corrected $\gamma/\theta = 1$ $\hat{R}_\Delta = 0.65$ (5)
Years of education	0.114	0.127	0.030	0.048	0.052
	(0.003)	(0.004)	(0.006)	(0.009)	(0.010)
Age	0.100	0.104	0.063	0.066	0.066
	(0.007)	(0.011)	(0.010)	(0.010)	(0.010)
Age squared	−0.0011	−0.0011	−0.0007	−0.0007	−0.0007
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Female	−0.524	−0.444	−0.472	−0.463	−0.461
	(0.023)	(0.036)	(0.027)	(0.027)	(0.027)
R^2	0.33	0.36	0.25	0.27	0.27
Sample size	4,773	2,051	2,051	2,051	2,051

Notes: Coefficients from regressions of log monthly wage earnings against the variables listed, in levels and differences, with and without measurement-error correction. Conventional standard errors appear in parentheses: no adjustment is made for within-household correlations between errors. The measurement-error correction in column (4) includes an adjustment for the estimated difference in the degree of mean-reversion in schooling measurement errors in 1993 versus 1998, yielding a higher estimated reliability of schooling differences ($\hat{R}_\Delta$). For columns (3) through (5) the reported R^2 measures the share of within-household variance that is explained by the model.

Table 4 produces qualitatively similar results for the parent-child pairs. OLS estimates are on the order of 0.11, while the measurement-error-corrected fixed-effects result (with $\hat{R}_\Delta = 0.71$) is 0.045. The cross terms in the S.U.R.E. model are again large and significant and the hypothesis of symmetry is not rejected. Similarly, neither vertical nor horizontal uniformity leads to a rejection of the equality of returns for parents versus children.

For the sibling pairs (Table 5) the pooled OLS estimates are between 0.09 and 0.11, and the measurement-error-corrected fixed-effects result stands at 0.049 (with $\hat{R}_\Delta$ set at 0.68). Note that in the measurement-error-corrected S.U.R.E. model only one of the two cross terms, $\hat{\tau}_{21}$, is clearly nonzero. Once again, symmetry cannot be rejected, nor can the implications of vertical or horizontal uniformity. Also as before, the estimates based on horizontal uniformity diverge from one another to an implausible degree (0.085 for the older siblings versus 0.010 for the younger), yet equality cannot be rejected due to the poor precision of the estimates.

Table 6 presents my preferred estimates, using all 2,051 people in households of two or more employed people; as argued above, pooling is justified by the failure to reject symmetry in the disaggregated analyses. The OLS estimate for this subset is 0.127, slightly larger than for the full data set that includes single-observation families (0.114). The application of household fixed effects reduces this figure to 0.030, and the measurement-error correction in column (4) raises it back up to 0.048, which is still 62 percent below the OLS estimate. This estimate incorporates the finding that the degree of mean reversion in schooling measurement errors was different in 1998 than in 1993 ($\gamma/\theta > 1$). The last column demonstrates that without this adjustment the final estimate of the returns to schooling would be slightly higher, at 0.052.

IV. Conclusions

I find that the private gain in earnings associated with an extra year of African education in South Africa was on the order of 5 percent, or less than half as large as Mincerian estimates would suggest.[21] Ordinary estimates of the returns to schooling in South Africa would thus appear to suffer from significant bias due to omitted variables. One likely source of this bias in the South African context is the omission of a measure of the quality of education, which has been shown to be not only low but also extremely variable across African schools, and to have a clear impact on labor market outcomes (Anne Case and Angus Deaton, 1999; Case and Motohiro Yogo, 1999).

My results imply that education may be a less powerful tool for raising incomes in South Africa than previously believed.[22] However, these findings, and this interpretation, are subject to a number of important caveats. First, the standard errors reported in Table 6 overstate the precision of the measurement-error-corrected estimates. This is because they are calculated under the assumption that $\hat{R}_{\Delta}$ is a known constant whereas, in fact, this statistic is itself derived from regression estimates of γ/θ, R, ρ_S, and Ω; standard errors for each of these components appear in Table 2, column (2).

Second, provided these point estimates are valid, there are still several reasons to be cautious in basing policy conclusions upon them. For one, as I have argued elsewhere (Hertz, 2001), the bias-corrected Mincerian estimates appear to fall below the true private internal rate of return to investment in education in South Africa, because the actual labor market opportunity costs of schooling are very low at young ages. Moreover, as is well known, policy decisions should be based on the *social* rate of return to schooling, which may differ from the private rate, and which must in turn be judged in relation to alternative social investment strategies. These include investments designed to raise the quality, as opposed to the quantity, of schooling.

[21] This figure rises somewhat, to roughly 6 percent, if the customary measure of labor market experience (age less years of schooling less six) is used in place of age. Age is used instead of experience to prevent errors in reported schooling from entering via this variable.

[22] For an example of a more optimistic conclusion, see Germano Mwabu and T. Paul Schultz (1998). Note also that Mwabu and Schultz (1996) find no evidence of "ability bias." However, their measure of ability is idiosyncratic: they equate it with the residual from the wage equation thus precluding a correlation with schooling.

Finally, these results suggest that continued research into the question of bias in Mincerian estimates from developing countries is warranted, despite the apparent resolution of this debate in the U.S. context. Moreover, researchers should be aware that the reliability of measured schooling in household surveys may sometimes be much lower than the U.S. rule of thumb of 0.9; that the reliability of within-household differences may fall as low as 0.6; and that measurement errors may be strongly correlated within the household, with important consequences for fixed-effects modeling.

REFERENCES

Ashenfelter, Orley; Harmon, Colm and Oosterbeek, Hessel. "A Review of Estimates of the Schooling/Earnings Relationship with Tests for Publication Bias." *Labour Economics*, November 1999, *6*(4), pp. 453–70.

Ashenfelter, Orley and Rouse, Cecilia. "Income, Schooling and Ability: Evidence from a New Sample of Identical Twins." *Quarterly Journal of Economics*, February 1998, *113*(1), pp. 253–84.

Ashenfelter, Orley and Zimmerman, David J. "Estimates of the Returns to Schooling from Sibling Data: Fathers, Sons and Brothers." *Review of Economics and Statistics*, February 1997, *79*(1), pp. 1–9.

Behrman, Jere R. and Rosenzweig, Mark R. " 'Ability' Biases in Schooling Returns and Twins: A Test and New Estimates." *Economics of Education Review*, April 1999, *18*(2), pp. 159–67.

Bound, John and Solon, Gary. "Double Trouble: On the Value of Twins-Based Estimation of the Return to Schooling." *Economics of Education Review*, April 1999, *18*(2), pp. 169–82.

Card, David. "The Causal Effect of Education on Earnings," in Orley Ashenfelter and David Card, eds., *Handbook of labor economics*, Vol 3A. Amsterdam: Elsevier Science Publishers, 1999, pp. 1801–63.

______**.** "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems." *Econometrica*, September 2001, *69*(5), pp. 1121–60.

Case, Anne and Deaton, Angus. "School Inputs and Educational Outcomes in South Africa." *Quarterly Journal of Economics*, August

1999, *114*(3), pp. 1047–84.

Case, Anne and Yogo, Motohiro. "Does School Quality Matter? Returns to Education and the Characteristics of Schools in South Africa." Unpublished manuscript, Department of Economics, Princeton University, September 1999.

Duflo, Esther. "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence From an Unusual Policy Experiment." *American Economic Review*, September 2001, *91*(4), pp. 795–813.

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

Hertz, Thomas. *Education, inequality and economic mobility in South Africa*. Ph.D. dissertation, University of Massachusetts, Amherst, 2001.

Isacsson, Gunnar. *Essays on the twins approach in empirical labor economics*. Ph.D. dissertation, Swedish Institute for Social Research, Stockholm University, 1999.

Kane, Thomas J.; Rouse, Cecilia Elena and Staiger, Douglas. "Estimating Returns to Schooling When Schooling Is Misreported." Working Paper No. 419, Industrial Relations Section, Princeton University, June 1999.

Kish, Leslie. *Survey sampling*. New York: John Wiley, 1965.

Lam, David and Schoeni, Robert F. "Effects of Family Background on Earnings and Returns to Schooling: Evidence From Brazil." *Journal of Political Economy*, August 1993, *101*(4), pp. 710–40.

Maluccio, John; Haddad, Lawrence and May, Julian. "Social Capital and Income Generation in South Africa, 1993–98." Unpublished manuscript, International Food Policy Research Institute (Washington, DC) and Center for Social and Development Studies, University of Natal, (Durban, South Africa), April 1999.

Mwabu, Germano and Schultz, T. Paul. "Education Returns Across Quantiles of the Wage Function: Alternative Explanations for Returns to Education by Race in South Africa." *American Economic Review*, May 1996 (*Papers and Proceedings*), *86*(2), pp. 335–39.

______**.** "Wage Premia for Education and Location, by Gender and Race in South Africa." Discussion Paper No. 785, Yale Economic Growth Center, February 1998.

SALDRU. *South Africans rich and poor: Baseline household statistics*. Cape Town: SALDRU, Project for Statistics on Living Standards and Development, 1994.