The Review *of* Economics *and* Statistics

VOL. LXXIX FEBRUARY 1997 NUMBER 1

ESTIMATES OF THE RETURNS TO SCHOOLING FROM SIBLING DATA: FATHERS, SONS, AND BROTHERS

Orley Ashenfelter and David J. Zimmerman*

Abstract—Data on brothers and on fathers and sons from the National Longitudinal Survey are used to consider the impact of omitted variables and measurement errors on the economic returns to schooling. The analysis suggests that the upward bias in estimated returns due to omitted variables is likely offset by an equal downward bias resulting from measurement errors in reported schooling. Controlling for both of these potential sources of bias yields results comparable to conventional regression estimates of the economic return to schooling.

IN THIS paper we use a representative random sample of data on brothers and on fathers and sons to estimate the economic returns to schooling. The basic idea is to contrast the differences between the schooling of brothers, and of fathers and sons, with the differences in their respective earnings. Since individuals linked by family affiliation are more likely to have similar innate ability and family backgrounds than randomly selected individuals, our procedure provides a straightforward control for unobserved family attributes. Our goal is to determine whether the correlation between earnings and schooling is due, in part, to the correlation between family backgrounds and schooling. Since intrafamily estimates of the return to schooling may be biased downward by measurement error in schooling (see Grilliches (1979)), we also explicitly examine the sensitivity of the results to the presence of measurement error.[1]

Our empirical results indicate that in the sample of brothers the ordinary least-squares estimates of the return to schooling may be biased upward by some 25% by the omission of family background factors. Adjustments for measurement error, however, imply that the intrafamily estimate of the returns to schooling is biased downward by about 25% also, so that the ordinary least-squares estimate suffers from very little overall bias. For contrasts between fathers and sons the empirical results are more complex, as specification tests indicate that simple models of the omitted family background factors are rejected by the data. Estimated returns to schooling for fathers may be biased upward by the omission of family background factors by about 30%, but adjustments for measurement error are of a similar magnitude, so that the ordinary least-squares (OLS) estimate suffers from little overall bias. Estimated returns to schooling for sons, however, are reduced dramatically when family background characteristics are controlled.

Received for publication March 21, 1994. Revision accepted for publication November 22, 1995.

* Princeton University and Williams College, respectively.

We are grateful for helpful comments from David Card, Alan Krueger, Phil Levine, and three anonymous referees.

[1] A large literature exists that attempts to control for omitted variables in estimating the returns to education. (See, for example, Angrist and Krueger (1991), Ashenfelter and Krueger (1994), Behrman et al. (1980), Blackburn and Neumark (1993a, b), Griliches (1977, 1979), and Taubman (1977).) One advantage of the sibling data we use is that they come from a representative national sample. Such a sample is not available, for example, in studies of twin siblings.

I. Empirical Framework

A. Basic Specification

Our analysis begins with the standard relationship (see Mincer (1974)) between the logarithm of hourly wages (Y) and observed schooling (X):

$$Y_{1j} = \beta_1 X_{1j} + F_j + v_{1j} \tag{1}$$

$$Y_{2j} = \beta_2 X_{2j} + F_j + v_{2j} \tag{2}$$

where Y_{ij} and X_{ij} represent the log wage and schooling of the ith brother in the jth family. (In father–son contrasts we write Y_{ij} and X_{ij} for the log wage and schooling of the ith son (or father) in the jth family.)

The error term in each equation is composed of a person-specific component (v_{ij}) and a family-specific component (F_j). The family-specific effect captures unchanging characteristics that are common to all family members. Thus, F varies across families, but is the same for all individuals within a family.

To model the potential correlation between the family effect F and the explanatory variables X we assume that (see Chamberlain (1982))

$$F_j = \lambda_1 X_{1j} + \lambda_2 X_{2j} + \xi_j . \tag{3}$$

The residual term ξ_j is assumed to be uncorrelated with the explanatory variables, and the λ's are parameters that may take different values. Substituting equation (3) into equations (1) and (2) yields the following reduced-form equations:

$$\begin{aligned} Y_{1j} &= (\beta_1 + \lambda_1)X_{1j} + \lambda_2 X_{2j} + w_{1j} \\ &= \Pi_{11}X_{1j} + \Pi_{12}X_{2j} + w_{1j} \end{aligned} \tag{4}$$

$$\begin{aligned} Y_{2j} &= \lambda_1 X_{1j} + (\beta_2 + \lambda_2)X_{2j} + w_{2j} \\ &= \Pi_{21}X_{1j} + \Pi_{22}X_{2j} + w_{2j} \end{aligned} \tag{5}$$

where
$\Pi_{11} = \beta_1 + \lambda_1$
$\Pi_{12} = \lambda_2$
$\Pi_{21} = \lambda_1$
$\Pi_{22} = \beta_2 + \lambda_2$
$w_{ij} = v_{ij} + \xi_{ij}$, which is uncorrelated with X_{ij}.

In this model the schooling of each family member enters into both family members' reduced-form equations. The magnitude of the λ parameters (the coefficient of the siblings' schooling) measures the extent to which estimated returns to schooling are biased due to the omission of family background factors. Equations (4) and (5) comprise a model with correlated random effects.

Least squares provides a simple estimator of the reduced-form equations (4) and (5). As specified, this model is exactly identified and has four structural parameters. Assuming, for example, that the returns to education are the same for both family members ($\beta_1 = \beta_2$) makes this model overidentified. In this case, there would be four reduced-form coefficients with which to identify the three structural parameters. There is a straightforward relationship between the estimates of this model with correlated random effects and the conventional ''fixed-effects'' estimator. For the exactly identified case, the implied structural coefficients found by differencing the estimated reduced-form coefficients, $\pi_{11} - \Pi_{21}$ and $\Pi_{22} - \Pi_{12}$ will be numerically identical to the fixed-effects estimates found by estimated the difference between equations (4) and (5),

$$\begin{aligned} Y_{2j} - Y_{1j} &= (\Pi_{21} - \Pi_{11})X_{1j} + (\Pi_{22} - \Pi_{12})X_{2j} + u \\ &= \Delta_1 X_{1j} + \Delta_2 X_{2j} + w_{1j} - w_{2j}. \end{aligned} \tag{6}$$

That is, $\hat{\Delta}_1 = \hat{\Pi}_{21} - \hat{\Pi}_{11}$ and $\hat{\Delta}_2 = \hat{\Pi}_{22} - \hat{\Pi}_{12}$. Thus, the (unrestricted) reduced-form estimates for the correlated random-effects model will always allow the estimation of the fixed-effects model. This suggests that there is never any harm in fitting the correlated random-effects model when it is not overidentified, and indeed, the correlated random-effects formulation has the benefit of allowing an interpretation of any bias in the OLS estimates that results from ignoring the family effect. For example, estimates of λ_1 and λ_2 should be positive if more ''able'' families obtain more schooling. Since the more general model with different β's (returns to education) is identified, the restriction implied by the commonly estimated fixed-effects model ($\beta_1 = \beta_2$ or, equivalently, $-\Delta_1 = \Delta_2$) is testable. The fixed-effects estimator can therefore be regarded as nested within the (unrestricted) correlated random-effects model. Some of the limitations associated with this empirical framework are discussed below.

Other overidentifying restrictions, such as $\lambda_1 = \lambda_2$, may also be tested. This restriction implies that the sum of the explanatory variables $X_{1j} + X_{2j}$ provides an adequate parameterization of the family effect and may well be appropriate in some applications.[2]

One potential issue is that differences in the family effect across family members could also lead to a rejection of the fixed-effect specification. While the assumption of a common family effect seems plausible for brothers, it may not be true for fathers and sons. For example, the literature on the intergenerational correlation to earnings (see Solon (1992) or Zimmerman (1992)) has suggested a correlation on the order of 0.4 between fathers and sons. This would suggest that the family factor for sons is some fraction of the family factor of the fathers (i.e., αF_j). In this case, the reduced-form equations would be

$$\begin{aligned} Y_{1j} &= (\beta_1 + \alpha\lambda_1)X_{1j} + \alpha\lambda_2 X_{2j} + w_{1j} \\ &= \Pi'_{11}X_{1j} + \Pi'_{12}X_{2j} + w_{1j} \end{aligned} \tag{4$'$}$$

$$\begin{aligned} Y_{2j} &= \lambda_1 X_{1j} + (\beta_2 + \lambda_2)X_{2j} + w_{2j} \\ &= \Pi'_{21}X_{1j} + \Pi'_{22}X_{2j} + w_{2j} \end{aligned} \tag{5$'$}$$

where
$\Pi'_{11} = \beta_1 + \alpha\lambda_1$
$\Pi'_{12} = \alpha\lambda_2$
$\Pi'_{21} = \lambda_1$
$\Pi'_{22} = \beta_2 + \lambda_2$
$w_{ij} = v_{ij} + \xi_{ij}$, which is uncorrelated with X_{ij}.

Such a misspecification would exaggerate the difference in education between fathers and sons, with sons' returns being too small and fathers' returns too large, possibly leading to rejection of the fixed-effects model (i.e., $\Pi'_{11} - \Pi'_{21} < \beta_1$ and $\Pi'_{22} - \Pi'_{12} > \beta_2$). Adding this additional parameter would leave the model underidentified. The model is, however, identified if we incorporate the additional restriction $\lambda_1 = \lambda_2$.

B. *Correlated Random Effects with Measurement Error*

The correlated random-effects model may be easily expanded to allow for the possibility of measurement error in observed schooling. Suppose that both X_1 and X_2 are

[2] For example, $\lambda_1 = \lambda_2$ is a common assumption in studies of twin siblings, as in Behrman et al. (1980) or Ashenfelter and Krueger (1994).

TABLE 1.—CHARACTERISTICS OF BROTHER AND FATHER–SON SAMPLES

	NLS			CPS[a]	
Characteristic	Mean	Standard Deviation		Mean	Standard Deviation
A. Brother Sample					
Brother 1					
Log hourly wage	6.79	0.460		6.67	0.400
Highest grade	13.56	2.72		13.31	2.39
Age	34.17	2.35		33.38	3.10
Brother 2					
Log hourly wage	6.75	0.455			
Highest grade	13.36	2.18			
Age	31.39	1.70			
Brother correlation in schooling			0.51		
Brother correlation in log wages			0.31		
Sample size			143		
B. Father–Son Sample					
Father					
Log hourly wage	6.69	0.507			
Highest grade	10.09	3.87			
Age	50.6	3.94			
Son					
Log hourly wage	6.81	0.419			
Highest grade	14.02	2.45			
Age	33.26	2.67			
Father–son correlation in schooling			0.385		
Father–son correlation in log wages			0.359		
Sample size			332		

[a] CPS means are for full-time working men aged 29–39 in the 1981 outgoing rotation group. Provided for comparison purposes.

measured with error so that

$$X_1 = X_1^* + m_1 \tag{7}$$

$$X_2 = X_2^* + m_2 \tag{8}$$

where X_1^* and X_2^* are the true levels of the explanatory variables and m_1 and m_2 are measurement error terms that are mutually uncorrelated and uncorrelated with the true values of the explanatory variables. Given these assumptions, OLS estimates of equations (4) and (5) would yield inconsistent estimates with[3]

$$\text{plim } \hat{\Pi}_{11} = \Pi_{11} - \frac{\Pi_{11}\psi_1 - \dfrac{\text{cov}(X_1, X_2)}{\text{var}(X_1)}\Pi_{12}\psi_2}{1 - \rho^2} \tag{9}$$

$$\text{plim } \hat{\Pi}_{12} = \Pi_{12} - \frac{\Pi_{12}\psi_2 - \dfrac{\text{cov}(X_1, X_2)}{\text{var}(X_2)}\Pi_{11}\psi_1}{1 - \rho^2} \tag{10}$$

$$\text{plim } \hat{\Pi}_{21} = \Pi_{21} - \frac{\Pi_{21}\psi_1 - \dfrac{\text{cov}(X_1, X_2)}{\text{var}(X_1)}\Pi_{22}\psi_2}{1 - \rho^2} \tag{11}$$

$$\text{plim } \hat{\Pi}_{22} = \Pi_{22} - \frac{\Pi_{22}\psi_2 - \dfrac{\text{cov}(X_1, X_2)}{\text{var}(X_2)}\Pi_{21}\psi_1}{1 - \rho^2} \tag{12}$$

where $\psi_1 = \text{var}(m_1)/\text{var}(X_1)$ and $\psi_2 = \text{var}(m_2)/\text{var}(X_2)$ are the ratios of noise to total variance for X_1 and X_2, respectively, and ρ is the correlation between the observed X_1 and X_4.[4] Bielby et al. (1977), using various repeated measures on schooling, report estimates of ψ of 0.199, 0.162, and 0.079. Siegel and Hodge (1968) report an estimate of 0.0668, and Ashenfelter and Krueger (1992) an estimate of 0.098. Illustrative calculations with formulas (9)–(12) indicate that even small amounts of measurement error may lead to considerable biases in the estimated returns to schooling in the model if the correlation ρ in sibling schooling is large. If we substitute plim $\hat{\Pi}_{ij}$ from (9)–(12) for Π_{ij} in (4) and (5), we obtain the population regression equations (13) and (14)

$$Y_{1j} = \left[\Pi_{11} - \frac{\Pi_{11}\psi_1 - \dfrac{\text{cov}(X_1, X_2)}{\text{var}(X_1)}\Pi_{12}\psi_2}{1 - \rho^2}\right]X_{1j} + \left[\Pi_{12} - \frac{\Pi_{12}\psi_2 - \dfrac{\text{cov}(X_1, X_2)}{\text{var}(X_2)}\Pi_{11}\psi_1}{1 - \rho^2}\right] \times X_{2j} + w_{1j}$$

$$\Downarrow$$

$$Y_{1j} = \left[(\beta_1 + \lambda_1) - \frac{(\beta_1 + \lambda_1)\psi_1 - \dfrac{\text{cov}(X_1, X_2)}{\text{var}(X_1)}\lambda_2\psi_2}{1 - \rho^2}\right]X_{1j} + \left[\lambda_2 - \frac{\lambda_2\psi_2 - \dfrac{\text{cov}(X_1, X_2)}{\text{var}(X_2)}(\beta_1 + \lambda_1)\psi_1}{1 - \rho^2}\right] \times X_{2j} + w_{1j} \tag{13}$$

[3] See Maddala (1992) for a derivation of this result. Maddala uses the normalization Var $(X) = 1$ in his derivation.

[4] Notice that for $\rho = 0$ these expressions reduce to the standard bias formula when only one explanatory variable is measured with error.

$$Y_{2j} = \left[\Pi_{21} - \frac{\Pi_{21}\psi_1 - \dfrac{\text{cov}(X_1, X_2)}{\text{var}(X_1)}\Pi_{22}\psi_2}{1 - \rho^2}\right] X_{1j} + \left[\Pi_{22} - \frac{\Pi_{12}\psi_2 - \dfrac{\text{cov}(X_1, X_2)}{\text{var}(X_2)}\Pi_{21}\psi_1}{1 - \rho^2}\right] \times X_{2j} + w_{2j}$$

$$\Downarrow$$

$$Y_{2j} = \left[\lambda_1 - \frac{\lambda_1\psi_1 - \dfrac{\text{cov}(X_1, X_2)}{\text{var}(X_1)}(\beta_2 + \lambda_2)\psi_2}{1 - \rho^2}\right] X_{1j} + \left[(\beta_2 + \lambda_2) - \frac{(\beta_2 + \lambda_2)\psi_2 - \dfrac{\text{cov}(X_1, X_2)}{\text{var}(X_2)}\lambda_1\psi_1}{1 - \rho^2}\right] \times X_{2j} + w_{2j}. \tag{14}$$

With known values for ρ and $\psi_1 = \psi_2$, equations (13) and (14) are linear in the parameters and may be estimated jointly using a seemingly unrelated regression (SURE) estimator (Zellner (1962)) on the sibling and the father–son data.

II. The Data

The data used in this study are from the National Longitudinal Survey (NLS). The NLS was initiated in 1966 and consisted of four groups, each with approximately 5000 respondents.[5] Several households in the survey yielded more than one respondent. Given household and relationship identifiers, it is possible to match related pairs of individuals. We were able to match 332 father–son pairs and 143 brother pairs for this study. The data used for the sons and brothers are extracted from the 1978 and 1981 cross sections of the NLS.[6] The data used for the fathers are extracted from the 1966 cross section. These dates are selected to capture the brothers at their oldest observation dates and to minimize the difference between the ages of the fathers and the sons. For families yielding more than one father–son or brother match, the eldest son or brother is retained. This preserves independence across observations and attempts to reduce the potential life cycle bias by retaining the son or brother farther out on his earnings life cycle. The analysis uses measures of the average log hourly wage for the brother or son for 1978 and/or 1981 (in cents), along with their ages and years of schooling. Wage rates are converted into 1981 dollars using the consumer price index (CPI). Only fully employed individuals were selected.[7]

Summary statistics for the sample may be found in table 1. The mean age for brother 1 (the elder brother) is 34.17, whereas the mean age for the younger brother is 31.39 years. The highest grade of schooling attained differs only slightly for the two brothers, with brother 1 and brother 2 both possessing on average 13.56 and 13.36 years, respectively. Hourly wages are higher for the elder brother, as would be expected from his greater age and education. The correlation in schooling for the two brothers is 0.51, whereas the correlation in log wages is 0.31.

One possible concern with the NLS brothers data is that they may not be representative of the population of brothers. For example, to be included in the sample, the brother must be living at home in 1966. This could cause an oversampling of men leaving home at a late age. The NLS also oversamples poor neighborhoods. To the extent that the schooling–earnings relationship differs for this group, biased estimates would result. To provide a simple measure of the representativeness of the brothers data, table 1 provides summary statistics for a comparable age range of full-time working men from the Current Population Survey (CPS). Table 2 and the table notes provide log wage regressions for the NLS brothers and from the CPS data. Both the means and the regressions are broadly comparable, though returns to education are somewhat lower in the CPS population whereas the mean and the variance of earnings and education are somewhat higher in the NLS.

Table 1 also contains the summary statistics for the father–son sample. For this sample, the mean age for the fathers in 1966 is 50.6 years, whereas the mean age for the sons in 1981 is 33.26 years. Again, this represents the earliest observation date we could obtain for the fathers and the latest observation date we could obtain for the sons. The difference in observed ages underscores the need to control for life-cycle differences in the reported data. Fathers have considerably less schooling than sons, with the highest grade attained being 10.09 years for the fathers and 14.02 years for the sons. Sons also earn about 12% more per hour than fathers. The correlation in father–son schooling is 0.385, whereas the correlation in log wages is 0.359.

III. Empirical Results

Our empirical results are organized as follows. First we report simple least-squares estimates for the structural

[5] The original cohorts used in this study were men aged 45–59 and young men aged 14–24 in 1966.

[6] Using either 1978 and 1981 or the average produces similar results. Using the average has the advantage of increasing the sample size slightly and raising the precision of the estimates. Data from earlier years capture many workers at the beginning of their working careers and seem less representative of permanent earnings.

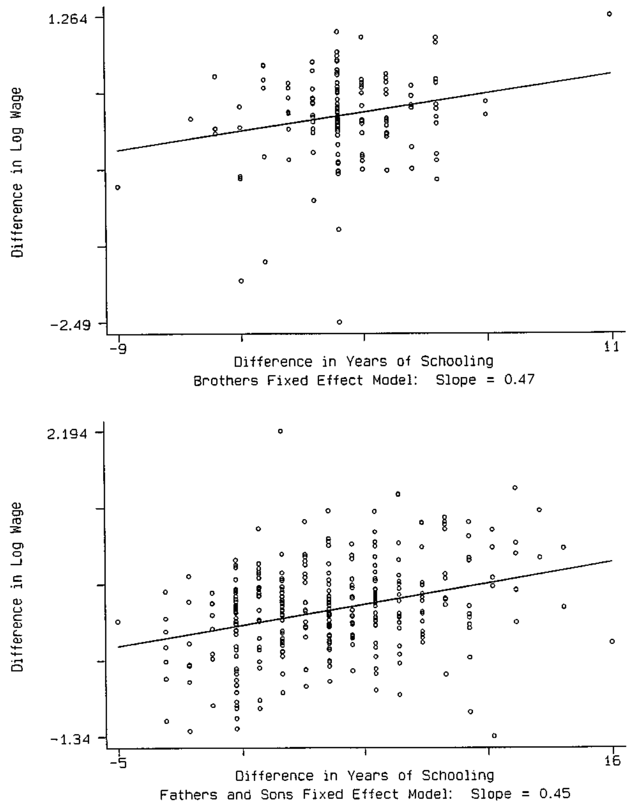
[7] For the purposes of this study, an individual working an average of 30 hours per week, at least 30 weeks per year, and not enrolled in school, was defined as fully employed. Individuals reporting hourly wages of less than one dollar were excluded.

TABLE 2.—SIMPLE CROSS-SECTION AND UNRESTRICTED CROSS-SECTION OLS ESTIMATES

	Log Wage for Years of Schooling			
	Brother 1	Brother 2	Father	Son
Brother 1	0.059 (0.014)	—	—	—
	0.052 (0.015)	0.018 (0.020)	—	—
Brother 2	—	0.071 (0.017)	—	—
	0.006 (0.015)	0.068 (0.019)	—	—
Father	—	—	0.075 (0.006)	—
	—	—	0.065 (0.006)	0.038 (0.010)
Son	—	—	—	0.057 (0.009)
	—	—	0.014 (0.006)	0.049 (0.009)

Note: Log wage regressions also include controls for age and age squared. CPS outgoing rotation group data (1981) for sample comparable to brothers (full-time men aged 29–39) yield a return to education of 0.043 with standard error (0.0015).

FIGURE 1.



equations and the reduced-form equations. These estimates provide the baseline against which to compare other estimates. Next, one brother's (father or son) schooling is used as an instrumental variable for his sibling's (son or father) schooling. This procedure provides a consistent estimator in the presence of measurement error in schooling, but it is not consistent in the presence of an omitted family effect. Finally, we present estimates of the correlated random-effects model with and without measurement error and under a variety of overidentifying restrictions.

A. Basic Estimates: Brothers

Table 2 contains the least-squares estimates of the structural equations (1) and (2) as well as the reduced-form equations (4) and (5) for the sample of brothers. These provide simple estimates of the returns to schooling controlling only for age differences between the two brothers. The results in rows 1 and 3 indicate the returns to schooling to be 5.9% and 7.1% for brothers 1 and 2, respectively.[8] Rows 2 and 4 present the reduced-form estimates. In this specification a brother's wage depends on both his own education and that of his brother. As shown in equations (3)–(5), the coefficient of the brother's sibling in the log wage equation provides a measure of the parameter λ_1 and λ_2 from equation (3). As anticipated, these coefficients are positive. They are, however, small in magnitude ($\lambda_1 = 0.018$ and $\lambda_2 = 0.006$) and statistically insignificant at conventional levels (see column 1 in table 4). This suggests that estimated returns for brothers are only slightly upward biased due to omitted family background factors. Figure 1 presents this basic result using a scatter diagram of the intrapair log wage differentials against the intrapair schooling differences. The return to an additional year of education, assuming equal returns for the brothers (this is the fixed-effects estimate found in table 4, column 2), is calculated to be 4.7%. Of course, this estimate could be biased downward if there is a measurement error in reported schooling.

TABLE 3.—INSTRUMENTAL-VARIABLES ESTIMATES

	Log Wage for Years of Schooling			
	Brother 1	Brother 2	Father	Son
Brother 1	0.080 (0.027)	IV	—	—
Brother 2	IV	0.083 (0.034)	—	—
Father	—	—	0.127 (0.017)	IV
Son	—	—	IV	0.109 (0.025)

Note: Log wage regressions also include controls for age and age squared. IV indicates the variable used to instrument for own schooling in regression.

Table 3 reports the instrumental-variables estimate of the return to schooling by using the education of each brother as an instrument for the education of his sibling. As noted above, if the measurement error in the brothers' schooling is uncorrelated with the true level of schooling, and uncorrelated across brothers, then this instrumental variable would

[8] Interestingly, these estimates are similar to those calculated by Chamberlain and Griliches (Taubman, 1977) using the NLS brothers. Chamberlain and Griliches used the sample when most of the brothers were still in school and relied on "expected" occupation to develop their earnings variable and "expected total schooling" for their measure of schooling. Despite these simplifications, their results (estimated returns to schooling of 5.7%) are remarkably similar to those ultimately attained by the brothers.

TABLE 4.—SEEMINGLY UNRELATED REGRESSION ESTIMATES: NO MEASUREMENT ERROR—BROTHER SAMPLE

	Model			
	Unrestricted (1)	$\beta_1 = \beta_2$ FE (2)	$\lambda_1 = \lambda_2$ (3)	$\beta_1 = \beta_2$, $\lambda_1 = \lambda_2$ (4)
β_1, brother 1	0.046	—	0.045	—
	(0.019)		(0.019)	
β_2, brother 2	0.050	—	0.054	—
	(0.024)		(0.023)	
$\beta_1 = \beta_2 = \beta$	—	0.047	—	0.048
		(0.018)		(0.018)
λ_1, brother 1	0.006	0.005	—	—
	(0.015)	(0.015)		
λ_2, brother 2	0.018	0.019	—	—
	(0.020)	(0.017)		
$\lambda_1 = \lambda_2 = \lambda$	—	—	0.010	0.011
			(0.011)	(0.011)
N^* objective	275.98	276.02	276.16	276.34
Prob $> \chi^2$	—	0.8415	0.6714	0.8353

Note: Models also include controls for age and age squared.

TABLE 5.—SEEMINGLY UNRELATED REGRESSION ESTIMATES: MEASUREMENT ERROR $\psi = 0.199$—BROTHER SAMPLE

	Model			
	Unrestricted (1)	$\beta_1 = \beta_2$ FE (2)	$\lambda_1 = \lambda_2$ (3)	$\beta_1 = \beta_2$, $\lambda_1 = \lambda_2$ (4)
β_1, brother 1	0.076	—	0.076	—
	(0.030)		(0.030)	
β_2, brother 2	0.087	—	0.089	—
	(0.038)		(0.037)	
$\beta_1 = \beta_2 = \beta$	—	0.077	—	0.078
		(0.030)		(0.030)
λ_1, brother 1	−0.006	−0.007	—	—
	(0.023)	(0.024)		
λ_2, brother 2	0.007	0.012	—	—
	(0.031)	(0.029)		
$\lambda_1 = \lambda_2 = \lambda$	—	—	−0.0007	0.001
			(0.017)	(0.017)

Note: Models also include controls for age and age squared.

TABLE 6.—SEEMINGLY UNRELATED REGRESSION ESTIMATES: MEASUREMENT ERROR $\psi = 0.0668$—BROTHER SAMPLE

	Model			
	Unrestricted (1)	$\beta_1 = \beta_2$ FE (2)	$\lambda_1 = \lambda_2$ (3)	$\beta_1 = \beta_2$, $\lambda_1 = \lambda_2$ (4)
β_1, brother 1	0.052	—	0.052	—
	(0.022)		(0.022)	
β_2, brother 2	0.057	—	0.061	—
	(0.027)		(0.026)	
$\beta_1 = \beta_2 = \beta$	—	0.054	—	0.054
		(0.021)		(0.021)
λ_1, brother 1	0.004	0.003	—	—
	(0.017)	(0.017)		
λ_2, brother 2	0.017	0.018	—	—
	(0.022)	(0.020)		
$\lambda_1 = \lambda_2 = \lambda$	—	—	0.009	0.009
			(0.013)	(0.013)

Note: Models also include controls for age and age squared.

provide a consistent estimate of the returns to schooling if there were no omitted family effect bias. The estimates for the instrumental-variables estimator rise to 8% for brother 1 and to 8.3% for brother 2.

These results suggest that estimated returns to education do suffer from a moderate upward omitted-variable bias.

B. *Correlated Random Effects and Measurement Error: Brothers*

Table 4 presents the estimates of the returns to schooling using the correlated random-effects framework. Column 1 gives the estimates for the unrestricted model. Returns to schooling are calculated to be 4.6% for brother 1 and 5% for brother 2. These estimates are the same (except for rounding error) as those implied by table 4. Column 2 presents the estimates restricting the returns to schooling to be the same for both of the brothers. This is the restriction implied by the standard fixed-effects model. The estimated (common) return to schooling is 4.7%. Again, λ_1 and λ_2 are insignificant. The chi-squared statistic for the joint restriction $\beta_1 = \beta_2$ has a *p*-value of 0.84. Thus, the fixed-effects specification cannot be rejected. Column 3 restricts the λ's to be the same. This hypothesis also cannot be rejected. Finally, column 4 restricts both the returns to schooling to be the same for both brothers and the λ's to be the same. It is not possible to reject this restriction, and the resulting estimates closely resemble those of the fixed-effects estimator.

Tables 5 and 6 reestimate the correlated random-effects specification, assuming different magnitudes for the measurement error in reported schooling. We provide estimates using the largest ratio of estimated measurement error to total variance in schooling, which was reported by Bielby et al. (1977) at $\psi = 0.199$, and the smallest estimate reported by Siegel and Hodge (1968) at $\psi = 0.0668$. We assume that measurement error variances are the same for both siblings. As expected, the downward bias in estimated returns to schooling is related positively to the magnitude of the measurement error. (See table 10 for a summary of the returns to schooling for a range of assumptions about the magnitude of the measurement error.) If there is no measurement error, estimated returns are 4.6% for brother 1 and 5% for brother 2. If the measurement error is as much as 20% of the schooling variance, returns are 7.4% for brother 1 and 8.6% for brother 2. Interestingly, comparing these estimates to the instrumental-variables estimates reported earlier suggests that a ratio of noise to total variance of 0.2 seems consistent with the NLS data. This comparison is plausible, given the evidence that, for brothers, the estimated returns are not biased downward due to omitted family characteristics.

The fixed-effects estimator indicates returns to schooling to be 7.7% when $\psi = 0.199$, compared to the estimate of 4.6% under the assumption of no measurement error.

These estimates of the returns to schooling imply that the upward bias in estimated returns to schooling due to omitted family background factors is fairly small and may be smaller than the downward bias in estimated returns due to measurement error.

C. *Basic Estimates: Fathers and Sons*

Table 2 contains the least-squares estimates of the structural equations (1) and (2) as well as the reduced-form equations (4) and (5) for the sample of fathers and sons. The results in rows 3 and 4 indicate the returns to schooling to be 7.5% for the father and 5.7% for the son. The reduced-form estimates are found in rows 6 and 8. For fathers and sons the estimated λ_1 and λ_2 parameters are both statistically significant (see column 1 in table 7) and large. Indeed, the implied structural estimates of the returns to schooling drop from 5.7% to 1.1% for the son. The estimate for the father drops from 7.5% to 5.2%. A scatter diagram of the intrapair log wage differentials against the intrapair schooling differentials is found in figure 1. Here it is apparent that in a comparison of two sons, both of whom have equally educated fathers, the son who is better educated has the higher earnings. The slope of the least-squares line drawn through these data (the fixed-effects estimate; see column 2 in table 7) indicates that an additional year of schooling results in a 4.5% increase in earnings.

Table 3 investigates the effect of measurement error on returns to schooling by using the education of the father (son) as an instrument for the education of his son (father). The instrumental-variables estimates of the returns to schooling are calculated to be 12.7% for the father and 10.9% for the son. These estimates must be regarded with caution, however, as they assume the absence of any omitted-variable bias due to family background effects.

These estimates suggest that when using matched father–son data, estimated returns may be biased upward by omitted family factors. It is also possible, however, that the assumption of a common family effect is less appropriate in the father–son data. Brothers, for example, would typically be exposed to a similar family environment, whereas fathers and sons do not grow up in the same household.

D. *Correlated Random Effects and Measurement Error: Fathers and Sons*

Table 7 presents the results for the correlated random-effects model using the father–son data. Column 1 simply separates the structural coefficients implied by the reduced-form estimates found in table 2. As noted above, the λ terms are both positive and significant, indicating an omitted-variable bias. The fixed-effect estimator, which restricts returns to fathers' and sons' schooling to be the same, is presented in column 2. The (common) estimated return is 4.5%, which is very similar to that found for the brothers. It is, however, possible to reject the fixed-effect restriction for the father and son data. It is also possible to reject the equality of the λ terms.

As noted above, the rejection of the fixed-effects specification could be the result of the family effect varying between fathers and sons. We explore this possibility by allowing the factor loading on the family effect to differ between fathers and sons. To identify the model, we restrict the λ terms to be the same. This reformulation of the model results in the estimated returns to education being much closer for the fathers and sons. The estimated coefficient linking family background for the sons and fathers is statistically significant and is estimated to be 0.351, very similar to that found in the literature on intergenerational earnings correlations.

Tables 8 and 9 reestimate the correlated random-effects specification using the estimated measurement error in reported schooling suggested by Bielby et al. (1977) and Siegel and Hodge (1968). Table 10 summarizes the estimated returns to schooling for a variety of measurement errors. It may be seen that the sons' returns to schooling

TABLE 8.—SEEMINGLY UNRELATED REGRESSION ESTIMATES: MEASUREMENT ERROR $\psi = 0.199$—FATHER–SON SAMPLE

	Model				
	Unrestricted (1)	$\beta_1 = \beta_2$ FE (2)	$\lambda_1 = \lambda_2$ (3)	$\beta_1 = \beta_2$, $\lambda_1 = \lambda_2$ (4)	α free, $\lambda_1 = \lambda_2$ (5)
β_1, son	0.027 (0.017)	—	0.036 (0.016)	—	0.049 (0.019)
β_2, father	0.069 (0.010)	—	0.070 (0.011)	—	0.047 (0.023)
$\beta_1 = \beta_2 = \beta$	—	0.067 (0.011)	—	0.067 (0.011)	—
λ_1, son	0.035 (0.014)	0.015 (0.012)	—	—	—
λ_2, father	0.013 (0.008)	0.014 (0.008)	—	—	—
$\lambda_1 = \lambda_2 = \lambda$	—	—	0.019 (0.007)	0.014 (0.006)	0.035 (0.014)
α	—	—	—	—	0.371 (0.290)

Note: Models also include controls for age and age squared.

TABLE 7.—SEEMINGLY UNRELATED REGRESSION ESTIMATES: NO MEASUREMENT ERROR—FATHER–SON SAMPLE

	Model				
	Unrestricted (1)	$\beta_1 = \beta_2$ FE (2)	$\lambda_1 = \lambda_2$ (3)	$\beta_1 = \beta_2$, $\lambda_1 = \lambda_2$ (4)	α free, $\lambda_1 = \lambda_2$ (5)
β_1, son	0.011 (0.012)	—	0.022 (0.011)	—	0.036 (0.013)
β_2, father	0.052 (0.008)	—	0.051 (0.008)	—	0.027 (0.014)
$\beta_1 = \beta_2 = \beta$	—	0.045 (0.008)	—	0.045 (0.008)	—
λ_1, son	0.038 (0.010)	0.020 (0.009)	—	—	—
λ_2, father	0.014 (0.006)	0.017 (0.006)	—	—	—
$\lambda_1 = \lambda_2 = \lambda$	—	—	0.021 (0.005)	0.018 (0.005)	0.038 (0.010)
α	—	—	—	—	0.351 (0.184)
N* objective	653.89	665.76	657.77	665.85	654.00
Prob > χ^2	—	0.0006	0.0489	0.00256	—

Note: Models also include controls for age and age squared.

TABLE 9.—SEEMINGLY UNRELATED REGRESSION ESTIMATES: MEASUREMENT ERROR $\psi = 0.0668$—FATHER–SON SAMPLE

	Model				
	Unrestricted (1)	$\beta_1 = \beta_2$ FE (2)	$\lambda_1 = \lambda_2$ (3)	$\beta_1 = \beta_2$, $\lambda_1 = \lambda_2$ (4)	α free, $\lambda_1 = \lambda_2$ (5)
β_1, son	0.015 (0.013)	—	0.026 (0.012)	—	0.419 (0.015)
β_2, father	0.058 (0.009)	—	0.057 (0.008)	—	0.032 (0.015)
$\beta_1 = \beta_2 = \beta$	—	0.051 (0.009)	—	0.051 (0.009)	—
λ_1, son	0.038 (0.011)	0.019 (0.009)	—	—	—
λ_2, father	0.012 (0.007)	0.015 (0.007)	—	—	—
$\lambda_1 = \lambda_2 = \lambda$	—	—	0.020 (0.005)	0.016 (0.005)	0.038 (0.011)
α	—	—	—	—	0.312 (0.202)

Note: Models also include controls for age and age squared.

TABLE 10.—SURE ESTIMATES OF RETURNS TO SCHOOLING FOR VARIOUS VALUES OF ψ

	ψ	β_1	β_2
Brother sample	0.00	0.046	0.050
	0.05	0.051	0.056
	0.10	0.058	0.064
	0.15	0.065	0.073
	0.20	0.074	0.086
	0.25	0.087	0.104
	0.30	0.104	0.130
Father–son sample ($\alpha = 1$)	0.00	0.011	0.052
	0.05	0.013	0.055
	0.10	0.017	0.059
	0.15	0.022	0.064
	0.20	0.027	0.070
	0.25	0.032	0.077
	0.30	0.045	0.086
	0.35	0.060	0.098
	0.40	0.083	0.115
	0.45	0.120	0.142
Father–son sample ($\lambda_1 = \lambda_2$, α free)	0.00	0.036	0.027
	0.05	0.039	0.030
	0.10	0.042	0.035
	0.15	0.045	0.040
	0.20	0.050	0.047
	0.25	0.055	0.057
	0.30	0.062	0.069
	0.35	0.071	0.087
	0.40	0.085	0.113
	0.45	0.106	0.157

remain low, even for relatively high levels of measurement error, when a common family effect is assumed. Allowing the family effect to differ between fathers and sons increases the sons' returns to schooling and reduces the differences in the fathers' and sons' returns.

E. *Caveats*

When interpreting the above results, two principal caveats must be born in mind. First, while we explicitly control for family-specific factors, we do not control for individual-specific factors. We cannot use individual fixed effects since education does not vary over time in our sample. Second, we do not allow for the possibility that the education variable is endogenous. Both of these difficulties can be illustrated by considering the simple fixed-effects model without measurement error, which is equation (6) with $-\Delta_1 = \Delta_2$,

$$Y_{2j} - Y_{1j} = \Delta(X_{1j} - X_{2j}) + (w_{1j} - w_{2j}). \tag{15}$$

The possibility that there are individual components of ability not eliminated in differencing the two wage equations would imply a correlation between $(X_{1j} - X_{2j})$ and $(w_{1j} - w_{2j})$. As Griliches (1979) notes, if this is true, differencing may actually result in greater bias than the conventional OLS estimator. The sign of this correlation depends on the structural model employed. Typically, it is assumed that the correlation is positive, with the more ''able'' sibling acquiring more education. If, however, education is endogenous, this need not be the case. For example, Behrman et al. (1982) developed a model incorporating parental preferences in how educational resources are distributed among children within a family. Families may choose to compensate for differences in individual-specific abilities by allocating more resources to the child with lower ability.[9] This could impart a negative correlation between $(X_{1j} - X_{2j})$ and $(w_{1j} - w_{2j})$. Alternatively, they could reinforce the differences by allocating extra resources to enhance the development of the gifted child. In this case the correlation could be positive. Notice that such a correlation also causes inconsistent estimates to be produced when one sibling's education is used as an instrument for the others since each sibling's educational attainment is correlated with the residual in the other sibling's wage equation via the parental reallocation rule.[10]

IV. Conclusions

In this paper we have used matched pairs of brothers and of fathers and sons from the National Longitudinal Survey to estimate the economic returns to schooling. Our empirical findings are strongest when using data on brothers. The

[9] The empirical results in Behrman et al. (1982) support the notion that parents follow a compensating strategy.

[10] We have attempted simple controls for these two difficulties. First, we have attempted to incorporate crude measures of ability into the estimating equations using IQ scores from the NLS. Unfortunately, missing observations cause the sample to be reduced to the point where estimates are highly imprecise. Second, it is very difficult to control for the possible endogeneity arising from parents' choices concerning the intrahousehold allocation of resources. An instrumental-variables estimation scheme would require instruments differing for each child within the household. Following Griliches (1979), we hypothesized that any parental compensatory or reinforcement strategies would be strongest for children closest in age. Hence, selecting sibling pairs farther apart in age should reduce the potential endogeneity. Unfortunately, we again were confronted with a sample so diminished as to be unhelpful. As such, we simply indicate that these caveats must be remembered in considering the results.

evidence suggests that any upward bias in estimated returns to schooling due to omitted family background factors is no larger than the downward bias due to errors in the measurement of schooling. Using data on fathers and sons introduces some ambiguity into these findings, as commonly used specification tests reject our simplest models of the role of family background in the determination of earnings. It seems likely that this is partly due to different family factors for the father and son, suggesting that a more complex model of the father–son relationship may be necessary for analyzing these data.

As is well known, the return to schooling has increased substantially in the past decade, so that some care must be taken in the comparison of our results with analyses of other time periods.[11] Our estimates of the returns to schooling are generally comparable to other estimates in the literature for the time period (1981) for which we measure sibling wage rates.[12]

It seems that additional studies of sibling data could provide a useful source of information on the economic returns to schooling. In principle the necessary data should be available to study the returns to schooling for different groups at different time periods and in different locations.

[11] See, for example, Katz and Murphy (1992) or Boozer et al. (1992), and, especially for the National Longitudinal Survey data, Blackburn and Neumark (1993a, b).

[12] For example, our results are similar to the finding by Behrman et al. (1980) that the intrafamily return to schooling for fraternal twins is about 25% lower than the estimate that does not control for unobserved family background differences. Fraternal twins of the same sex bear the same genetic relationship as do brothers or sisters.

REFERENCES

Angrist, Joshua, and Alan Krueger, ‘‘Does Compulsory School Attendance Affect Schooling and Earnings?,’’ *Quarterly Journal of Economics* (Nov. 1991).

Ashenfelter, Orley, and Alan Krueger, ‘‘Estimates of the Economic Returns to Schooling from a New Sample of Twins,’’ *American Economic Review* (Dec. 1994).

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

Behrman, Jere, Robert A. Pollak, and Paul Taubman, ‘‘Parental Preferences and Provision of Progeny,’’ *Journal of Political Economy* (Feb. 1982).

Bielby, William, Robert Hauser, and David Featherman, ‘‘Response Errors of Non-Black Males in Models of the Stratification Process,’’ in D. J. Aigner and A. S. Goldberger (eds.), *Latent Variables in Socio-Economic Models* (Amsterdam: North-Holland, 1977), 227–251.

Blackburn, McKinley, and David Neumark, ‘‘Omitted-Ability Bias and the Increase in the Return to Schooling,’’ *Journal of Labor Economics* 11 (July 1993a), 521–544.

——— ‘‘Are OLS Estimates of the Return to Schooling Biased Downward? Another Look,’’ National Bureau of Economic Research, Working Paper 4259 (Jan. 1993b).

Boozer, Michael, Alan Krueger, and Shari Wolkon, ‘‘Race and School Quality Since Brown v. Board of Education,’’ *Brookings Papers: Microeconomics* (1992), 269–338.

Chamberlain, Gary, ‘‘Multivariate Regression Models for Panel Data,’’ *Journal of Econometrics* 18:5 (1982), 5–46.

——— ‘‘Estimating the Returns to Schooling—Some Econometric Problems,’’ *Econometrica* 45 (Jan. 1977), 1–22.

——— ‘‘Sibling Models and Data in Economics: Beginnings of a Survey,’’ *Journal of Political Economy* 87(Supplement), (1979), S37–S64.

Katz, Lawrence, and Kevin Murphy, ‘‘Changes in Relative Wages, 1963–1987: Supply and Demand Factors,’’ *Quarterly Journal of Economics* 107 (Feb. 1992), 35–78.

Maddala, G. S., *Introduction to Econometrics*, 2nd ed. (New York: Macmillan, 1992), 456–459.

Mincer, Jacob, *Schooling, Experience, and Earnings* (New York: National Bureau of Economic Research, Columbia University Press, 1974).

Siegel, Paul, and Robert Hodge, ‘‘A Causal Approach to the Study of Measurement Error,’’ in Hubert Blalock and Ann Blalock (eds.), *Methodology in Social Research* (New York: McGraw-Hill, 1968), 28–59.

Solon, G., ‘‘Intergenerational Income Mobility in the United States,’’ *American Economic Review* (June 1992).

Taubman, Paul (ed.), *Kinometrics: Determinents of Socioeconomic Success within and between Families* (Amsterdam: North-Holland, 1977).

Zellner, Arnold, ‘‘An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias,’’ *Journal of the American Statistical Association* 57 (1962), 348–368.

Zimmerman, David J., ‘‘Regression Towards Mediocrity in Economic Stature,’’ *American Economic Review* (June 1992).