Provided by the author(s) and University College Dublin Library in accordance with publisher policies. Please cite the published version when available.

| Title | Estimates of the economic return to schooling for the United Kingdom |
| :--- | :--- |
| Authors(s) | Walker, lan; Harmon, Colm |
| Publication date | $1995-12$ |
| Publication information | American Economic Review, 85 (5): 1278-1286 |
| Publisher | American Economic Association |
| Link to online version | http://search.ebscohost.com/login.aspx?direct=true\&db=eoh\&AN=0375629\&site=ehost-live |
| Item record/more information | http://hdl.handle.net/10197/647 |

# Estimates of the Economic Return to Schooling for the United Kingdom 

By Colm Harmon and Ian Walker*

The rate of return to schooling is an important factor in determining educational attainment and participation and, ultimately, wages and incomes. There are a variety of sources of bias associated with ordinary least-squares (OLS) estimates of the return to schooling. First, in an optimizing investment model we would expect a positive correlation between schooling and its return, which would imply that OLS estimates of the rate of return would be biased upward. Secondly, "ability bias" arises in the presence of an unobservable factor which is correlated with schooling and yet also correlated with wages, and this results in upward bias in OLS estimates. A third source of potential bias is measurement error, and the conventional wisdom suggests that in this case the bias in OLS is downward (see Zvi Griliches [1977] and McKinley L. Blackburn and David Neumark [1995] for further discussion). Recent research by Kevin Lang (1993) and David Card (1994) suggests that

[^0]OLS estimates may be subject to discount-rate bias arising from individuals with higher discount rates choosing less education in an optimizing model. OLS provides an estimate of the rate of return to education on average while instrumental variables (IV) provides an estimate of the rate of return for marginal individuals with high discount rates. Hence we expect the OLS estimates to be lower than IV estimates.

Previous work has dealt with the endogeneity issue using one of four methods. First, one could include an explicit proxy for ability, for example, IQ tests (Griliches and William M. Mason, 1972; Griliches, 1977). Recent results using this approach suggest an upward bias in least-squares estimates (see Blackburn and Neumark, 1993). A second approach uses twins (or siblings) to attempt to eliminate endogeneity bias by exploiting the differences between twins in levels of schooling and earnings, on the grounds that this eliminates differences in innate ability or motivation. Paul Taubman (1976) provided an early example of such a study, which found that the return to schooling was only 3 percent. However bias from measurement error is likely to be greater in any method that identifies the return to education from differences in education. The work by Orley Ashenfelter and Alan B. Krueger (1994) attempted to reduce the potential for measurement-error-induced bias by instrumenting the education of one twin with an estimate obtained from responses from the other twin. This yielded estimates of schooling returns of some 16 percent per year of schooling; a threefold increase over typical findings. The data for their paper were collected at an annual "twins day" festival and may therefore suffer from the nonrandom nature of the sample, but the innovation of having two estimates of education for each individual does make this an important methodological contribution to the literature. Moreover, Ashenfelter and

David J. Zimmerman (1993) use sibling data to provide support for these twins results. Recent work by David G. Blanchflower and Peter Elias (1993) on U.K. twins data (extracted from the U.K. National Child Development Survey panel) estimates the returns to schooling using test scores as instruments for ability. While finding evidence of upward bias in standard least-squares estimates Blanchflower and Elias show strong evidence that, in comparison to nontwins of a similar age and environment, a sample of twins can be quite different. Thus, generalizing to the nontwin population may be difficult. ${ }^{1}$ A third procedure for addressing endogeneity is to treat ability as a fixed effect and to use panel data. However, the rate of return from panel data can only be identified from individuals who return to school, which may also involve endogeneity ( see Joshua D. Angrist and Whitney K. Newey [1991], who find increases in schooling levels for some 19 percent of continuously employed males in the National Longitudinal Survey of Youth [NLSY]).

The fourth approach exploits natural variation in data caused by exogenous influences on the schooling decision. Angrist and Krueger (1991) explore how an individual's season of birth may imply that some students reach school-leaving age after fewer months of compulsory education than others. Angrist and Krueger (1992) exploit the Vietnam-era draft lottery in the United States, which induced a change in educational participation. The essence of this "natural experiment" approach is to provide a suitable instrument for schooling. The conclusions in Angrist and Krueger (1991, 1992) suggest only a limited impact of endogeneity. However Card (1993) exploits data on proximity to educational institutions on the grounds that people living near a college are more likely to avail themselves of the facility than someone living farther from college. In addition, Kristin F. Butcher and Anne Case (1994) use the sex of children in the household to provide instruments for schooling for women. In both cases the evidence points toward an increase in the

[^1]estimated return to schooling after correcting for endogeneity via IV procedures, and although their results appear to be fairly robust to changes in specification, the IV estimates are not well determined, and the rate of return is much larger than, but not significantly different from, the OLS estimate.

In this paper we also adopt an IV approach. We rely on the exogenous changes in the educational distribution of individuals caused by the raising of the minimum school-leaving age in the United Kingdom (which has occurred twice over the age-spread of the working-age individuals in our data) to provide instruments for schooling. In addition to the IV approach, we also estimate a selectivity model which allows both for the endogeneity of schooling and the fact that it is recorded as an integer in our data.

Section I outlines the data used in the paper. Section II gives our econometric specifications and estimates. Our results, whether from conventional IV estimates or from the selectivity model, suggest quite dramatic bias in OLS, and our estimate of the return to schooling is of the order of 16 percent.

## I. Schooling and Earnings Distributions in the United Kingdom

The data set used in this analysis is the U.K. Family Expenditure Survey (FES). Our sample consists of 34,336 employed males aged 18-64 in the year of interview, obtained from pooling the nine consecutive annual FES cross sections from 1978 to 1986 . $^{2}$ Some summary statistics are given in Table A1 in the Appendix.

Important legislation on schooling was introduced in the 1944 Education Act. Of particular interest in this study are the provisions for raising the minimum school-leaving age (SLA). The first increase, from 14 to 15 ,

[^2]

Figure 1. Schooling Distribution for the United Kingdom
occurred in 1947. A further increase in the SLA subsequently occurred in 1973 (see Micklewright et al. [1989] for further discussion). ${ }^{3}$ Figure 1 provides the distribution of schooling broken down by minimum SLA and suggests that the 1947 change was particularly influential in raising participation in postcompulsory education. ${ }^{4}$ That is, many of those who would otherwise have left at the old minimum stayed on beyond the new minimum. Since our data are pooled over nine years, we can compare age-specific cohorts who faced a minimum SLA of 14 with the same age cohorts who faced the minimum of 15 . Figure 2 shows the relationship between log-earnings residuals and schooling for 46 - 53 -year-old males who faced a minimum SLA of 14 (with confidence intervals also plotted), compared to

[^3]the corresponding group who faced the minimum SLA of 15 , partialling out all other effects. ${ }^{5}$ That is, the graph controls for all other observable differences between the two cohorts and shows that the two cohorts have an expected wage differential of only 3 percent at school-leaving age 15 , and this is not statistically significant. Moreover this regression of log earnings against schooling including the age, year, and a dummy for the change in the minimum SLA reveals that the latter is insignificant, with a $t$ value of 0.807 .

## II. Econometric Analysis and Estimates

The conventional approach estimates the following two-equation system describing log earnings ( $y_{i}$ ) and years of schooling ( $S_{i}$ ):

$$
\begin{gather*}
y_{i}=\mathbf{X}_{i}^{\prime} \boldsymbol{\delta}+\beta S_{i}+u_{i}  \tag{1}\\
S_{i}=\mathbf{Z}_{i}^{\prime} \boldsymbol{\alpha}+v_{i} \tag{2}
\end{gather*}
$$

[^4]

Figure 2. Schooling, Residual Earnings, and Minimum School-Leaving Age
where $\mathbf{X}$ and $\mathbf{Z}$ are vectors of observed attributes, $E\left(\mathbf{X}_{i}, u_{i}\right)=E\left(\mathbf{Z}_{i}, v_{i}\right)=0$, and $\beta$ is interpreted as the return to schooling (Card, 1993). Estimation of equation (1) by OLS will yield an unbiased estimate of $\beta$ only if $S_{i}$ is exogenous. Following previous work we therefore estimate (1) via IV methods.

Schooling is not, however, a continuous variable. In particular, in any data set, it seems likely that there will be a considerable proportion of observations at the minimum level of $S$. In addition, our data code school-leaving as an integer, and in any case, there is inflexibility on the date of leaving in any year that gives rise to some discreteness. Robert J. Willis and Sherwin Rosen (1979) consider college education as an endogenous dummy variable; Lawrence W. Kenny et al. (1979) extend this to the case in which schooling is a Tobit specification; and John Garen (1984) considers the case in which education is coded as an ordered integer and shows that the discrete nature of the choice set implies that standard simultaneous-equations estimators are not consistent due to the nature of the disturbances. We therefore also pursue an alternative method to deal with this. This is an extension of the Heckman two-step approach, whereby we replace equation (2) above with the latent model below. $S^{*}$ is the latent variable corresponding to $S$ and we treat this as an ordered probit model (James J. Heckman, 1979, 1990). Thus, our alternative model structure is

$$
\begin{equation*}
y_{i}=\mathbf{X}_{i}^{\prime} \boldsymbol{\delta}+\beta S_{i}+u_{i} \tag{3}
\end{equation*}
$$

$$
\begin{gather*}
S_{i}^{*}=\mathbf{Z}_{i}^{\prime} \boldsymbol{\alpha}+v_{i}  \tag{4}\\
S_{i}=j \quad \text { if } \mu_{j-1}<S_{i}^{*} \leq \mu_{j}
\end{gather*}
$$

where $j=0,1,2, \ldots$, and $\mu_{j-1}<\mu_{j}$ ensures that the probabilities are positive. The $\mu$ 's are unknown parameters, estimated together with $\alpha$, which indicate the threshold values for moving through the schooling participation decision.

As discussed by Heckman (1990), Charles F. Manski (1989), and Card (1993), identification in IV and in the alternative selectivity model is provided by including variables in the vector $\mathbf{Z}$ that are not contained in $\mathbf{X}$. That is, there must exist a variable which is a determinant of schooling that can legitimately be omitted from the earnings equation. Here identification is achieved by the inclusion of dummy variables that record the exogenous change in the minimum school-leaving age in the set of variables $\mathbf{Z}$. In particular, dummy variables are defined for individuals who entered their 14th year between 1947 and 1971 and hence faced a minimum SLA of 15, and for those entering their 14th year after 1971 who therefore faced a minimum SLA of 16. The minimum SLA of 14 is our omitted category. Both $\mathbf{Z}$ and $\mathbf{X}$ include age and age squared, which are used to capture the impact of experience, and year and region dummies to capture time- and geographic-specific effects. ${ }^{6}$
Table 1 summarizes the estimates for the OLS and conventional IV approaches. OLS yields an estimated return of over 6 percent, which is of a similar order of magnitude to results in other U.K. studies. ${ }^{7}$ However, it is clear from Table 1 that the OLS estimated return to schooling is markedly smaller than that estimated by IV. Note that identification here is achieved only through the inclusion of the two minimum SLA dummies and that these

[^5]Table 1-OLS and IV Estimates

| Independent variable | OLS log earnings |  | Reduced-form schooling |  | IV log earnings |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Coefficient | SE | Coefficient | SE | Coefficient | SE |
| Constant | -0.0609 | 0.023 | 2.2928 | 0.171 | -0.1825 | 0.031 |
| Years of schooling | 0.0613 | 0.001 | - | - | 0.1525 | 0.015 |
| Minimum SLA $=15$ | - | - | 0.5409 | 0.055 | - | - |
| Minimum SLA $=16$ | - | - | 0.1100 | 0.076 | - | - |
| Age | 0.0836 | 0.001 | -0.0058 | 0.009 | 0.0765 | 0.002 |
| Age squared | -0.0009 | 0.000 | -0.0005 | 0.000 | -0.0008 | 0.000 |
| Yorkshire | 0.0063 | 0.011 | 0.1402 | 0.058 | -0.0065 | 0.012 |
| Northwest | 0.0194 | 0.010 | 0.2600 | 0.056 | -0.0047 | 0.012 |
| East Midlands | 0.0044 | 0.011 | 0.1751 | 0.060 | -0.0120 | 0.012 |
| West Midlands | 0.0001 | 0.011 | 0.1778 | 0.057 | -0.0171 | 0.011 |
| East Anglia | -0.0140 | 0.014 | 0.1507 | 0.073 | -0.0286 | 0.014 |
| Southeast | 0.1194 | 0.009 | 0.8582 | 0.049 | 0.1392 | 0.016 |
| Southwest | -0.0141 | 0.011 | 0.4821 | 0.061 | -0.0592 | 0.014 |
| Wales | -0.0172 | 0.012 | 0.3201 | 0.066 | -0.0462 | 0.014 |
| Scotland | 0.0150 | 0.011 | 0.2518 | 0.058 | -0.0124 | 0.012 |
| Northern Ireland | -0.1092 | 0.020 | 0.5890 | 0.105 | -0.1642 | 0.022 |
| Year $=1979$ | 0.0749 | 0.008 | -0.0146 | 0.045 | 0.0754 | 0.009 |
| Year $=1980$ | 0.0962 | 0.008 | 0.0295 | 0.045 | 0.0929 | 0.009 |
| Year $=1981$ | 0.1131 | 0.008 | 0.1670 | 0.045 | 0.0969 | 0.009 |
| Year $=1982$ | 0.0990 | 0.008 | 0.1536 | 0.046 | 0.0841 | 0.009 |
| Year $=1983$ | 0.1137 | 0.009 | 0.2576 | 0.048 | 0.0887 | 0.010 |
| Year $=1984$ | 0.1268 | 0.009 | 0.2878 | 0.048 | 0.0995 | 0.010 |
| Year $=1985$ | 0.1504 | 0.009 | 0.4158 | 0.049 | 0.1114 | 0.011 |
| Year $=1986$ | 0.1516 | 0.009 | 0.5468 | 0.049 | 0.1010 | 0.012 |
| $\overline{R^{2}}$ : | 0.27134.336 |  | 0.147 |  | $0.197$ |  |
| $N$ : |  |  | 34,336 |  | $34,336$ |  |

are significant in the reduced-form schooling equation in Table $1 .{ }^{8}$ These results strongly suggest the endogeneity of schooling with respect to the wage, and this conclusion is sup-

[^6]ported when we include the residual from the schooling equation in the OLS earnings equation to provide a Hausman $t$ test for the endogeneity of schooling (see Garen, 1984; Richard Smith and Richard Blundell, 1986), resulting in a $t$ statistic of 6.5 .

John Bound et al. (1993) urge caution in the selection of instruments because a weak correlation between potential instruments and wages, even in apparently large samples, can result in a large bias in the IV estimates. They propose that the researcher check the quality of IV estimates in two ways. First, the $F$ statistic on excluded instruments in the reducedform schooling equation needs to indicate statistical significance. Secondly, they essentially argue that the addition of the instrument to the reduced-form equation needs to improve the $\overline{R^{2}}$ of that equation. A test on excluding our potential instruments from the reducedform equation yielded an $F$ statistic of 90.81 , and the partial $R^{2}$ obtained from regressing

Table 2-Selectivity Model

|  | Ordered-probit schooling |  | Selectivity corrected log earnings |  |
| :---: | :---: | :---: | :---: | :---: |
|  | Coefficient | SE | Coefficient | SE |
| Constant | 0.7881 | 0.101 | -0.0552 | 0.023 |
| Years of schooling | - | - | 0.1688 | 0.007 |
| Minimum SLA $=15$ | 0.8960 | 0.027 | - | , |
| Minimum SLA $=16$ | 0.9316 | 0.052 | - |  |
| Age | -0.0009 | 0.004 | 0.0779 | 0.001 |
| Age squared | -0.0004 | 0.000 | -0.0008 | 0.000 |
| Yorkshire | 0.1021 | 0.034 | 0.0068 | 0.011 |
| Northwest | 0.1643 | 0.032 | 0.0207 | 0.010 |
| East Midlands | 0.1109 | 0.035 | 0.0051 | 0.011 |
| West Midlands | 0.1252 | 0.033 | 0.0008 | 0.011 |
| East Anglia | 0.1218 | 0.041 | -0.0130 | 0.014 |
| Southeast | 0.4932 | 0.028 | 0.1324 | 0.009 |
| Southwest | 0.3117 | 0.034 | -0.0097 | 0.011 |
| Wales | 0.2442 | 0.036 | -0.0168 | 0.012 |
| Scotland | 0.1445 | 0.033 | 0.0149 | 0.011 |
| Northern Ireland | 0.3565 | 0.056 | -0.1032 | 0.020 |
| Year $=1979$ | 0.0302 | 0.022 | 0.0736 | 0.008 |
| Year $=1980$ | 0.0680 | 0.023 | 0.0950 | 0.008 |
| Year $=1981$ | 0.1231 | 0.022 | 0.1134 | 0.008 |
| Year $=1982$ | 0.1078 | 0.023 | 0.0992 | 0.008 |
| Year $=1983$ | 0.1827 | 0.024 | 0.1144 | 0.009 |
| Year $=1984$ | 0.2064 | 0.024 | 0.1278 | 0.009 |
| Year $=1985$ | 0.2378 | 0.026 | 0.1537 | 0.009 |
| Year $=1986$ | 0.3083 | 0.025 | 0.1573 | 0.009 |
| $\lambda$ | - | - | -0.0526 | 0.003 |
| $\mu(1)$ | 1.2790 | 0.015 | . | , |
| $\mu(2)$ | 2.1216 | 0.017 | _ | - |
| $\mu(3)$ | 2.3857 | 0.017 | - | - |
| $\mu(4)$ | 2.6639 | 0.018 | - | - |
| $\mu(5)$ | 2.7591 | 0.018 | - | - |
| $\mu(6)$ | 2.8269 | 0.018 | - |  |
| $\mu(7)$ | 3.0500 | 0.019 | - | - |
| $\overline{R^{2}}$ : |  |  |  |  |
| Log-likelihood: | -54,7 |  |  |  |
| $N$ : | 34,3 |  | 34, |  |

schooling against our two potential instruments, once common exogenous variables have been partialled out, has a value of 0.0046 . Both of these compare favorably with those reported in Bound et al. (1995). Thus, our estimates strongly support the findings of Ashenfelter and Krueger (1994) and suggest an estimate of schooling returns of more than 15 percent. We tested for the time stability of our estimates both by interacting the schooling variable with the year dummies and by estimating over each cross section separately. In no year did the estimated return to schooling vary by more than 1 percent from the esti-
mated values for 1978. The year dummies indicate rising wages, and the region dummies suggest lower wages in the southwest and northwest than in the northern region, but higher wages in the southeast, which includes London.

Table 2 gives comparable results for our alternative model, whereby the schooling equation is estimated by ordered probits and the log earnings equation is selectivity corrected by including the relevant hazard from the schooling probit. The minimum SLA dummies have highly significant effects on schooling. The estimated return to schooling again shows a high

Table A1-Family Expenditure Survey, 1978-1986

| Variable | All |  | Minimum SLA $=14$ |  | Minimum SLA $=15$ |  | Minimum SLA $=16$ |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| $\ln$ (WAGE) | 1.913 | 0.445 | 1.902 | 0.434 | 1.995 | 0.416 | 1.584 | 0.426 |
| Age left education | 16.15 | 2.203 | 14.91 | 1.836 | 16.55 | 2.289 | 16.73 | 1.440 |
| Age | 38.74 | 12.67 | 55.84 | 4.466 | 35.60 | 7.316 | 21.59 | 2.697 |
| Yorkshire | 0.088 | 0.283 | 0.088 | 0.284 | 0.085 | 0.279 | 0.101 | 0.301 |
| Northwest | 0.110 | 0.312 | 0.105 | 0.307 | 0.109 | 0.312 | 0.119 | 0.324 |
| East Midlands | 0.075 | 0.264 | 0.073 | 0.260 | 0.074 | 0.262 | 0.082 | 0.274 |
| West Midlands | 0.099 | 0.298 | 0.103 | 0.304 | 0.099 | 0.298 | 0.090 | 0.286 |
| East Anglia | 0.037 | 0.188 | 0.038 | 0.190 | 0.037 | 0.189 | 0.032 | 0.176 |
| Greater London | 0.306 | 0.461 | 0.317 | 0.465 | 0.300 | 0.458 | 0.311 | 0.463 |
| Southwest | 0.074 | 0.261 | 0.070 | 0.256 | 0.074 | 0.262 | 0.080 | 0.271 |
| Scotland | 0.051 | 0.219 | 0.050 | 0.218 | 0.050 | 0.218 | 0.054 | 0.227 |
| Northern Ireland | 0.089 | 0.285 | 0.081 | 0.273 | 0.101 | 0.302 | 0.050 | 0.217 |
| Wales | 0.013 | 0.115 | 0.013 | 0.112 | 0.013 | 0.113 | 0.017 | 0.129 |
| Year $=1979$ | 0.116 | 0.321 | 0.143 | 0.351 | 0.118 | 0.322 | 0.062 | 0.241 |
| Year $=1980$ | 0.116 | 0.321 | 0.140 | 0.347 | 0.116 | 0.320 | 0.075 | 0.264 |
| Year $=1981$ | 0.121 | 0.326 | 0.129 | 0.335 | 0.122 | 0.327 | 0.102 | 0.302 |
| Year $=1982$ | 0.118 | 0.322 | 0.113 | 0.317 | 0.119 | 0.324 | 0.118 | 0.323 |
| Year $=1983$ | 0.101 | 0.301 | 0.084 | 0.277 | 0.105 | 0.307 | 0.113 | 0.316 |
| Year $=1984$ | 0.104 | 0.306 | 0.091 | 0.287 | 0.103 | 0.304 | 0.135 | 0.342 |
| Year $=1985$ | 0.101 | 0.302 | 0.069 | 0.253 | 0.102 | 0.302 | 0.159 | 0.365 |
| Year $=1986$ | 0.102 | 0.303 | 0.058 | 0.234 | 0.100 | 0.300 | 0.193 | 0.395 |
| $N$ | 34,336 |  | 8,717 |  | 20,224 |  | 5,102 |  |

value, comparable to the IV results in order of magnitude. The strong conclusion of endogeneity is highlighted by the Mill's ratio ( $\lambda$ ), which is significant and negative, implying that least squares produces estimates that are biased downward. Notice that, apart from the schooling variable, the coefficients are essentially unchanged from the IV results reported in Table $1,{ }^{9}$ and the smaller standard errors point to a gain in efficiency, as suggested in Garen (1984).

## III. Conclusion

This paper estimates the returns to schooling for men in a standard human-capital model, using a large U.K. sample to address the endogeneity of schooling. We exploit the experimental nature of two changes in the minimum SLA to provide instruments for education.

In both conventional IV and in a selection model in which years of education is treated

[^7]as an ordered probit, the corrected estimates of the return to education indicate the presence of a large and negative bias in the least-squares estimate of the schooling-earnings relationship. This confirms recent U.S. work which uses IV methods and in which the OLS bias has been found to be large although not significantly different from zero. For example, Card (1993) suggests that the estimates of schooling returns using IV methods are almost double those found using OLS. Butcher and Case (1994) also show that IV methods produce double the OLS estimates. However, in both cases, the IV estimates are much less precise, and the differences from OLS are not statistically significant. Thus our results provide greater precision to substantiate U.S. evidence of much larger rates of return to education than OLS suggests. Moreover our results are consistent with the findings in Ashenfelter and Krueger (1994) and Ashenfelter and Zimmerman (1993) which use twins and siblings, respectively.

## Appendix

Summary statistics from the U.K. Family Expenditure Survey are given in Table A1.

## REFERENCES

Angrist, Joshua D. and Krueger, Alan B. 'Does Compulsory Schooling Attendance Affect Schooling and Earnings?', Quarterly Journal of Economics, November 1991, 106(4), pp. 979-1014.
. 'Estimating the Payoff to Schooling Using the Vietnam-Era Draft Lottery.' Na tional Bureau of Economic Research (Cambridge, MA) Working Paper No. 4067, 1992.

Angrist, Joshua D. and Newey, Whitney K. "Over-Identification Tests in Earnings Functions with Fixed Effects." Journal of Business and Economic Statistics, July 1991, 9(3), pp. 317-23.
Ashenfelter, Orley and Krueger, Alan B. 'Estimates of the Economic Return to Schooling from a New Sample of Twins." American Economic Review, December 1994, 84(5), pp. 1157-73.
Ashenfelter, Orley and Zimmerman, David J. 'Estimates of the Returns to Schooling from Sibling Data: Fathers, Sons, and Brothers.' National Bureau of Economic Research (Cambridge, MA) Working Paper No. 4491, 1993.
Atkinson, Anthony B. and Micklewright, John. "On the Reliability of Income Data in the Family Expenditure Survey." Journal of the Royal Statistical Society, January 1983, 146(1), pp. 33-61.
Blackburn, McKinley L. and Neumark, David. "Omitted Ability Bias and the Increase in the Return to Schooling." Journal of Labor Economics, July 1993, 11 (3), pp. 521-43. . "Are OLS Estimates of the Return to Schooling Biased Downward? Another Look.' Review of Economics and Statistics, May 1995, 77(2), pp. 217-29.
Blanchflower, David G. and Elias, Peter. '"Ability, Schooling and Earnings: Are Twins Different?'’ Mimeo, Dartmouth College, U.S., 1993.

Bound, John; Jaeger, David A. and Baker, Regina. '"Problems with Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogenous Explanatory Variable Is Weak." Journal of the American Statistical Association, June 1995, 90(430), pp. 443-50.

Butcher, Kristin F. and Case, Anne. "The Effect of Sibling Composition on Women's Education and Earnings.' Quarterly Journal of Economics, Fall 1994, 109 (3), pp. 531-63.
Card, David. 'Using Geographic Variation in College Proximity to Estimate the Return to Schooling.' National Bureau of Economic Research (Cambridge, MA ) Working Paper No. 4483, 1993.
. 'Earnings, Schooling and Ability Revisited.' National Bureau of Economic Research (Cambridge, MA) Working Paper No. 4832, 1994.
Garen, John. '"The Returns to Schooling: A Selectivity Bias Approach with a Continuous Choice Variable.'" Econometrica, September 1984, 52(5), pp. 1199-1218.
Griliches, Zvi. "Estimating the Returns to Schooling: Some Econometric Problems." Econometrica, January 1977, 45(1), pp. 122.

Griliches, Zvi and Mason, William M. 'Education, Income and Ability." Journal of Political Economy, May/June 1972, Part 2, 80(3), pp. S74-S103.
Halsey, Albert H.; Heath, Anthony and Ridge, John. Origins and destinations-Family, class and education in modern Britain. Oxford: Oxford University Press, 1980.
Heckman, James J. 'Sample Selection Bias as a Specification Error." Econometrica, January 1979, 47(1), pp. 153-61.
. "Varieties of Selection Bias." American Economic Review, May 1990 (Papers and Proceedings), 80(2), pp. 313-18.
Kenny, Lawrence W.; Lee, Lung F.; Maddala, G. S. and Trost, R. P. 'Returns to College Education: An Investigation of Self-Selection Bias Based on the Project Talent Data.' International Economic Review, October 1979, 20(3), pp. 775-89.
Lang, Kevin. "Ability Bias, Discount Rate Bias, and the Return to Education.' Mimeo, Boston University, 1993.
Machin, Stephen; Meghir, Costas and Whitehouse, Edward. '"Changes in the U.K. Wage Structure 1968-90: Identifying Macroeconomic Effects Across and Within Groups.' Mimeo, Institute for Fiscal Studies, London, 1993.
Manski, Charles F. "Anatomy of the Selection Problem.' Journal of Human Resources, Fall 1989, 24(3), pp. 343-60.

Micklewright, John; Pearson, Mark and Smith, Stephen. "Has Britain an Early School Leaving Problem?'’ Fiscal Studies, January 1989, 10(1), pp. 1-16.
Miller, Paul; Mulvey, Charles and Martin, Nick. "What Do Twin Studies Reveal About the Economic Returns to Education? A Comparison of Australian and U.S. Findings. American Economic Review, June 1995, 85(3), pp. 586-99.
Moghadam, Reza. "Wage Determination: An Assessment of the Rate of Returns to Education, Occupation, Region and Industry in Great Britain." Mimeo, London School of Economics, 1990.

Nickell, Stephen. "Cohort Size and Earnings." Mimeo, Oxford University, 1993.
Smith, Richard and Blundell, Richard. "An Exogeneity Test for a Simultaneous Equation Tobit Model with an Application to Labor Supply." Econometrica, June 1986, 54(3), pp. 679-85.
Taubman, Paul. "The Determinants of Earnings: Genetics, Family, and Other Environments: A Study of White Male Twins." American Economic Review, December 1976, 66(5), pp. 858-70.
Willis, Robert J. and Rosen, Sherwin. "Education and Self-Selection." Journal of Political Economy, October 1979, Part 2, 87(2), pp. S7-S36.

Copyright of American Economic Review is the property of American Economic Association and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.


[^0]:    * Harmon: Department of Economics, Maynooth College, Maynooth, County Kildare, Ireland; Walker: Department of Economics, Keele University, Keele, Staffordshire ST5 5BG, United Kingdom, and Institute for Fiscal Studies (IFS). Material from the Family Expenditure Survey, made available by the Central Statistical Office (U.K.) through the Economic and Social Research Council's (ESRC) Data Archive, has been used with the permission of the Controller of Her Majesty's Stationary Office. The authors gratefully acknowledge ESRC support for the Centre for Microeconomic Analysis of Fiscal Policy at IFS, and European Union Human Capital Mobility Programme grant number 930225. Helpful comments were received from Bruce Chapman, Michael Devereux, Richard Freeman, Patrick Geary, Eric Hanuschek, Arie Kapteyn, Gauthier Lanot, Costas Meghir, Elizabeth Symons, and participants in seminars at the University of Aarhus, the Australian National University at Canberra, Keele University, the 1994 conference of the Royal Economic Society, and the 1994 European Meeting of the Econometric Society. An anonymous referee did much to improve the content and its exposition. The usual disclaimer applies.

[^1]:    ${ }^{\text {' }}$ However, Paul Miller et al. (1995) also find large bias in OLS estimates.

[^2]:    ${ }^{2}$ Data prior to 1978 did not contain education, and data since 1986 does not contain some other characteristics. Anthony B. Atkinson and John Micklewright (1983) make extensive comparisons between FES data and other surveys, including large surveys of employers, and they find that the FES earnings data match other sources very closely.

[^3]:    ${ }^{3}$ This act applied to England, Wales, and Northern Ireland only. Scotland comes under a different governing regulation, and it experienced changes in the schoolleaving age in 1946 and 1976, rather than 1947 and 1973.
    ${ }^{4}$ The raising of the school-leaving age in 1947 and 1973 imposed real constraints on behavior, borne out by contemporary reports of overcrowding in schools and labor-market shortages at the time (see Stephen Nickell [1993] and Albert H. Halsey et al. [1980] for further discussion).

[^4]:    ${ }^{5}$ Earnings profiles comparing the young cohort who faced an SLA of 16 with the comparable cohort who faced the SLA of 15 are not shown because such a comparison would be contaminated by the shorter job tenures of the SLA $=16$ cohort .

[^5]:    ${ }^{6}$ We use age rather than the more conventional experience measure because measurement error in education will induce error in experience.
    ${ }^{7}$ For example, Stephen Machin et al. (1993) estimate a return to a year of schooling of 6.9 percent over a similar sample period. The results of Reza Moghadam (1990) also support this base estimate.

[^6]:    ${ }^{8}$ SLA is recorded in the data as the age when continuous full-time education ceased. We have dropped 1,300 missing observations on this variable. We also exclude 421 observations on individuals who are recorded to have left school at less than the prevailing minimum age. As the compulsory-schooling laws were strictly adhered to and tied to employment and social-welfare legislation this could suggest reporting error. More typically this problem and the appearance of missing values for schooling would occur if the individual is an immigrant and not educated in the U.K. system. We have also recoded everyone reporting school-leaving ages in excess of 23 to be 23, because we suspect that some individuals have reported interrupted spells of education, rather than the end of the continuous spell. This recoding affects 686 observations. Neither dropping those with low education nor the recoding of those with very high levels of education makes any effective difference to the reported results.

[^7]:    ${ }^{9}$ Pairwise comparisons of IV and selection-model estimates show that only the coefficients on Southwest, year $=1985$, and year $=1986$ exhibit significant differences.

