Education Policy Reform and the Return to Schooling from Instrumental Variables*

KEVIN J. DENNY

University College Dublin & Institute for Fiscal Studies, London

COLM P. HARMON

University College Dublin, University of Warwick & CEPR, London

Version 1.03, January 28th 2000

Abstract:

We exploit an unusual policy reform which had the effect of reducing the direct cost of schooling in Ireland in the early 1970's. This gave rise to an increased level of schooling but with effects that vary across family background. This interaction generates a set of instrumental variables which we use to estimate the return to schooling allowing for the endogeneity of schooling. We find a large and well determined rate of return of the order of 11 or 12% substantially higher than the OLS estimates of around 7%.

* Material from the Survey of Income Distribution, Poverty and Usage of State Services Survey is used with permission of the Economic and Social Research Institute in Dublin. The providers bear no responsibility for the calculations contained herein nor for any interpretation made by the authors. Our thanks to Ian Walker for the encouragement to pursue this paper, and to participants at seminars in Athens and NUI Maynooth for comments. This research has been funded by the EU Commission under the auspices of the TSER programme, contract number SOE2-CT98-2044 . Corresponding author: Dr. Kevin Denny, Department of Economics, University College Dublin, Belfield, Dublin 4, Ireland. Phone (+353 1) 7068399. Fax (+3531) 2830068. E-mail colm.harmon@ucd.ie.

The rate of return to schooling is an important factor in determining educational attainment but in recent years considerable research effort has been invested in both the manner of estimation and the magnitude of the return to schooling. This is due to a variety of sources of bias associated with OLS estimates of the return to schooling making recovery of an estimate of the return to schooling difficult.

Recent research on the returns to schooling (along with other areas of labor research) has drawn analogies with 'treatment' and 'control' group concepts in the medical/psychology literature. While it is clearly not feasible to pursue a random assignment exercise for years of schooling, the literature on returns to schooling is now largely focused on finding real-life (as opposed to experimental) events which can be considered as events that assign individuals randomly to different treatments, thus exploiting 'natural' variation in data. For example Joshua Angrist and Alan Krueger (1991) explores how an individual's season of birth may imply that some students reach school leaving age after fewer months of compulsory education than others, allowing for the creation of suitable instruments to exploit in an Instrumental Variables (IV) approach.

While the conventional wisdom had been that Ordinary Least Squares (OLS) was biased upwards (due to omitted ability measures for example) many of the IV studies show returns that are at best unchanged and often considerably larger than those found by OLS. Orley Ashenfelter *et al.* (1999) provide a survey of this literature and show that the average difference between IV and OLS is around 3% per year of schooling. Recent research (see Kevin Lang (1993), David Card (1999)) suggests that OLS estimates may be subject to discount rate bias arising from individuals with higher discount rates choosing less education in an optimizing model. In this framework OLS provides an estimate of the rate of return to education on average while IV provides an estimate of the rate of return for marginal individuals with high discount rates, so, in principle, the OLS estimates can be higher or lower than those found under IV[1].

In this paper we also adopt an IV approach. We rely on the exogenous changes in the educational distribution of individuals caused by a policy innovation in Ireland in the early 1970's whereby secondary schooling was made free for all school-going youths. This reform had a dramatic effect on the participation rate in education - a summary indication of the changes in the system can be gleaned from the fact that the proportion of the age-cohort taking the first set of secondary schooling examinations rose from about 40% in 1967 to close to 100% in 1994, while the proportion taking higher level post compulsory secondary schooling rose from about 21% to about 82% (see Joseph Durkan *et al* (1997) for a detailed analysis of the education composition of the labor force).

I. Reforms as Instruments and the Identification Strategy

In this paper we estimate the following two-equation system describing log earnings, (y_i), and years of schooling, (S_i)

(1) $$y_i = \mathbf{X}'_i \delta + \beta S_i + u_i$$

(2) $$S_i = \mathbf{Z}'_i \alpha + v_i$$

[1] Ashenfelter *et al*. (1999) presents an alternative explanation for the dominance of IV results that are higher than OLS. By a meta-analysis of some 100 estimated rates of return they find that the average premium of around 3% of IV over OLS may be partly (1.8%) explained by selective reporting of results by researchers.

where **X, Z** are a vectors of observed attributes, $E(\mathbf{X_i} u_i) = E(\mathbf{Z_i} v_i) = 0$, and b is interpreted as the return to schooling (Card (1993)). Estimation of equation (1) by OLS will yield an unbiased estimate of b only if the S_i is exogenous.

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 **Z** that are not contained in **X**. That is, there must exist a variable which is a determinant of schooling that can legitimately be omitted from the earnings equation. In the paper by Esther Duflo (1999) estimation is based on the exposure of individuals to a massive investment program in education in Indonesia in the early 1970's. Individuals were assigned to the treatment on the basis of their date of birth (pre and post reform) and the district they lived in (as investment was a function of local level needs assessment). Costas Meghir and Maarten Palme (1999) pursue a similar strategy in their examination of reforms in Sweden in the 1950's which aimed to extend the schooling attainment nationally. This was piloted in a number of school districts prior to national adoption and it is from this pre-trial experiment that the variation in attainment comes.

In this paper identification is achieved by the inclusion of dummy variables that record the exogenous influence on schooling caused by the abolition of fees at secondary school in Ireland. This remarkable policy shift which occurred in the late 1960's in Ireland where fees, which were payable upon entry to secondary school of any type (private or state), were removed and full state funding at secondary level was introduced (see Tussing (1978) for a complete description). The prevailing fee-paying aspect to secondary education was a major hurdle for families, so typically among the older generation those that received secondary (and by implication third level) education came from a wealthier socio-economic background.

Assignment to the treatment is a function of your date of birth. In particular a dummy variable is defined for individuals who entered were born on or after 19XX and hence faced a regime of no fees at secondary school. minimum schooling age of 15, and for those entering their fourteenth year after 1971 who therefore faced a minimum SLA of 16. The minimum SLA of 14 is our omitted category. Both **Z** and **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.[2]

The data used in this study is a sample of employed males from a household survey conducted by the Economic and Social Research Institute in 1987 (hereafter ESRI87)[3]. A total of 3,300 households were interviewed, generating information on over 6,500 adults. The total number of male employees aged 18-64 in year of interview for whom the necessary information on pay, hours of work and labour market experience is available is 1158. More complete information on this data can be found in Callan and Harmon (1999). The principal focus is the gross hourly wage rate. We estimate a simple earnings function using years of education as determined from information in the data on the age the respondent first left full time education. As with other studies we include a quadratic in age to proxy for experience given the possible endogeneity of labour market experience. We also include the participation rate in education when the individual was 15 to proxy for trends in participation and to eliminate the potential interpretation of the free fees dummy as a cohort effect alone. The remaining variable in the earnings function include a geographic dummy for those resident in Dublin.

[2] We use age rather than the more conventional experience measure because measurement error in education

II. Results

Table 1: Estimated Schooling and Earnings Functions – ESRI87

	OLS Earnings		OLS Schooling		IV Earnings	
	Co-eff.	***Std. Err.***	**Co-eff.**	***Std. Err.***	**Co-eff.**	***Std. Err.***
Years of Schooling	.0762	*.0053*	--	--	.1213	*.0254*
Age	.1537	*.0117*	.1662	*.0675*	.1496	*.0123*
Age ^{2}	-.0014	*.0001*	-.0013	*.0005*	-.0013	*.0001*
Trend in Secondary Participation	.0128	*.0024*	.0659	*.0202*	.0117	*.0025*
Resident in Dublin	.1157	*.0246*	.0668	*.1342*	.1139	*.0253*
Parental Class 2 (= Admin/Clerical)	.0046	*.0567*	-.0500	*.3743*	.0232	*.0593*
Parental Class 3 (= Other Non-Manual)	-.0508	*.0483*	-.8752	*.3239*	-.0022	*.0565*
Parental Class 4 (= Skilled Manual)	-.1104	*.0455*	-2.635	*.2979*	-.0105	*.0724*
Parental Class 5 (= Semi-Skilled Manual)	-.0792	*.0504*	-3.214	*.3252*	.0432	*.0853*
Parental Class 6 (= Unskilled Manual)	-.1290	*.0488*	-3.686	*.3057*	.0124	*.0929*
No Fees? (1 if born > 1954)	--	--	-1.803	*.5938*	--	--
Parental Class 2 * No Fees	--	--	-1.217	*.6625*	--	--
Parental Class 3 * No Fees	--	--	-.3414	*.5488*	--	--
Parental Class 4 * No Fees	--	--	1.230	*.5043*	--	--
Parental Class 5 * No Fees	--	--	1.468	*.5510*	--	--
Parental Class 6 * No Fees	--	--	1.643	*.5283*	--	--
Constant	-3.425	*.3988*	6.742	*2.499*	-3.888	*.4844*
N	1158		1158		1158	
$\overline{R}^2$	0.4623		0.2244		0.4284	
					2.952	
					(.7074)	

The first column gives the OLS estimates of the earnings equation with an estimated rate of return equal to 7.6% an estimate that is consistent with the existing literature. Other specifications lead to fairly similar results, for example including a trade union dummy, which

will induce error in experience.

itself has a large and well determined coefficient (of the order of .15) does not change the estimated schooling return. Given the possible endogeneity of union status we omit this variable. The second set of estimates is the schooling education used in the IV estimation. What is notable here is the importance of parents socio-economic background and how it change with the introduction of the reform. The coefficients on parental class show that does from poorer backgrounds have lower education *ceteris paribus* with non-manual (or blue collar) backgrounds at a distinct disadvantage compared to those with non-manual backgrounds. This is consistent with a wide body of other, mostly sociological, research. The policy reform has the effect of reducing these socio-economic penalties, approximately by a half. This effect is consistent with studies of the reform (Tussing (1978)). For our purposes this gives us a set of instruments. The final column is the IV estimate of the earnings equation. The instruments are the interaction of the reform and socio-economic background. The direct effects of family background are included in the earnings equation since they may be correlated with unobserved characteristics such as motivation, see Isaac Rischall (1999).[4] Card (1999) has argued that family background may be a poor instrument and in an application to Finnish data, Conneely and Uusitalo (1999) reject the hypothesis that it is uncorrelated with the error term in the earnings equation. The outcome is that there is a large rise in the estimated return to education, over 12%. This is at the high end of estimated return though not as high as the 15% estimated by Harmon and Walker (1995) who also use a "natural experiment" with British data. Using the standard test (e.g. Davidson & McKinnon, (1993)) we are unable to reject the over-identifying restrictions.

[3] More recent data at the individual level is not publicly available.

[4] Omitting them from the earnings equation , that is treating them as instruments, reduces the estimated return slightly to .117 while continuing to pass the test for over-identifying restrictions.

Other estimates similar to those presented in the above table were computed and are available on request. In summary however the estimated return to schooling under IV ranges from a low of 0.1154 to a high of 0.1306, compared to the range of OLS returns of 0.076 to 0.079.

Much of the recent literature on schooling returns has emphasised the possibility of heterogeneous returns typically using a variant of a model due to Gary Becker (e.g., Conneely & Uusitalo (1999), Kling (1999)). One of the appealing aspects of this approach is that as Card shows, with heterogeneous returns and a binary treatment the IV return can be written as a weighted average of the returns of the sub groups.

Note that one advantage of the standard IV estimator with a binary treatment is that it may identify the Average Causal Response (Angrist & Imbens 1995 JASA). However a necessary condition is monotonocity, that in this case, the reform does not decrease educational attainment for any individual. As the schooling equation shows this does *not* hold since those from better off backgrounds had lower attainments. This is essentially because they were partially "crowded out" by higher attainment at the other end of the socio-economic distribution. Given the presence of binding supply side constraints it is not surprising that a fall in relative prices to one sub-group should have this impact on the others. The results here are closest in spirit to that Card (1995) (and subsequently Kling(1999)) who uses an interaction between family background and proximity to college as an instrumental variable with bigger effects on schooling for those from poorer backgrounds.

III. Conclusion

In a standard model of education and earnings we exploit an unusual policy reform which had the effect of reducing the direct cost of schooling. This gave rise to an increased level of schooling but with effects that vary across family background. This interaction generates a set of instrumental variables which we use to estimate the return to schooling allowing for the endogeneity of schooling. We find a large and well determined rate of return of the order of 11 or 12% substantially higher than the OLS estimates of around 7%.

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.

Callan, Tim and Colm Harmon (1999), "The Economic Return to Schooling in Ireland", forthcoming Labour Economics, December 1999.

Card, D. (1995). "Using Geographic Variation in College Proximity to Estimate the Return to Schooling." In Christofides, Louis; Grant, E. Kenneth; Swidinsky, Robert (eds.) Aspects of Labour Market Behaviour: Essays in Honour of John Vanderkamp. Toronto; Buffalo and London: University of Toronto Press, 1995, pages 201-22.

_____. (1999) "The Causal Effect of Education." In Ashenfelter, O. and Card, D. (eds) Handbook of Labor Economics. Amsterdam: Elsevier.

Conneely, K. and Uusitalo, R. (1999). "Estimating Heterogeneous Treatment Effects in the Becker Schooling Model." Mimeo, Princeton University.

Davidson, Russell and McKinnon, James (1993). Estimation and Inference in Econometrics. Oxford: Oxford University Press.

Duflo, Esther (1999). "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment." *Mimeo*, Department of Economics, MIT.

Harmon, C. and Walker, I. (1995). "Estimates of the Economic Return to Schooling for the UK, American Economic Review, 85, pp. 1279-1286.

Kling, Jeffrey (1999). "Interpreting Instrumental Variables Estimates of the Returns to Schooling." Industrial Relations Section, Princeton University Working Paper 415, January.

Meghir, Costas and Palme, Marten (1999). "Assessing the Effect of Schooling on Earnings Using a Social Experiment." Institute for Fiscal Studies Working Paper 99/10, November.

Rischall, Isaac (1999). "The Roles of Education, Skill and Parental Income in Determining Wages." *Mimeo*, McMaster University.

Tussing, Dale (1978). "*Irish Educational Expenditures - Past, Present and Future.*" E.S.R.I. General Research Series no. 92.