Journal of Applied Economics, Vol. VIII, No. 1 (May 2005), 191-202

NONLINEARITY IN THE RETURN TO EDUCATION

Philip A. Trostel*

University of Maine

Submitted February 2003; accepted May 2004

This study estimates marginal rates of return to investment in schooling in 12 countries. Significant systematic nonlinearity in the marginal rate of return is found. In particular, the marginal rate of return is increasing significantly at low levels of education, and decreasing significantly at high levels of education. This may help explain why estimates of the return to schooling are often considerably higher when instrumenting for education.

JEL classification codes: I20, J24

Key words: return to education, nonlinearity, instrumental variables

I. Introduction

The rate of return to education has been estimated in literally hundreds of studies (see the surveys by Psacharopoulos, 1985, 1994; Ashenfelter et al., 1999; and Harmon et al. 2000). The vast majority of this work implicitly assumes that the marginal rate of return is constant over all levels of education. Some studies, however, found significant nonlinearity in the rate of return to schooling. Most of this work focused on deviations from nonlinearity at particular levels of education; that is, sheepskin effects (see, for example, Hungerford and Solon, 1987; Belman and Heywood, 1991; and Jaeger and Page, 1996). Perhaps as a result, evidence on the general nonlinearity in the return to schooling appears inconsistent. Mincer (1974), Psacharopoulos (1985, 1994), and Harmon and Walker (1999) showed significant diminishing

* Philip Trostel: Department of Economics & Margaret Chase Smith Center for Public Policy, University of Maine, Orono, ME 04469-5715; email: philip.trostel@maine.edu.
For constructive comments I am grateful to the editor, a referee, Ian Walker, and seminar participants at the University of Warwick.

returns to education.[1] Heckman and Polachek (1974), Card and Krueger (1992), and Card (1995, 1999) argued that the rate of return appears roughly constant. One could, however, interpret Card and Krueger's (1992) results as indicative of increasing returns at low levels of education. The results in Heckman et al. (2003) suggest increasing returns at low levels of education followed by diminishing returns at high levels of education. The general nature of possible nonlinearity in the return to education is unclear.

This study tests for the general nonlinearity in the (private) rate of return to education for working-age men using comparable micro data in 12 countries. The data indicate that the marginal rate of return is essentially nil for the first several years of schooling, it then increases rapidly until about year 12, and then it declines.

II. Data

Data from the International Social Survey Programme (ISSP) are used. The ISSP contains comparable cross-sectional data on individuals in 33 countries from 1985 through 1995 (most of the countries, however, only participated in a few of the years). Only 13 of the countries have at least 1,000 observations of labor-market data for men, and measured schooling is truncated between 10 and 14 in one of these countries (Great Britain). Thus observations from Great Britain are excluded, leaving samples from 12 countries.

The 12 samples consist of men within the ages of 18 to 64; without missing information on wage rates or education; and not self-employed, retired, or in school. A handful of observations with more than 22 years of measured education are also excluded.[2] Table 1 lists for each country its: sample size, number of cross sections, mean years of education, and standard deviation of education.

---

[1] Although not directly comparable to this literature, diminishing returns to schooling are also suggested by the relatively high return to early interventions relative to later remedial interventions (see Carneiro and Heckman, 2003).

[2] The results are essentially invariant to either censoring or truncating the schooling data at 20 years or any reasonable higher level.

Table 1. Summary Statistics

Country	N	Years	$\bar{S}$	s_s
West Germany	3,396	9	10.53	3.07
United States	3,347	11	13.54	2.90
Australia	3,090	6	11.64	2.78
Norway	2,751	7	12.52	2.97
Russia	2,537	5	13.09	3.39
Netherlands	2,215	6	13.21	3.72
Austria	1,755	8	11.01	2.57
Poland	1,456	5	11.07	2.66
Italy	1,347	6	11.87	3.89
East Germany	1,238	5	10.86	2.86
Ireland	1,176	6	12.10	3.07
New Zealand	1,126	5	12.69	3.14

III. Evidence of Nonlinearity

The equations to be estimated are simple nonlinear extensions of the standard Mincer wage equation. That is, an education polynomial is used in a log-wage equation rather than just a linear term:

$$\ln(w_i) = \beta_0 + \sum \beta_j S_i^j + \sum \beta_{Eh} E_i^h + \beta'_Y Y_i + \varepsilon_i \quad (1)$$

where w_i is the hourly wage rate of individual i, S is years of schooling, j is the order of the education polynomial, E is potential experience (age minus years of schooling minus six), h is the order of the experience polynomial (following Murphy and Welch, 1990, this is a fourth-order polynomial), and Y is a vector of indicator variables for each year.[3]

[3] Earnings are measured in categories in many of the countries. Thus, all of the results reported are from maximum-likelihood interval regressions. The results, however, are essentially the same for OLS regressions on category midpoints.

To determine the appropriate order of the education polynomial, likelihood ratio tests are conducted for different versions of equation (1). In particular, the λ_{kl} reported in Table 2 test for the difference in the model when using an education polynomial of order l compared to order k. The evidence indicates that a third-order polynomial is generally necessary to adequately describe the education profile of wages. At the 90% confidence level, the addition of S^2 significantly improves the fit of equation (1) in only four of the 12 countries. The addition of S^3, however, is statistically significant at this level in nine of the 12 countries. But the addition of S^4 is significant in only one of the 12 cases. This evidence indicates that the estimated private marginal rate of return, ρ, is a quadratic function of years of schooling:

$$\hat{\rho}(S) = \hat{\beta}_1 + 2\,\hat{\beta}_2 S + 3\,\hat{\beta}_3 S^2 \tag{2}$$

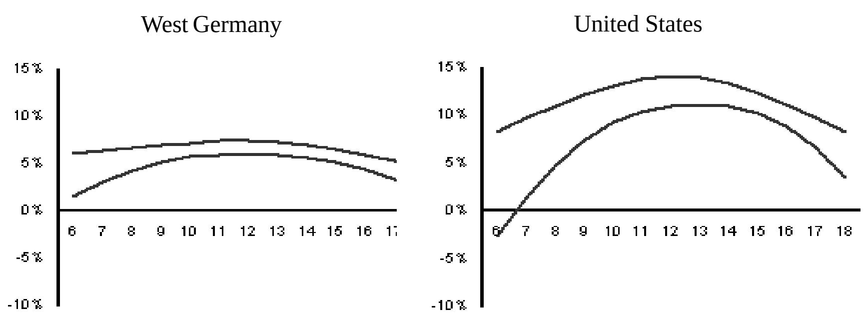
Table 2. Likelihood Ratio Tests

Country	λ_{12}	λ_{23}	λ_{34}
West Germany	4.14	9.68	0.25
United States	0.02	20.06	0.14
Australia	2.33	16.18	0.33
Norway	0.05	10.37	0.28
Russia	0.00	1.71	1.48
Netherlands	0.38	3.54	0.10
Austria	4.20	4.06	0.46
Poland	0.06	3.01	0.04
Italy	3.68	2.55	0.10
East Germany	0.02	2.40	1.25
Ireland	1.06	5.40	10.39
New Zealand	8.07	8.85	1.28
Weighted average	1.77	8.84	0.92

Note: These likelihood ratio tests are χ^2 statistics with one degree of freedom.

The results of estimating equation (1) as a third-order polynomial in education are summarized in Table 3. The estimated coefficients on the education polynomials are reported along with the implied marginal rates of return for 8, 12, and 16 years of education (for comparison, estimates of the standard linear rate of return are also shown). The equation (2) results are illustrated in Figure 1 for West Germany and U.S.A.

Figure 1. 95% Confidence Interval of the Marginal Rate of Return to Education



As emphasized in Trostel et al. (2002), there is considerable cross-country variation in the linear rate of return to education.[4] Yet there is considerable cross-country similarity in the nonlinearity in the rate of return to education. In all 12 countries the coefficient estimates on S and S^3 are negative, and the coefficient on S^2 is positive. Moreover, $\hat{\beta}_2$ and $\hat{\beta}_3$ are statistically significant with at least 95% confidence in eight of the 12 countries.

Although the levels of the estimates of the nonlinear marginal rates of return to education vary considerably across countries, their nonlinear pattern is quite consistent. $\hat{\rho}(8)$ and $\hat{\rho}(16)$ are lower than $\hat{\rho}(12)$ in all 12 countries. Indeed, the marginal rates of return at 8 and 16 years of schooling are usually substantially below those at 12 years. The weighted average $\hat{\rho}(8)$ and $\hat{\rho}(16)$ are 70.4% and 75.4% of the weighted average $\hat{\rho}(12)$. Moreover, the differences in the estimated marginal rates of return are even greater at schooling levels below 8 and above 16. The levels where the estimated marginal rates of return reach a maximum lie in the narrow range between 11.5 and 12.9 years of

[4] Trostel et al. (2002) also use ISSP data. Their estimates of the rate of return are somewhat lower than the linear estimates in Table 3 because their estimation used an age polynomial instead of an experience polynomial.

Table 3. Rate of Return Estimates

Country	$\hat{\beta}_1$ x 10^2	$\hat{\beta}_2$ x 10^3	$\hat{\beta}_3$ x 10^4	$\hat{\rho}$ (8)	$\hat{\rho}$ (12)	$\hat{\rho}$ (16)	Linear $\hat{\rho}$
West Germany	-5.45	10.31	-2.91	0.06	0.07	0.05	0.06
	(4.29)	(3.47)	(0.90)	(0.01)	(0.00)	(0.00)	(0.00)
United States	-23.28	28.58	-7.59	0.08	0.13	0.10	0.10
	(9.24)	(7.21)	(1.83)	(0.02)	(0.01)	(0.01)	(0.01)
Australia	-14.43	18.99	-5.41	0.06	0.08	0.05	0.06
	(3.83)	(3.76)	(1.17)	(0.01)	(0.01)	(0.01)	(0.00)
Norway	-12.34	14.10	-3.65	0.03	0.06	0.05	0.05
	(4.32)	(3.50)	(0.92)	(0.01)	(0.00)	(0.00)	(0.00)
Russia	-6.71	9.53	-2.55	0.04	0.05	0.04	0.05
	(7.38)	(6.13)	(1.67)	(0.01)	(0.01)	(0.01)	(0.00)
Netherlands	-1.98	5.63	-1.53	0.04	0.05	0.04	0.04
	(3.51)	(2.83)	(0.73)	(0.01)	(0.00)	(0.00)	(0.00)
Austria	-7.90	13.13	-3.75	0.06	0.07	0.05	0.06
	(6.78)	(5.51)	(1.42)	(0.01)	(0.01)	(0.01)	(0.00)
Poland	-17.36	23.36	-6.77	0.07	0.10	0.05	0.08
	(16.51)	(14.60)	(4.15)	(0.02)	(0.01)	(0.02)	(0.01)
Italy	-3.03	8.09	-2.47	0.05	0.06	0.04	0.05
	(6.79)	(5.80)	(1.55)	(0.01)	(0.01)	(0.01)	(0.00)
East Germany	-21.67	19.95	-5.15	0.00	0.04	0.03	0.03
	(13.39)	(10.58)	(2.71)	(0.02)	(0.01)	(0.01)	(0.00)
Ireland	-6.89	16.07	-4.49	0.10	0.12	0.10	0.11
	(10.30)	(8.06)	(2.08)	(0.02)	(0.01)	(0.01)	(0.01)
New Zealand	-12.25	13.88	-3.43	0.03	0.06	0.06	0.05
	(3.44)	(3.42)	(1.02)	(0.01)	(0.01)	(0.01)	(0.01)
Weighted average	-11.34	15.41	-4.22	0.05	0.07	0.06	0.06
	(6.82)	(5.67)	(1.52)	(0.01)	(0.01)	(0.01)	(0.00)

Note: Robust standard errors are in parentheses. The regressions include controls for potential experience (fourth-order polynomial) and for each year.

education in all but two of the 12 countries. Even the two outliers in this respect (10.9 years Italy and 13.5 years in New Zealand) are not far from the others.

Thus, the evidence indicates that the marginal rate of return is increasing significantly at relatively low levels of education, and decreasing significantly at relatively high levels of education. Hence, linear estimates of the rate of return (that is, weighted average marginal rates of return within countries) noticeably understate the maximum marginal rates of return around 12 years of schooling, and substantially overstate the rates of return at both the low and high levels of education. Indeed, the return to investment in education is insignificant for the first several years. Evidently, the initial increasing returns in human capital production are substantial.

Because the marginal rate of return is lower at both ends of the education distribution, a first-pass test for a non-constant rate of return does not reveal much nonlinearity. The l_{12} in Table 2 are generally not significant (and the coefficient estimates on S^2 in a quadratic version of equation (1) are generally not significant) because the initial increasing returns are offset by the later diminishing returns.[5]

Various versions of equations (1) and (2) were estimated to check the sensitivity of the nonlinearity in the rate of return.[6] The nonlinearity results were found to be robust. The marginal rate of return to education for women displays a nonlinear relationship that is similar to men. Indeed, the nonlinearity in the rate of return is somewhat more pronounced for women than men. The results are essentially the same when using (log) monthly earnings as the dependent variable rather than the (log) hourly wage rate. Similarly, the results are essentially the same when using a second-order, instead of a fourth-order, polynomial in potential experience (as is common in the literature). Similar results are also found when using an age polynomial instead of a potential experience polynomial. The nonlinearity results are unaffected when including a schooling- experience interaction term, which allows for schooling to affect

[5] Similarly, Box-Cox estimates of the relationship between w and S, such as in Heckman and Polachek (1974), are extremely close to log-linearity (thus suggesting a near constant rate of return).

[6] These results are available by request to the author.

the experience profile of wages.[7] The nonlinearity relationship remains, albeit somewhat weaker, when dropping observations from either or both tails of the education distribution. It is not just the extremes of the education distribution that produce the estimated nonlinearity. A similar picture also generally emerges when estimating equations (1) and (2) for ISSP countries with less than 1,000 observations, and for each year separately.[8]

IV. Discussion

One problem with the above estimates of marginal rates of return is that education is potentially endogenous. For this reason numerous recent studies have used natural experiments as instruments to identify the causal effect of education (see, for example, the recent survey by Card 1999). The ISSP, however, does not contain good instruments for education. But even if there were good instruments in the dataset, it is unlikely that they could yield unbiased estimates of marginal rates of return. As stressed by Card (1995, 1999), typical instruments for education capture the causal effect of education only at one point or over a small range of education outcomes. In the present context where nonlinearity is explicitly examined, one needs valid instruments for the entire range of education outcomes. Such instruments might not be available in any dataset.

Moreover, the magnitude of the rate of return to education is not the primary issue in this study. The issue is nonlinearity in the rate of return. Hence, the primary concern in the present context is whether potential endogeneity of

---

[7] The schooling-experience interaction coefficient is negative in nine of the 12 countries (i.e., the experience profile of wages usually flattens as education rises). Four of the nine negative instances and two of the three positive cases are statistically significant (at 90%). In no instances, though, does the inclusion of the interaction term appreciably affect the nonlinearity in the rate of return.

[8] A similar nonlinear relationship can also be found in larger datasets. Comparable estimates from the U.K. Family Resources Survey in 1995 (N = 9,037) are $\hat{\beta}_1 \times 10^2 = -11.71$ (5.05), $\hat{\beta}_2 \times 10^3 = 26.79$ (4.00), and $\hat{\beta}_3 \times 10^4 = -8.88$ (1.06). Comparable estimates from the U.S. Current Population Survey in 1991 (the last year that education was measured as years of schooling instead of credentials) (N = 62,493) are $\hat{\beta}_1 \times 10^2 = 3.29$ (0.92), $\hat{\beta}_2 \times 10^3 = 4.19$ (0.92), and $\hat{\beta}_3 \times 10^4 = -0.84$ (0.28). Moreover, the nonlinearity in the CPS estimates occurs despite the measure of education being top-coded at 18 years.

education can explain the observed nonlinearity in the rate of return. It appears that it cannot.

Endogeneity of education can potentially explain the rising rate of return at the low end of the education distribution, but it works against observing a declining rate of return at the high end. As again stressed by Card (1995, 1999), unobserved heterogeneity in ability, family background, etc. is likely to cause schooling and wages to be positively correlated independent of the causal effect of schooling. Thus, independent of the causal effect of schooling, observed wages are likely to rise with the level of education. To the extent that this is true, the observed marginal rate of return is rising with education independently of its causal effect. The extent that this can explain the observed rising return at low levels of education is, of course, unclear. In any event, the direction of bias caused by this endogeneity is in one direction. Hence, it cannot explain the observed diminishing returns at relatively high levels of education.

Given that typical instruments for schooling capture the causal effect of education only at small middle ranges in the distribution of schooling outcomes, and that the (OLS) marginal rate of return is higher over the middle range than over the entire education distribution; there is reason to expect IV estimates of the rate of return to be greater than OLS estimates, a result frequently found in the literature. This is essentially the problem stressed by Card (1995, 1999), although in slightly different form. Card conjectured that the marginal rate of return is declining throughout. Thus, instruments that affect relatively low levels schooling will produce upwardly- biased estimates of the average marginal rate of return. If, however, the marginal rate of return is lower at both ends of the schooling distribution, then any instrument that truncates this distribution toward the middle will yield an upwardly-biased estimate of the average marginal rate of return, even those that affect relatively high levels of schooling. Perhaps this can explain the finding in Harmon and Walker (1999) that IV estimates are higher than OLS estimates even when the instruments affect different schooling levels (a result that should not occur if the marginal rate of return is declining throughout).

Another issue in the preceding empirical work is whether educational sorting can explain the observed nonlinearity in the marginal rate of return.

Although not completely decisive, there is considerable evidence that years of education are sorted by ability, tastes, work attitudes, and so forth (e.g., Hungerford and Solon, 1987; Belman and Heywood, 1991; Kroch and Sjoblom, 1994; Groot and Oosterbeeck, 1994; Weiss, 1995; and Jaeger and Page, 1996). The preceding estimates subsumed possible sheepskin effects. That is, the continuously-estimated nonlinearity could simply be reflecting discrete changes at degree-completion years. This issue in this context, however, is essentially the same as the possible endogeneity of schooling. In particular, for essentially the same reason as above, educational sorting can potentially explain at least some of the rising return at the low end of the education distribution, but it works against finding diminishing returns at the high end (unless there is some a priori reason for there to be sheepskin effects at the secondary and undergraduate levels, while not at the graduate level).

V. Conclusion

Private marginal rates of return to education were estimated from comparable micro data from 12 countries. Economically and statistically significant nonlinearity was found in the return to education. Substantial increasing returns were generally found in primary and secondary education. Substantial diminishing returns were generally found in higher education. Standard linear estimates of the rate of return to education substantially overstate the marginal rates of return at both low and high levels of schooling, and they noticeably understate the maximum rate of return at middle levels of education.

The results also suggest that estimating the return to education is even more problematic than perhaps previously believed. Using natural experiments as instruments to identify the causal effect of education is particularly problematic. Indeed, as argued by Card (1995, 1999), significant variation in the marginal causal effect can explain why IV estimates of the rate of return are usually noticeably higher than OLS estimates. Instruments that pick up exogenous variation in education near the middle of the education distribution (where the marginal causal effect is the highest) can be expected to yield estimates of the rate of return greater than OLS estimates.

References

Ashenfelter, Orley, Colm Harmon, and Hessel Oosterbeek (1999), "A Review of Estimates of the Schooling/Earnings Relationship, with Tests for Publication Bias," *Labour Economics* **6**: 453-470.

Belman, Dale, and John S. Heywood (1991), "Sheepskin Effects in the Returns to Education: An Examination of Women and Minorities," *Review of Economics and Statistics* **73**: 720-724.

Card, David (1995), "Earnings, Schooling, and Ability Revisited," *Research in Labor Economics* **14**: 23-48.

Card, David (1999), "The Causal Effect of Education on Earnings," in O. Ashenfelter and D. Card, eds., *Handbook of Labor Economics Vol. 3A*, Amsterdam, Elsevier Science.

Card, David, and Alan B. Krueger (1992), "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States," *Journal of Political Economy* **100**: 1-40.

Carneiro, Pedro, and James Heckman (2003), "Human Capital Policy," Working Paper 9495, Cambridge, MA, NBER.

Groot, Wim, and Hessel Oosterbeek (1994), "Earnings Effects of Different Components of Schooling; Human Capital versus Screening," *Review of Economics and Statistics* **76**: 317-321.

Harmon, Colm, Hessel Oosterbeek, and Ian Walker (2000), *The Returns to Education: A Review of Evidence, Issues and Deficiencies in the Literature*, London, Center for the Economics of Education.

Harmon, Colm, and Ian Walker (1999), "The Marginal and Average Returns to Schooling in the UK," *European Economic Review* **43**: 879-887.

Heckman, James, and Solomon Polachek (1974), "Empirical Evidence on the Functional Form of the Earnings-Schooling Relationship," *Journal of the American Statistical Association* **69**: 350-354.

Heckman, James J., Lance J. Lochner, and Petra E. Todd (2003), "Fifty Years of Mincer Earnings Functions," Working Paper 9732, Cambridge, MA, NBER.

Hungerford, Thomas, and Gary Solon (1987), "Sheepskin Effects in the Return to Education," *Review of Economic and Statistics* **69**: 175-177.

Jaeger, David A., and Marianne E. Page (1996), "Degrees Matter: New

Evidence on Sheepskin Effects in the Returns to Education," *Review of Economics and Statistics* **78**: 733-740.

Kroch, Eugene A., and Kriss Sjoblom (1994), "Schooling as Human Capital or a Signal," *Journal of Human Resources* **29**: 156-180.

Murphy, Kevin M., and Finis Welch (1990), "Empirical Age-Earnings Profiles," *Journal of Labor Economics* **8**: 202-229.

Mincer, Jacob (1974), *Schooling, Experience, and Earnings*, New York, Columbia University Press.

Psacharopoulos, George (1985), "Returns to Education: A Further International Update and Implications," *Journal of Human Resources* **20**: 583-604.

Psacharopoulos, George (1994), "Returns to Investment in Education: A Global Update," *World Development* **22**: 1325-1343.

Trostel, Philip, Ian Walker, and Paul Woolley (2002), "Estimates of the Economic Return to Schooling for 28 Countries," *Labour Economics* **9**: 1-16.

Weiss, Andrew (1995), "Human Capital vs. Signalling Explanations of Wages," *Journal of Economic Perspectives* **9**: 133-154.