Institute for the Study of Social Change

ISSC DISCUSSION PAPER SERIES

DISPERSION IN THE ECONOMIC RETURN TO SCHOOLING

Dr. Colm Harmon
Dr. Vincent Hogan
Prof. Ian Walker

Colm Harmon is a member of the Economics Department at University College Dublin and an associate at the Centre for Economic Policy Research (CEPR). Vincent Hogan is a member of the Economics Dept. in UCD. Ian Walker is an Economics Professor at Warwick University, an associate of the Institute for Fiscal Studies (IFS) and a Research Fellow at the Institute for the Study of Social Change (ISSC).

ISSC WP 2002/08

Dispersion in the Economic Return to Schooling*

Version 4.10 18th October 2002

Colm Harmon

University College Dublin, University College London, CEPR and IZA

Vincent Hogan

University College Dublin and University of Warwick

and

Ian Walker

University of Warwick, IFS and IZA

JEL Classification: J31, J24, C31

Abstract: We extend the standard human capital earnings function to include dispersion in the return to schooling by treating the return as a random coefficient. If the rapid expansion in participation in higher education has been brought about by dipping further into the ability distribution, we should observe a rise in the variance of returns. Alternatively, if the expansion has come about through relaxing credit constraints then we might expect to see an increase in both the mean and variance of returns. Our estimates suggest that the variance in returns has not risen over time.

* We are grateful to Yu Zhu for research support. Valuable comments were received from seminar participants at NUI Maynooth and at EALE 2002, Paris. Financial support which facilitated this research is acknowledged from the UK Department for Education and Skills (Walker), the EU Commission for the award of a Marie Curie Fellowship at the University of Warwick (Hogan) and to the Nuffield Foundation for the award of a New Career Development Scholarship in the Social Sciences at UCL (Harmon). This paper is part of Policy Evaluation programme of the Institute for the Study of Social Change, UCD. Data used with permission of the ESRC data archive (www.data-archive.ac.uk). The usual disclaimer applies. Corresponding author: Vincent Hogan, Department of Economics, University College Dublin, Belfield, Dublin 4, Ireland. Phone +353 1 716 8300, Fax +353 1 283 0068, Email vincent.hogan@ucd.ie

1. Introduction and Context

In estimating the standard model of human capital accumulation it is usual for the econometrician to assume that the return to schooling is constant across individuals. As we discuss below, however, there are good reasons to think that the true return to schooling may vary across individuals and that the degree of this dispersion may itself vary over time. This variability of returns is of independent interest, containing information useful for policy makers.

In this paper we extend the standard human capital earnings function (Mincer, 1974) to include dispersion in the rate of return to schooling. We allow the return to education estimated on a sample of UK data to vary across individuals by treating it as a random coefficient. Thus we estimate both the mean return and the variance around this mean and show how both vary through time.

We can think of two broad reasons why observed rates of returns may vary across individuals: risk and heterogeneity. Heterogeneity refers to differences in the returns to education across individuals due to factors unobservable to the econometrician, but known to the individual at the time of their decisions. More generally each individual faces a distribution of returns that is conditional (in its mean and possibly higher moments) on individual characteristics observable to them, but some of which are not observable by the econometrician. Thus even if each individual were to receive their expected return to education with absolute certainty, returns would still appear to vary randomly across the population.

In contrast, we use "risk" to refer to variation across states of nature. This variation results from individuals receiving good or bad draws from their (individual specific) wage distributions at a given point in time. This leads to variation in the returns to education due to factors unobserved by both the individual and the econometrician. Note that while both risk and heterogeneity will lead to dispersion in the measured economic returns to education across individuals, the policy implications would be radically different. One assumes that policy makers would be more concerned with genuine ex-ante individual differences than with than with differences due to temporary shocks in the labour market.

Much of the recent work that considers the dispersion in the returns to education does so from point of view of examining its implications for IV estimators. For example, Carneiro, Heckman and Vytlacil (2001) and Card (2001) note that using school reforms as instruments will produce estimates that are weighted averages of the returns to each individual with higher weight placed on those individuals most likely to have been affected by the reform. If returns vary across individuals, the resulting IV estimate will be a biased estimate of the average return to education and also a biased estimate of the return to education of the group affected by the reform. We argue below that ability bias and measurement error more or less cancel out so there is no advantage to IV. Furthermore even if the mean return is affected by ability bias, the variance need not be.

In this paper we focus on the dispersion in the returns to schooling as we see it as being of interest in its own right. Dispersion, and how it changes over time, has interesting implications for policy with a number of other economic phenomena. In recent years there has been an increase in the provision of advanced education places in many western countries. The U.K., for example, has a policy goal of providing a university place for fifty percent of school leavers. [2] Policies like this are clearly motivated by the existing research that estimates

[2] For details on the educational reforms in the UK see Harmon and Walker (1999) and the detailed analysis in Office of National Statistics (2001).

the return to schooling to be positive and to be quite high (see Card, 2001, for a survey). However, the existing research has focused on estimating the mean return to education. It is possible that the mean may have remained constant but that the variance of returns has changed over time. Thus there could be groups who have low returns to education and for whom expansion in education provision has not proved particularly productive.[3]

Why might the variance of returns have changed overtime? There are several scenarios. Firstly, in a standard human capital model, the expansion of education provision should have lead to an increase in the supply of skilled workers. Given demand, this would imply that the returns to skills would fall. The weight of evidence suggests that the there has been little tendency for the mean return to education to fall even where there has been a large expansion in post-compulsory education.[4] However, the increase in the supply of educated workers could have been brought about by dipping further into the lower tail of the ability distribution. If innate ability and human capital were complementary, then we should observe a rise in the variance of returns as more and more low ability individuals acquire education.

Similarly, if the increase in education supply occurred in response to an increase in the demand for skilled labour, it is possible that the mean return to education would remain unchanged reflecting roughly equal and offsetting increases in supply and demand. But the observed dispersion in returns would rise if education and ability were complements.

Changes in education finance could also affect the distribution of returns over time. For example, if credit constraints faced by poor individuals were relieved there is every

[3] These need not be low ability individuals. It is at least possible that formal education and innate ability are substitutes so that getting higher ability individuals into further education is unnecessary and wasteful.

[4] In the US see, in particular, Card and Lemieux (2001). In the UK see Dearden et al. (2000) and see Trostel, Walker and Woolley (2002) for an analysis of comparable data across 28 countries.

reason to believe that the mean return to education would rise, as more high ability but poor individuals continued in education. We might also expect the dispersion of returns to rise as reducing the cost of private borrowing to fund education would encourage those with low returns to undertake more education.

In this paper we conduct an examination of these issues by augmenting the standard human capital earnings function (Mincer, 1974) to include dispersion in the rate of return to schooling by treating the return to schooling as a random coefficient. Martins and Pereira (2002) use a quantile regression version of the Mincer equation to examine similar questions. They approximate the distribution of returns by (in their case) 10 discrete points. They then define the difference between the first and last deciles as their measure of dispersion and track how it varies across countries in a given year. They further identify their dispersion parameter with risk – but clearly their method would also pick up what we have called heterogeneity in returns. Chen (2002) uses US panel data (NLSY) to separate the dispersion in the returns to college into what she terms as "permanent" and "transitory" components which we interpret as being equivalent to heterogeneity and risk respectively. She finds that almost all the dispersion in returns is accounted for by the permanent component.

In this paper we remain agnostic as to the cause of the variation in returns, hence our use of the term "dispersion" rather than either "risk" or "heterogeneity".

2. Specification

We specify the basic Mincer-type earnings function as

$$\ln Y_i = (\beta + u_i)S_i + \gamma X_i + v_i \quad (1)$$

where Y is the log wage, **X** is a vector of explanatory variables, including a constant term and a quadratic in age to proxy for experience, S is years of schooling and ν is the usual residual assumed to be normally distributed with mean zero and variance σ^2. We explicitly allow the individual specific coefficient on schooling to be a have a normal distribution (independent of ν) so that β is the mean return to education and θ^2 is its variance. This is equivalent to

$$\ln Y_i = \beta S_i + \gamma X_i + \varepsilon_i \qquad (2)$$

where $\varepsilon_i = u_i S_i + \nu_i$, and $\Sigma_i = E(\varepsilon_i^2) = \sigma^2 + \theta^2 S_i^2$. Thus, equation (1) is a specific example of general heteroscedastic model and therefore the contribution of the i[th] observation to sample log-likelihood is given by (3)

$$\ln L_i = -0.5\ln(2\pi) - 0.5\ln(\Sigma_i) - 0.5\left(\frac{\varepsilon_i^2}{\Sigma_i}\right) \qquad (3)$$

At this juncture it is worth clarifying that the parameter θ is the standard deviation of the distribution of returns. However this is different from the sampling error associated with the estimate of β, which is itself the estimate of the mean of the distribution of returns. We can see this difference clearly from equation (3) where the parameters θ and β enter into the likelihood function, whereas the sampling error of β will be calculated from the estimated information matrix associated with (3). In order to avoid confusion in what follows, we refer to θ as "dispersion" and we refer to the sampling error of β as being the "standard error" of β.

Our estimation procedure does not control for any possible ability bias or measurement error.[5] There is evidence for the US, UK and elsewhere that using instrumental variables results in larger estimates than OLS. IV estimation, however, is becoming more

[5] Conneely and Uusitalo (1999) estimates a random coefficient model using Finish data that allows for endogenous schooling and ability bias.

controversial (see, for example, Card (2001) and Carneiro, Heckman and Vytlacil (2001) for a discussion of the precise interpretation of IV estimates). Furthermore Ashenfelter, Harmon and Oosterbeek (1999) suggested that much of the difference between IV and OLS estimates was the result of publication bias. Manski and Pepper (2000) also suggest that IV may be biased upwards. Griliches (1977) suggested that measurement error and ability bias could cancel each other out; leaving OLS estimates of the return to education approximately correct. The results of Angrist and Krueger (1991) seem to confirm this. More recently, Hogan and Rigobon (2002) also confirm Griliches' proposition for the UK using the same data as here.

Furthermore, our concern is with the dispersion of returns, and as Koop and Tobias (2002) show, this need not be affected by ability bias even if the mean return is affected.[6] Alternatively, if we are unwilling to accept the RC estimates as unbiased estimates of the return to education, one can simply treat them as giving the dispersion of wages conditional on education.

3. Data and Results

We use the UK Labour Force Survey from 1993-2000 to estimate the model outlined in Section 2. The LFS is close in design to the US CPS data. It is a large sample survey with a 5-quarter rotating panel design which has contained earnings information since 1993. Full details on the data and descriptive statistics are available on request. We select employees aged 25 to 59 with positive recorded hours and earnings and define the wage as hourly earnings.

[6] This will be true if the structural disturbances are uncorrelated with the random component of returns.

Table 1 presents results from OLS and random coefficient (RC) models for men and women in the pooled data. We control for years of schooling, a quadratic in age as a proxy for experience, birth cohort through a cubic function of the year of birth (we can discriminate between birth cohort and age because the data is pooled over eight successive years), marital status (married or cohabiting versus divorced, widowed, separated and never married), ethnic background (white versus non-white), union membership (member versus non-member) and health (an impediment to work). In addition to the direct control for years of schooling, in this specification we also include interactions of schooling with the other covariates to allow the return to schooling to vary as much as possible by observable characteristics.

The return to schooling from OLS is about 4% for men and 7% for women for the default individual but varies significantly with observable characteristics. If we take the interaction term at their average values, then the estimated (mean) return is 6.8% for men and 7.3% for women. These results change little when we use RC to estimate the return - the returns for women rise only slightly. Our estimate of the dispersion in the return to schooling is about 4% for men and 3.3% for women.

These estimates are in line with some of the existing work. Conneely and Uusitalo (1999) estimates a similar model on a sample of Finish men.[7] They find a mean return of 5% and a dispersion of 2%. The quantile regression estimates of Martins and Pereira (2002) would imply a mean return and dispersion for UK men of 7% and 1.9% respectively, if we assume that the returns are normally distributed. Koop and Tobias (2002) apply the model to

[7] Conneely and Uusitalo (1999) estimate an IV version of the random coefficients model using proximity to university and parental background as instruments.

the NLSY and find a mean return of 12% with a dispersion of 7%. Chen (2002) finds that the dispersion in returns to a US college education is 7%.

Our results suggest that 95% of men have returns in the +/- 7.8% interval around the mean, while the dispersion for women is lower with 95% of women within +/- 6.5% of the estimated mean. Thus the dispersion is large, even though we have allowed for differences by observable characteristics and it implies that a large number of people have very low returns to education. In fact these results imply that 4.5% of men and 1.7% of women will have negative returns to education. Similarly, 42% of men and 35% of women have education returns of less than 6% -- the discount rate typically applied to public investment projects in the UK. Taken at face value, these results imply that, while public investment in education is productive on average, for many people the returns do not appear to justify the costs. This possibility is missed when we focus solely on the average return to education.

One other result that is worth noting is the large difference in returns to education between non-whites and whites. Our random coefficients model suggest that a non-white man will experience a return to education almost 4% less than an otherwise identical white man. Thus wages are not just lower on average for non-whites, they are lower conditional on every level of education. This is clearly worrying and worthy of more detailed investigation that is beyond the scope of this paper.[8]

Table 1 includes year fixed effects but the year dummies are not interacted with schooling. To account for this interaction we estimate the model separately for each year. We

[8] Blackaby et. al. (2002) conducted a more detailed analysis of this issue using LFS data. Their results are consistent with ours as they find that returns to education for blacks that are half those of whites.

plot the resulting mean return (β) and the dispersion parameter (θ) in Figures 1 and 2, for men and women respectively. In each case the top half of the graph plots the OLS and RC return to schooling while in the bottom half of the graph the dispersion parameter is plotted together with the 95% confidence interval for this parameter.[9]

For men the OLS and RC returns to schooling differ by about 1% over the range of years with a slight, insignificant, upward trend in the return. The corresponding dispersion figure behaves quite erratically but varies only between 3% and 5% over the period. For women the returns behave quite differently with a downturn in the return to schooling, albeit insignificant, in the later period. In contrast the dispersion parameter is relatively stable over time between 3% and 4%.

These results show that there has been no discernable trend in the dispersion of returns to education in UK during the 1990s. This suggests that the increased supply of education places and/or reduction in credit constraints did not result in an increase in the number of low return individuals attaining higher levels of education.

Ideally we would like to corroborate this result we some evidence of the ability of students in general and of college entrants in particular. We would expect to see that the test scores of those staying on beyond the minimum school leaving age had remained constant over our sample period. Unfortunately, there is relatively little evidence on how students have performed in standardised tests in the UK over time. Nevertheless Mulls et. al. (2000) provide

[9] For clarity we do not include confidence intervals for the estimated mean returns to education. In any case they are very tight. The standard error of OLS estimate is less 0.35% in every year for both men and women and the standard error of the RC estimate is always less than 0.4%.

some information in this regard. They applied the same standard maths test to secondary school students in several countries over a period of several years. They report that average test score in England in 1995 and 1999 were almost identical. This does not prove that there has been no change in the proportion of less able individuals continuing in education. But it is at least consistent with our general result that there has been not been any discernible change in the distribution of returns to education over the period of the sample.

4. Conclusion

This paper is motivated by the concern that examination of the mean return to schooling may overlook important information about the nature of the returns. We estimate a simple random coefficients version of the Mincer wage equation using UK data. We find that, at 4%, the dispersion of returns (as measured by the standard deviation of returns across persons) is quite high, given a mean return of 7 %. Perhaps surprisingly, we find that there has been no significant increase in either the mean return to education or the dispersion of individual returns during the 1990s. This is true for men and for women. The absence of a trend in either the mean or variance of returns suggests that the increased supply of graduates has been absorbed by the labour market, implying that the demand for, and supply of, skills have increased roughly in tandem.

Furthermore, our results suggest that the major changes in the provision of education in the UK have not caused the education system to dip any further into lower tail of the distribution of returns. There does not appear to be more individuals with systematically lower (or higher) returns staying in education. Similarly, the reforms in education finance, that reduced the extent of credit constraints faced by some individuals, do not appear to have resulted in an increase in the number of low return individuals continuing in education.

References

Angrist, Joshua D. and Alan B Krueger, 1991. Does Compulsory School Attendance Affect Schooling and Earnings? Quarterly Journal of Economics, 106, 979-1014.

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, 453-70.

Blackaby, D. H., D. G. Leslie, P. D. Murphy and N. C. O'Leary, 2002. White/Ethnic Minority Earnings and Employment Differentials in Britain: Evidence from the LFS. Oxford Economic Papers, 54, 270-297.

Card, David, 2001. Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems. Econometrica, 69, 1127-1160.

Card, David and Thomas Lemieux, 2001. Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis. Quarterly Journal of Economics, 116, 705-46.

Carneiro, Pedro, James J. Heckman and Edward Vytlacil, 2001. Estimating the Return to Education When it Varies Among Individuals. Mimeo, Department of Economics, University of Chicago.

Chen, Stacey, 2002. Is Investing in College Education Risky? Mimeo, Department of Economics, SUNY, Albany.

Conneely, Karen and Roope Uusitalo, 1999. Estimating Heterogeneous Treatment Effects in the Becker Schooling Model. Mimeo, Department of Economics, University of

Helsinki.

Dearden, Lorraine, Steven McIntosh, Michael Myck and Anna Vignoles, 2000. The Returns to Academic and Vocational Qualifications in Britain. LSE Centre for the Economics of Education, Discussion Paper 4.

Griliches, Zvi, 1977. Estimating Returns to Schooling: Some Econometric Problems. Econometrica, 45, 1-22.

Harmon, Colm and Ian Walker, 1999. The Marginal and Average Returns to Schooling in the UK. European Economic Review; 43, 879-87.

Hogan, Vincent and Roberto Rigobon, 2002. Using Heteroscedasticity to Estimate the Returns to Education. NBER working paper, 9145.

Koop, Gary and James Tobias, 2002. Learning about Unobserved Heterogeneity in Returns to Schooling. Mimeo, Dept. of Economics, University of Glasgow.

Manski, Charles and John Pepper, 2000. Monotone Instrumental Variables: with an Application to the Returns to Schooling. Econometrica, 68, 997-1013.

Martins, Pedro and Pedro Pereira, 2002. Is There a Return-Risk Link in Education? Economics Letters, 75, 31-37.

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

Mulls, Ina, Michael Martin, Eugenio Gonzalez, Kelvin Gregory, Robert Garden, Kathleen O'Connor, Steven Chrostowski and Teresa Smith, 2000. TIMISS 1999: International Mathematics Report. The International Association for the Evaluation of Educational

Achievement and Boston College.

Office of National Statistics, 2001. Estimating Participation in Education, Training and Employment. Labour Market Trends, June.

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

Table 1 OLS and Random Coefficient Models (Pooled Annual Cross Sections)

	MEN				**WOMEN**			
	OLS		RC		OLS		RC	
	Coeff.	*Std. Error*	Coeff.	*Std. Error*	Coeff.	*Std. Error*	Coeff.	*Std. Error*
Constant	2.05	*0.01*	2.03	*0.02*	1.91	*0.01*	1.86	*0.03*
Years of schooling	4.03	*0.23*	4.13	*0.45*	7.43	*0.24*	8.62	*0.62*
AGE /100	0.66	*0.19*	0.36	*0.33*	1.09	*0.18*	0.67	*0.32*
$AGE^2/10^4$	-5.88	*0.75*	-6.25	*1.32*	0.80	*0.76*	1.07	*1.31*
COHORT /100	0.06	*0.20*	-0.34	*0.36*	0.84	*0.19*	0.34	*0.34*
$COHORT^2/10^4$	1.46	*0.70*	2.07	*1.24*	0.33	*0.71*	0.33	*1.21*
$COHORT^3/10^5$	1.74	*0.29*	1.95	*0.52*	1.95	*0.28*	1.86	*0.51*
MARRIED	0.13	*0.01*	0.11	*0.01*	-0.05	*0.01*	-0.04	*0.01*
COHABITATING	0.08	*0.01*	0.09	*0.02*	0.02	*0.01*	0.03	*0.02*
NONWHITE	-0.18	*0.02*	-0.15	*0.03*	0.02	*0.02*	0.01	*0.03*
HEALTH	-0.14	*0.01*	-0.15	*0.02*	-0.11	*0.01*	-0.09	*0.02*
UNION	0.14	*0.00*	0.13	*0.01*	0.21	*0.00*	0.20	*0.02*
Yrs of Schooling*AGE /100	0.22	*0.03*	0.22	*0.07*	-0.01	*0.04*	0.04	*0.07*
Yrs of Schooling $*AGE^2/10^2$	-0.43	*0.19*	-0.27	*0.39*	-0.96	*0.20*	-1.45	*0.37*
Yrs of Schooling *COHORT/100	0.19	*0.04*	0.17	*0.08*	0.03	*0.04*	0.08	*0.08*
Yrs of Schooling $*COHORT^2/10^3$	-0.15	*0.17*	-0.46	*0.37*	-0.08	*0.19*	0.29	*0.34*
Yrs of Schooling $*COHORT^3/10^4$	-0.48	*0.07*	-0.32	*0.15*	-0.76	*0.08*	-0.68	*0.15*
Yrs of Schooling *MARRIED/100	0.53	*0.16*	1.24	*0.39*	0.50	*0.16*	0.37	*0.30*
Yrs of Schooling *COHAB/100	0.02	*0.24*	-0.09	*0.48*	-0.04	*0.25*	-0.16	*0.43*
Yrs of Schooling *NONWHITE/100	-2.45	*0.27*	-3.82	*0.59*	-3.88	*0.30*	-4.07	*0.66*
Yrs of Schooling *HEALTH/100	-0.31	*0.26*	0.31	*0.54*	-0.32	*0.27*	-0.86	*0.52*
Yrs of Schooling *UNION/100	1.90	*0.11*	1.95	*0.22*	1.01	*0.12*	0.58	*0.37*
Dispersion - θ	-		4.22	*0.42*	-		3.37	*0.34*
Sample Size	76,722		76,722		81,508		81,508	
R^2	0.22				0.27			

Note: regressions also include controls for region and year of sample.

Figure 1: Year-on-Year Estimates of the Return to Schooling for Men: OLS and RC

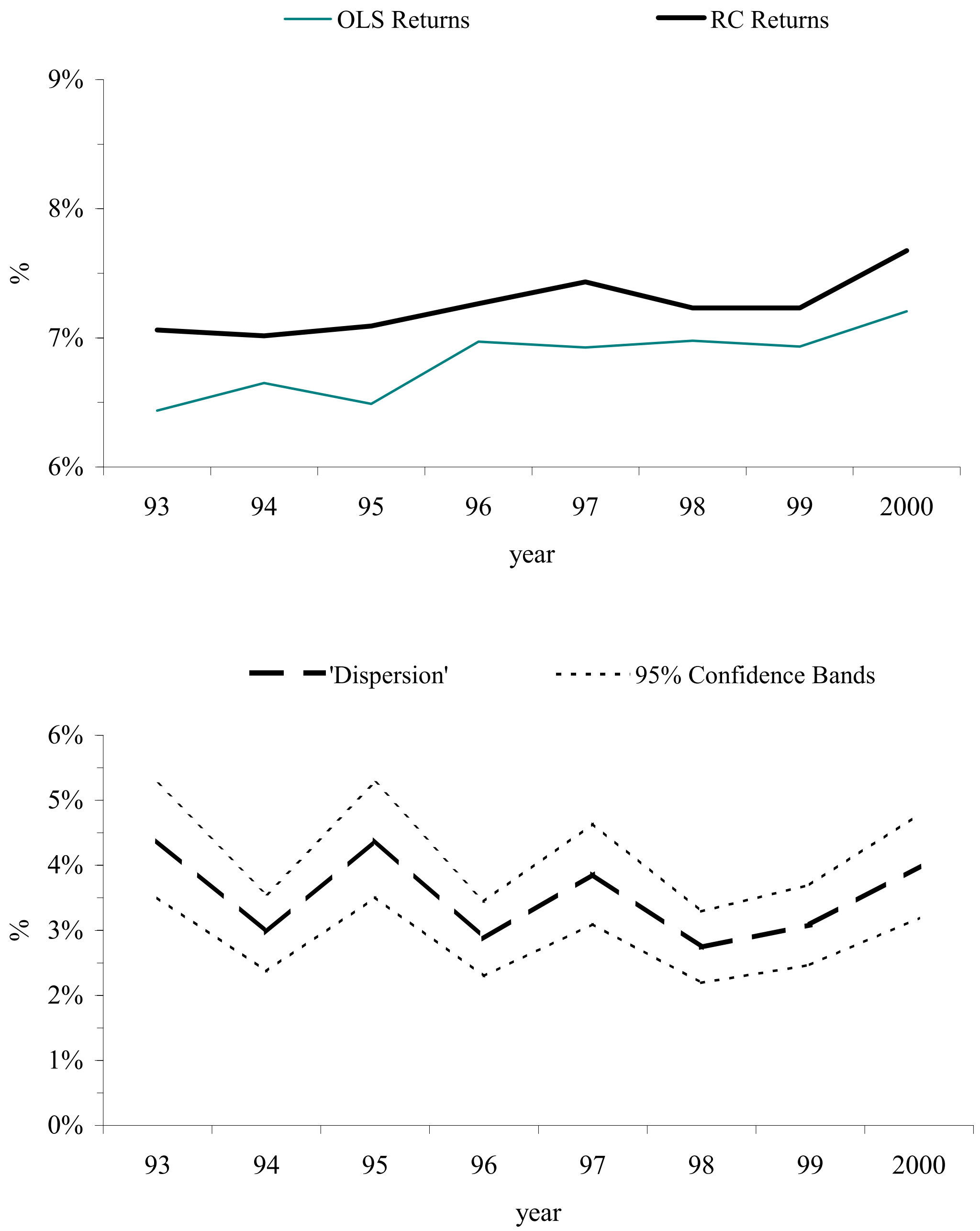

Figure 2: Year-on-Year Estimates of the Return to Schooling for Women: OLS and RC