RESEARCH

The Returns to Education: Evidence from the Labour Force Surveys

Ian Walker And Yu Zhu

Department of Economics
University of Warwick, Coventry CV4 7AL

Research Report RR313

department for
education and skills
creating opportunity, releasing potential, achieving excellence

The Returns to Education: Evidence from the Labour Force Surveys

Ian Walker and Yu Zhu

Department of Economics
University of Warwick, Coventry CV4 7AL

Date: 26 November 2001

Email: i.walker@warwick.ac.uk
Fax: 024765 523032
Tel: 024765 523054

List of Contents

List of figures		ii
List of tables		iii
Abstract		iv
Acknowledgements		v
1. Introduction		1
2. Review of the Theory and Methodological Issues		4
3. Review of Previous Empirical Evidence		11
3.1	International evidence	11
3.2	UK Evidence	19
3.3	A Summary of Existing Evidence	24
4.	Evidence from the LFS Data	26
5.	The Return to Qualifications	31
6.	The Variance in the Return to Education	39
6.1	Quantile Regression Results	39
6.2	Random Coefficient Results	44
7.	Conclusion	51
References		53

List of figures

Figure1	Returns to Schooling – A Meta Analysis	25
Figure 2	OLS Returns to Education Years – Men and Women	29
Figure 3	OLS Returns to Education Years – Men and Women	29
Figure 4	OLS Returns to Qualifications Relative to None – Women	33
Figure 5	OLS Returns to Degrees by subject – Women	34
Figure 6	OLS Returns to Qualifications Relative to None – Men	36
Figure 7	OLS Returns to Degree by Subject – Men	37
Figure 8	Quantile Regressions	41
Figure 9	Year-on-Year Estimates of the Return to Schooling for Men: OLS and RC	47
Figure 10	Year-on-Year Estimates of the Return to Schooling for Women: OLS and RC	48
Figure 11	Distribution of Returns – Men	50
Figure 12	Distribution of Returns – Women	50

List of tables

Table 1	Returns to Years of Schooling in Europe (year closest to 1995).	11
Table 2	OLS Estimates of Returns to Years of Education: ISSP Data	13
Table 3	OLS Estimates of the Trend in the Return to Years of Education: ISSP Data	15
Table 4	IV Estimates of Returns to Years of Education using Spouse's Education: ISSP Data	17
Table 5	IV Estimates of Return to Years of Schooling using Father's Education: ISSP Data	18
Table 6	IV Estimates of Return to Years of Schooling using Mother's Education : ISSP Data	18
Table 7	Summary of Previous Specifications	20
Table 8	Estimated returns to years of schooling	23
Table 9	Estimated Returns To Qualifications	23
Table 10	OLS Returns to Education Years - Women	28
Table 11	OLS Returns to Education Years - Men	28
Table 12	OLS Returns to Qualifications – Women	32
Table 13	OLS Returns to Degree by Subject – Women	33
Table 14	OLS Returns to Qualifications – Men	35
Table 15	OLS Returns to Qualifications – Men	36
Table 16	Quantile Regressions: Various Countries	41
Table 17	Quantile Regressions: Return to Year of Education - Women	43
Table 18	Quantile Regressions: Returns to Degree vs 2+ A Levels – Women	43
Table 19	Quantile Regressions: Return to Year of Education - Men	43
Table 20	Quantile Regressions: Returns to Degree vs 2+ A Levels – Men	43
Table 21	OLS and Random Coefficient Models	46

Abstract

This report details research on the relationship between education and wages. The work is largely based on the Labour Force Surveys 1992-2000 and the focus of the research is largely on academic education.

The report contains microeconometric estimates of the relationship between (log) wages and years of education and allows: for this relationship to be non-linear, so as to separately identify the effect of higher levels of education from the effect of earlier years; for the relationship to shift over time so we provide estimates that show how returns to education vary over time; and for the relationship to vary across individuals according to their observable and unobservable differences. Separate results for men and women are presented.

The LFS data is large and this enables separate analysis of particular groups of individuals. In particular, the report contains microeconometric estimates of the effects of a degree on wages and allows for: different degree subjects to have differential effects; "sheepskin" effects associated with years of education that yield a qualification; and different lengths of study. Separate results for men and women are presented.

In addition to estimating the *mean* effect of education on wages we also estimate the *variance* in returns around this mean. There are two complementary ways in which we pursue this. In the first method, estimation is by "quantile regression" methods to estimate the effect on different parts of the wage distribution. We are particularly concerned to show the extent to which the returns to education differ across the wage distribution. If the average ability of graduates has fallen over time then we might see this reflected in the size of the returns across quartiles of the wage distribution. The second method estimates a "random coefficients" model. Instead of assuming that the effect of education is the same for all individuals this model assumes that the effect differs (randomly) across individuals. The model estimates the mean effect of education and the variance around this mean. Again, by estimating the models for each separate year it is possible to see if the variance is getting larger over time. The modelling controls for observable differences in returns across individuals.

Acknowledgements

This work was supported by DfES Analytical Services but the views expressed are those of the authors alone. We are grateful to David Thompson of DfES Analytical Services for his advice and careful work that improved this report immeasurably.

We are also grateful to Vincent Hogan of University College Dublin and Philip Trostel of the University of Maine. Our work with them on this topic was supported through DfES funding and is cited here.

1. Introduction

This report is concerned with the effect of education on earnings and, in particular, with the financial return to education. The methodology used is to estimate, using large survey datasets that contain information on education, earnings and other characteristics, the relationship between (log) wages and education. The work largely uses straightforward regression methods to estimate coefficients that pick up the effect of either a year of education or the possession of a specific qualification (for example, a degree) on (log) wages[1]. The empirical specifications are based on the theory of human capital, whereby individuals make decisions on acquiring human capital, such as education, up to the point where the returns to education are driven down to the real return on other assets. With additional assumptions, such a framework implies that (log) wages are linearly related to education (or qualifications) and a quadratic function of work experience. Furthermore, the coefficient on education in such a model (i.e. the effect of a year of education on wages) can be interpreted as the financial rate of return to education providing the only costs of education are the opportunity costs of the forgone earnings. That is, the coefficient that we estimate is a measure both of the effect of a year of education on individual wages **and** the financial return to an individual investing in his/her human capital. Indeed, our analysis suggests that the returns estimated here are not very sensitive to whether we include the (modest) real costs that are associated with education – for example, including the recently introduced fees for higher education into a financial rate of return calculation makes little difference to the impression given by the coefficients that we estimate[2].

Our aim is to investigate not just the **level** of this financial return but also try to show **how it varies across individuals** and **across time**. Our motivation for being interested in the variance in returns across individuals and across time is to investigate whether the large increases in education participation that have occurred in recent years has been reflected in both the average return (because, other things equal, an increase in the supply of educated workers relative to less educated workers will decrease the wages of the former relative to the latter) **and** in the variance in the

[1] The work complements Dearden *et al* (2000).

[2] That is, typical fees lie within the range of costs that has little effect on the financial rate of return calculation.

return across individuals (perhaps because increased participation might be among those individuals who expect to gain lower returns).

In some of the research, we focus on education as measured by years of education. This summary measure of education, while crude, has the virtue that it allows us to concentrate on particular issues. In other work, we look explicitly at the returns to academic qualifications – especially of higher education degrees compared to A-levels. In this line of our research we are concerned to see how much of the heterogeneity in returns can be attributed to the heterogeneity in the education that individuals receive as their academic and/or vocational training proceeds, and on the subject matter of the degree.

Part of the variance across individuals might arise because individuals are different in **observable** ways – the return may be higher for women than men, for nonwhites than whites, etc. Thus we allow for the returns to differ according to observable characteristics. But we also investigate the extent to which returns vary for **unobserved** reasons. We investigate the variance in returns as well as the level of returns and, in particular, we estimate how these have changed over time. How the variance in returns might have changed is investigated using two complementary methods that generalise the elementary regression method that yields only estimates of the average return. In the first method, estimation is by "quantile regression" methods which estimates the effect of education on individuals at different parts (for example, quantiles) of the wage distribution. We are particularly concerned to show the extent to which, if the average ability of more highly educated individuals has fallen over time because education expansion has moved the marginal student further down the ability distribution, the returns to education rise or falls as we move up the quartiles of the wage distribution. That is, we estimate the returns to education for each quantile of the wage distribution and see whether the returns to individuals in different quantiles have changed differentially over time. The second method estimates a "random coefficients" model. Instead of assuming that the effect of education is the same for all individuals this model assumes that the effect differs (randomly) across individuals. The model estimates both the mean effect of education and the variance around this mean. Again, by estimating the models for each separate year it is possible to see if the variance is getting larger over time or not. Such modelling can include controls for observable differences in returns across individuals: for example men vs. women.

For much of the work the data used here is the Labour Force Surveys. We use this data, where earnings information is only available from 1992, because it is large and all of the hypotheses that we wish to address require us to divide the data into groups – sometimes groups who are quite a small proportion of the total. Thus, it is only with a large dataset that it is possible to get robust estimates of the effects of characteristics that are present in just a small proportion of workers. However, LFS has its drawbacks. In particular, years of education in the LFS can only be inferred from the age at which individuals left full-time **continuous** education. Thus, the data fails to record years of education accurately for those that interrupted their schooling. We attempt to identify such individuals in the data and to then control for interrupted schooling in our analysis.

We also spend some time making cross-country comparisons and here we rely on new data from the International Social Science Project (ISSP), which is comparable across countries. Throughout we try to place our new work in the context of the existing literature.

The structure of the report is as follows. Section 2 contains a brief summary of the theoretical framework and highlights the main methodological "issues" that we feel are important. Section 3 briefly reviews existing empirical research and includes evidence from the UK, US and a variety of other countries. Section 4 provides the new estimates from the LFS data for a simple model which shows, in broad terms, estimates that are broadly comparable with research on other datasets. Section 5 expands on this simple specification and incorporates information about qualifications obtained and, in particular, this section exploits information about degree subject that LFS contains. Finally, section 6 presents the results from the two methodologies that attempt to identify the extent to which the returns differ across individuals.

The broad conclusions from our analysis is that: despite the expansion in higher education that has occurred in recent years, there is no significant trend in the average return to education which remains high by international standards; and there is a large variance in returns around the average, but the unexplained component of this variance has not changed over time.

2. Review of the Theory and Methodological Issues

The analysis of the demand for education has been driven by the concept of human capital and has been pioneered by Gary Becker, Jacob Mincer and Theodore Schultz. The classic reference is Becker (1964). In human capital theory education is an investment of current resources (the opportunity cost of the time involved as well as any direct costs) in exchange for future returns. The benchmark model for the development of empirical estimation of the returns to education is the key relationship derived by Mincer (1974). The typical human capital theory (Becker (1964)) assumes that education, s, is chosen to maximise the expected present value of the stream of future incomes, from when work starts at date s+1 up to retirement at date T, net of the costs of education, c_s. So, at the optimum s, the present value of the s^{th} year of schooling just equals the costs of the s^{th} year of education, so equilibrium is characterised by: $\sum_{t=1}^{T-s} \frac{w_s - w_{s-1}}{\left(1+r_s\right)^t} = w_{s-1} + c_s$ where r_s is called the internal rate of return (we are assuming that s is infinitely divisible, for simplicity, so "year" should not be interpreted literally). Optimal investment decision-making would imply that one would invest in the s^{th} year of schooling if $r_s > i$, the market rate of interest. If T is large then the left hand side of the equilibrium expression can be approximated so that the equilibrium condition becomes $\frac{w_s - w_{s-1}}{r_s} = w_{s-1} + c_s$. Then, if c_s is sufficiently small, we can rearrange this expression to give $r_s \approx \frac{w_s - w_{s-1}}{w_{s-1}} \approx \log w_s - \log w_{s-1}$ (where $\approx$ means approximately equal to). This says that the return to the s^{th} year of schooling is approximately the difference in log wages between leaving at s and at s-1. Thus, one could estimate the returns to S by seeing how *log* wages varies with S. That is, the empirical approximation of the human capital theoretical framework is the familiar functional form of the earnings equation:

$$\log w_i = \mathbf{X_i}\beta + rS_i + \delta x_i + \gamma x_i^2 + u_i ,$$

where w_i is an earnings measure for an individual i such as earnings per hour or week, S_i represents a measure of their schooling, x_i is an experience measure (typically age minus age left education), $\mathbf{X_i}$ is a set of other variables assumed to affect earnings, and u_i is a disturbance term representing other unobservable factors which are not be explicitly measured, assumed independent of $\mathbf{X_i}$ and S_i. Note that experience is included as a quadratic term to capture the concavity of the earnings profile. Mincer's derivation of the empirical model implies that, under the assumptions made (particularly, the assumption, no tuition costs), r can be considered the private financial return to schooling as well as being the proportionate effect on wages of an increment to S.

The availability of microdata and the ease of estimation has resulted in many studies, which essentially estimate the simple Mincer specification. In the original study Mincer (1974) used 1960 US Census data and used an experience measure known as potential experience (i.e. current age *minus* age left full time schooling) and found that the returns to schooling were 10% with returns to experience of around 8%. In the earliest UK study, Layard and Psacharopolous (1979) used the GB General Household Survey (GHS) 1972 data and found returns to schooling of a similar level, around 10%. See Willis (1986) and Psacharopolous (1994) for many more examples of this simple specification.

It is useful, at this point, to consider the implications of *endogenous* schooling. In the human capital framework, on which the original Mincer work was based, schooling is an optimising investment decision based on future earnings and current costs: that is, on the (discounted) difference in earnings between undertaking and not undertaking a unit of education and the total cost of that unit of education including foregone earnings. Investment in education continues until the difference between the marginal cost and marginal return to education is zero.

A number of implications stem from considering schooling as an investment decision. Firstly, the internal rate of return (IRR, or r in this work) is the discount rate that equates the present value of benefits to the present value of costs. More specifically if r is greater than the market rate of interest then more education is a worthwhile investment for the individual. In making an investment decision an individual who places more (less) value on current income than future income streams will have a higher (lower) value for the discount rates so individuals with high discount rates (high r_i) are therefore *less* likely to undertake education[3]. Secondly,

[3] Thus the model implies that early schooling has a greater return than schooling later in life since there are fewer periods left to recoup the costs. The corollary of this is that schooling should precede work.

direct education costs (c_s) lower the net benefits of schooling. Finally, if the probability of being in employment were higher if more schooling is undertaken then an increase in unemployment benefit would erode the reward from undertaking education. However, should the earnings gap between educated and non-educated individuals widen or if the income received while in schooling should rise (say, through a tuition subsidy or maintenance grant) the net effect on the incentive to invest in schooling should be positive.

Clearly in this empirical implementation the schooling measure is treated as *exogenous*, although education is clearly an *endogenous* choice variable in the underlying human capital theory.

Moreover, in the Mincer specification the disturbance term captures unobservable individual effects and these individual factors may also influence the schooling decision, and hence induce a correlation between schooling and the error term in the earnings function. A common example is unobserved ability: ability is correlated with wages and with schooling but is not (usually) observed. Thus, a useful extension to the theory is to consider the role of the individual's ability on the schooling decision, whilst preserving the basic idea of schooling being an investment. Griliches (1977) introduces ability (A) explicitly into the derivation of the log-linear earnings function. In the basic model the r of schooling is partly determined by foregone income (less any subsidy such as parental contributions) and any educational costs. Introducing ability differences has two effects on this basic calculus. The more able individuals may be able to 'convert' schooling into human capital more efficiently[4] than the less able, and this raises r for the more able. One might think of this as inherent ability and education being complementary factors in producing human capital so that, for a given increment to schooling, a larger endowment of ability generates more human capital. On the other hand, the more able may have higher opportunity costs since they may have been able to earn more in the labour market, if ability to progress in school is positively correlated with the ability to earn, and this reduces the r. The net effect can therefore be ambiguous.

[4] In the Griliches model there is a subtle extension often overlooked but highlighted by Card (1994). There can exist a negative relationship between optimal schooling and the disturbance term in the earnings function by assuming the presence of a second unmeasured factor (call this energy or motivation) that increases income and by association foregone earnings while at school, but is otherwise unrelated to schooling costs.

This problem has been a preoccupation of the literature since the earliest contributions - if schooling is endogenous then estimation by least squares methods will yield biased estimates of the return to schooling. That is, if the omitted ability variable is correlated with *S* then it's least squares coefficient will be biased because *S* picks up the effect of ability as well as schooling. Since it has been reasonable to presume that *A* and *S* are positively correlated, as is the correlation between *A* and *w,* then least squares is biased upwards – that is, least squares is an **upper** bound on the true return[5].

However, it should not be assumed that least squares estimates would **necessarily** be biased upwards. There may be other unobservable (omitted) variables that affect both *S* and *w* but in **opposite** directions and this would result in least squares being biased downwards. For example, "impatient" individuals could be construed as having a high personal "rate of time preference" and so apply a high discount rate to benefits in the future. Such individuals would find it optimal to choose a low value of *S.* At the same time, impatience may be virtue that is rewarded in the labour market through higher *w* since individuals blessed with this characteristic may be perceived as having the drive to complete tasks, meet deadlines, etc.

Thus, the question is ultimately an empirical one and there have been a number of approaches put forward to deal with this problem. Firstly, measures of ability have been incorporated to proxy for unobserved effects. The inclusion of direct measures of ability should reduce the estimated education coefficient if it acts as a proxy for ability, so that the coefficient on education then captures the effect of education alone since ability is controlled for. Secondly one might exploit within-twins or within-siblings differences in wages and education if one were prepared to accept the assumption that unobserved effects are additive and common within twins (or other similar pairs) so that they can be differenced out by regressing the wage difference within twins against the education difference. This approach is a

[5] In addition, Griliches (1977) has shown that bias due to measurement error is necessarily negative, i.e. that least squares is biased downwards if *S* is measured with error. However, since *S* is normally quite

modification of a more general fixed effect framework using individual panel data, where the unobserved individual effect is considered time-invariant. Unlike panel data where schooling is a time invariant regressor which is removed by differencing, in twins data differencing leaves the differences in schooling in the model. But good data on twins is scarce and a more common approach deals directly with the schooling/earnings relationship in a two-equation system by exploiting instrumental variables that affect *S* but not *w*. To address this endogeneity bias in the absence of data on twins we need to instrument schooling by purging its correlation with unobservable influences on wages, using variables that are correlated with schooling but not with wage rates. Namely, the instrument needs to be orthogonal to the unobserved component of the wage equation error term that is correlated with schooling; i.e. the term that captures the individual's ability or discount rate. Such a joint model would be $y_i = \mathbf{X}_i'\alpha + \beta S_i + u_i,\ S_i = \mathbf{Z}_i{}'\delta + v_i$ where $\mathbf{Z}_i$ is the vector of observed instrumental variables with the properties suggested above and v captures variation in *S* across individuals that arises for unobservable reasons.

The empirical implications of this extension to the basic theory are most clearly outlined in Card (1999). This work embodies the usual idea that the optimal schooling level equates the marginal rate of return to additional schooling with the marginal cost of this additional schooling. Card (1999) allows the optimal schooling to vary across individuals for a further reason: not only can different returns to schooling arise from variation in ability, so that those of higher ability 'gain' more from additional schooling, but individuals may also have different marginal rates of substitution between current and future earnings. That is, there may be some variation in the discount rate across individuals. This variation in discount rates may come for example from variation in access to funds or taste for schooling.

If ability levels are similar across individuals then the effects are relatively unambiguous - lower discount rate individuals choose more schooling. However, one might expect a negative correlation between these two elements: high-ability parents, who would typically be wealthier, will tend to be able to offer more to their children in terms of resources for education. Moreover high education parents will

accurately measured this is probably not a great source of bias. See Dearden (1998) for UK estimates that confirm this.

have stronger tastes for schooling (or lower discount rates) and their children may "inherit" some of this. Indeed, if ability is partly inherited then children with higher ability may be more likely that average to have lower discount rates. The reverse is true for children of lower ability parents. This bias will be determined by the variance in ability relative to the variance in discount rates as well as the covariance between them. This "endogeneity" bias arises because people with higher marginal returns to education choose higher levels of schooling. If there is no discount rate variance then the endogeneity will arise solely from the correlation between ability and education and since this is likely to be positive the bias in OLS estimates will be upwards (if ability increases wages later in life more than it increases wages early in life). If there is no ability variance, then the endogeneity arises solely from the (negative) correlation between discount rates and OLS will be biased downwards if discount rates and wages are positively correlated (for example, if ambitious people earn higher wages and are more impatient). Thus, the direction of bias in OLS estimates of the returns to education is unclear and is, ultimately, an empirical question.

A parallel literature to the IV, as mentioned above, attempts to control for the unobservable determinants of schooling using data on twins[6] - identical twins have the advantage, relative to other siblings, of being genetically identical so that twins data are an obvious way test the argument that genetics determines economic success (see Herrstein and Murray (1994)). The twins methodology in recent research follows Ashenfelter and Krueger's (1994) innovation of asking one twin to report on the schooling of the other, in order to examine possible measurement error.

The methodology relies on differences in education between twins being random. Thus the correlation between differences in wages and differences in education reveals the effect of random variation in education on wages. However, Bound and Solon (1999) argue that whilst within pair differencing removes genetic variation, the differences in schooling might still reflect ability bias to the extent that ability is affected by more than just genes.

This begs the question - what causes the differences in schooling between identical twins? Ashenfelter and Rouse (1998), Bound and Solon (1999) and Neumark

[6] See also Ashenfelter and Zimmerman (1997) for a study based on samples of pairs of (non-twin) brothers and of father-son pairs.

(1999), following earlier arguments due to Griliches (1979) debate this at length in recent papers. As Bound and Solon (1999) point out, conventional OLS ability bias depends on the fraction of variance in schooling that is accounted for by variance in unobserved abilities that might also affect wages. Similarly, within pair ability bias depends on the fraction of within pair variance in schooling that is accounted for by within pair variance in unobserved abilities that also affect wages. Thus the within pair bias will be smaller if the endogenous variation within families is smaller than the endogenous variation between families. While plausible, Bound and Solon (1999) argue that there is no reason to suppose this is the case.

Ultimately the matter is of course an empirical one. Ashenfelter and Rouse (1998) present evidence that differences in schooling within-twin pairs are uncorrelated with birth order and a range of characteristics such as union status, self-employment, tenure and spouse's education. They therefore argue that within pair education differences are primarily due to random factors (luck, optimisation error) and not ability. But, they do find significant correlations between average levels of family education and characteristics. To the extent that these correlations capture differences in ability they therefore argue that most of the variation in ability is between families and not between twins within a family.

While these issues are still under debate recent research seems to come down in favour of IV estimation where instruments are drawn from education reforms that have occurred over time that affect some groups and not others – that is, where the instruments are generated by some "natural" experiment. Card (2001) gives a review of this, largely US, research. The only UK examples are Harmon and Walker (1995, 1999) which suggests that OLS suffers from significant downward bias.

Finally, it is worth considering for a moment the interpretation of the estimated returns under alternative estimation methods. Card argues that IV estimates based on reforms exceed OLS because IV estimates the return *for those individuals who are induced to stay on at school because of the reform*. These will typically be individuals with little schooling. If such individuals have high returns then the reason why they were not staying on at school is that they faced high costs. In contrast, IV estimates based on family background, which also tend to exceed OLS but by a smaller margin, are invalid because background is a proxy for ability and this affects wages directly as well as schooling.

3. Review of Previous Empirical Evidence

3.1 International evidence

We begin by presenting international evidence: estimates of rates of return to education in many countries obtained using comparable data and specifications. Our aim here is to investigate how widespread are the two important features of the recent, largely UK and US, literature: the rising rate of return to education that has been detected in both the UK and the US in data from the mid 1970's, and the finding that least squares estimates are biased downwards rather than, as had commonly been thought more plausible, upwards. The work referred to here complements and updates Psacharopoulos's (1994) summary of the evidence on rates of return to schooling across countries.

Firstly, in Harmon *et al* (2001) a number of studies are collected containing the findings from the EU funded "PURE" project that studied the returns to education across 15 European countries. The project attempted to use comparable specifications with data from the same time period, even though there were some differences in definitions. Table 1 shows their results for one simple comparable specification. The coefficients reported are all highly statistically significant and it is clear that the UK has amongst the highest estimated return.

Table 1 Returns to Years of Schooling in Europe (year closest to 1995).

	Men	Women
Denmark (95)	0.061	0.043
Germany (West) (95)	0.077	0.095
Netherlands (96)	0.057	0.042
Portugal (94)(95)	0.100	0.104
Sweden (91)	0.041	0.037
UK (94-96)	**0.096**	**0.122**
Ireland (94)	0.088	0.129
Italy (95)	0.058	0.070
Norway	0.045	0.047
Finland (93)	0.085	0.087
Spain (94)	0.069	0.079
Switzerland (95)	0.089	0.089
Mean	0.072	0.079

Source: Harmon *et al* (2001).

Secondly, Trostel *et* al (2000) is a 28-country study that has the added benefit of strictly comparable data collected as part of the International Social Survey Programme (ISSP). These data were collected in each of a large number of countries using a common questionnaire. National cross-sectional surveys are pooled across (up to) eleven years from 1985 to 1995. The dependant variable is the logarithm of hourly earnings, computed as weekly earnings divided by the number of hours usually worked per week. In some of the countries in some years, however, the weekly earnings variable is only observed to fall within particular intervals on a continuous scale. In these cases we use interval midpoints for weekly earnings[7]. The estimation does not correct for selectivity into employment since the literature has not been able to show that, despite the correlation between education and employment, this has important consequences for the rate of return. Estimates for males and females in each country are obtained using the conventional Mincer (1974) model of earnings (the human capital earnings function), which has log wage rates determined by years of schooling, age or experience and other explanatory variables: $y_i = \mathbf{X}_i'\alpha + \beta S_i + u_i$, where y_i is the log of hourly wages, S_i is years of schooling and $\mathbf{X}_i$ is a vector of observed exogenous explanatory variables including controls for age or experience and, where appropriate, country and year fixed effects. $\beta_{.}$ is interpreted as the return to schooling; the percentage change in wages due to an additional year of schooling.

This work complements the meta-analysis of Ashenfelter *et al.* (1999), which uses statistical criteria to pool previous findings and subjects that earlier work to a variety of tests designed to explain differences in results and uncover evidence of publication bias. A meta-analysis combines and integrates the results of several studies that share a common aspect so as to be 'combinable' in a statistical manner. The methodology is typical in the clinical trials in the medical literature. In its simplest form the computation of the average return across a number of studies is now achieved by weighting the contribution of an individual study to the average on the basis of the standard error of the estimate (see Ashenfelter, Harmon and Oosterbeek (1999) for further details) – so that studies that feature more precise estimates get a higher weight. Here, because we publish exactly the same specifications for each country using comparable data, the prospect of inducing bias through deliberate (or

[7] We extrapolate values for the top open-ended group. Our estimates, however, are not sensitive to this choice.

even inadvertent) decisions is unlikely. This work also complements the recent, and largely US, research surveyed by Card (1999) that is concerned with the sensitivity of instrumental-variable estimates to the choice of instruments. Thus it addresses the difficulty in applying meta-analysis methods to establish the extent of bias in ordinary-least-squares estimates when earlier studies used a variety of instruments. Since the work uses comparable data it is able to use consistent instruments across countries. Table 2 presents conventional estimates of Mincerian human capital wage functions: that is, OLS estimates of rates of return to schooling.

Table 2 *OLS Estimates of Returns to Years of Education: ISSP Data*

Country	Males		Females	
USA	0.074	*0.004*	0.096	*0.005*
Great Britain	0.127	*0.006*	0.130	*0.006*
West Germany	0.036	*0.002*	0.043	*0.004*
Russia	0.044	*0.004*	0.053	*0.004*
Norway	0.023	*0.002*	0.025	*0.003*
Australia	0.051	*0.004*	0.052	*0.006*
Netherlands	0.031	*0.002*	0.019	*0.004*
Austria	0.038	*0.004*	0.064	*0.006*
Poland	0.073	*0.005*	0.100	*0.005*
East Germany	0.026	*0.003*	0.045	*0.004*
New Zealand	0.033	*0.004*	0.029	*0.005*
Italy	0.037	*0.003*	0.053	*0.005*
Ireland	0.085	*0.006*	0.090	*0.008*
Japan	0.075	*0.007*	0.094	*0.014*
Hungary	0.075	*0.007*	0.077	*0.006*
N. Ireland	0.174	*0.011*	0.146	*0.011*
Sweden	0.024	*0.004*	0.033	*0.005*
Slovenia	0.080	*0.007*	0.101	*0.007*
Israel	0.053	*0.007*	0.061	*0.008*
Czech Rep.	0.035	*0.007*	0.043	*0.007*
Bulgaria	0.040	*0.009*	0.057	*0.010*
Slovak Rep.	0.052	*0.012*	0.064	*0.009*
Canada	0.038	*0.008*	0.045	*0.008*
Czechoslovakia	0.031	*0.010*	0.036	*0.007*
Spain	0.046	*0.005*	0.038	*0.010*
Switzerland	0.045	*0.007*	0.048	*0.012*
Latvia	0.067	*0.020*	0.078	*0.014*
Philippines	0.113	*0.015*	0.192	*0.030*
Pooled	0.048	*0.001*	0.057	*0.001*

Source: Trostel, Walker and Woolley (2001).

Notes: Robust standard errors are in italics. The estimating equations include year dummies, union status, marital status, age and age squared and, in the case of the aggregate equation, country-year dummies.

There is a great deal of heterogeneity in the rate-of-return estimates across these 28 countries[8]. Pooling the samples suggests a worldwide OLS estimate of the rate of return to schooling in the order of 4.8 percent for men, and 5.7 percent for women. Both the magnitude and the gender differential are consistent with the previous literature but the cross-country variation in the estimated rates of return is quite striking. The highest estimate (19.2 percent for females in the Philippines) is ten times higher than the lowest estimate (1.9 percent for females in the Netherlands). The UK is again amongst the highest. The estimates are generated using uniform procedures and comparable data across countries. And the rates of return are typically estimated with considerable precision. However, we find it difficult to explain much of this cross-country variation. Confronting these estimates with aggregate information about each country we find that there is tenuous evidence that the rate of return declines with average educational attainment (i.e., diminishing returns to schooling), a significantly negative effect of per capita income, and, surprisingly, no effect of relative spending on education[9]. The most puzzling aspect of the cross-country heterogeneity is the lack of obvious explanations for it. Few general patterns are apparent in the cross-country variation. It appears that the returns are generally higher outside Continental Europe but this seems to be the only clear pattern.

Some recent studies using US data suggest that the return to schooling is increasing over time, possibly due to increased returns to skill or ability[10]. ISSP also provides multiple cross sections of data for most of the 28 countries, which allows us to investigate how the returns vary over time. Contrary to evidence from data from the US in the mid 70's to late 1980s, there is essentially no evidence of a rising rate of return. In most countries there is no significant trend in the rate of return to education, and, overall, there is evidence of a slight decline in the worldwide rate of return over the 1985-95 period, particularly for women. By including a trend interaction with schooling we estimate how the return to schooling has grown over time, on average, over the years in the data. Table 3 shows these results for the countries with a

[8] In many countries the earnings data is interval midpoints but estimates obtained using Stewart's (1983) interval-regression technique show that the grouped nature of the dependent variable produces only minor differences in the estimates. As little is gained using maximum likelihood, we simply use the midpoint data in the subsequent estimates.

[9] One other variable included in the equation was whether the wage data was recorded gross or net of tax (which was significantly negative as we would expect with progressive tax systems).

[10] For example, Blackburn and Neumark (1993), Murnane *et al.* (1995) and Cawley *et al.* (1997).

minimum of five years of data. In most countries there are no significant increases (or indeed decreases) in the returns to schooling, including the USA. When we include a quadratic trend, however, the US return for males rises initially before falling in the later years. Moreover, both the linear and quadratic trend coefficients are marginally significant for US men (although not for US women). Thus, the evidence is consistent with the earlier US evidence, but it appears that the trend was reversed in first half of the 1990s.

Table 3 OLS Estimates of the Trend in the Return to Years of Education: ISSP Data

	Males				Females			
Country	Initial Rate of Return	Std error	Trend	Std error	Initial Rate of Return	Std error	Trend	Std error
USA	0.0742	*0.0078*	-0.0001	*0.0014*	0.0963	*0.0110*	-0.0001	*0.0018*
Great Britain	0.1158	*0.0099*	0.0026	*0.0019*	0.1501	*0.0101*	-0.0042	*0.0018*
W. Germany	0.0455	*0.0042*	-0.0020	*0.0007*	0.0546	*0.0089*	-0.0023	*0.0014*
Russia	0.0260	*0.0049*	0.0137	*0.0032*	0.0374	*0.0045*	0.0118	*0.0030*
Norway	0.0388	*0.0047*	-0.0048	*0.0012*	0.0344	*0.0057*	-0.0028	*0.0015*
Australia	0.0612	*0.0067*	-0.0022	*0.0012*	0.0863	*0.0115*	-0.0069	*0.0018*
Netherlands	0.0279	*0.0033*	0.0008	*0.0008*	0.0475	*0.0070*	-0.0072	*0.0015*
Austria	0.0397	*0.0089*	-0.0003	*0.0013*	0.0660	*0.0127*	-0.0004	*0.0017*
Poland	0.0660	*0.0080*	0.0035	*0.0034*	0.0983	*0.0085*	0.0010	*0.0031*
E. Germany	0.0199	*0.0044*	0.0037	*0.0024*	0.0391	*0.0052*	0.0032	*0.0025*
NewZealand	0.0403	*0.0066*	-0.0042	*0.0031*	0.0220	*0.0100*	0.0041	*0.0043*
Italy	0.0350	*0.0089*	0.0004	*0.0017*	0.0737	*0.0115*	-0.0046	*0.0023*
Ireland	0.0734	*0.0099*	0.0035	*0.0021*	0.0899	*0.0145*	0.0002	*0.0031*
N. Ireland	0.1960	*0.0171*	-0.0091	*0.0056*	0.1417	*0.0162*	0.0020	*0.0060*
Pooled	0.0532	*0.0023*	-0.0008	*0.0003*	0.0766	*0.0033*	-0.0028	*0.0005*

Source: Trostel, Walker and Woolley (2001).

Notes: Robust standard errors are in italics. The estimating equations include year dummies, union status, marital status, age and age squared and, in the case of the aggregate equation, country-year dummies.

In general, though, for males there is little evidence that the return to education is either increasing or decreasing appreciably worldwide. There are equal numbers of negative and positive trend coefficients for men, and most are not significantly different from zero. When the samples are pooled, however, the overall trend in the return to schooling is slightly downward for men over the 1985-95 period[11]. For women, there is slightly stronger evidence of a declining rate of return. But even for

[11] This is broadly consistent with previous evidence of a decreasing return in Europe; see for example, Goux and Maurin (1994) and Jarousse (1988).

women, the evidence is far from uniform across countries. Most trend coefficients for females are insignificant, and there are almost equal numbers of positive and negative coefficients.

It has become well known that the OLS estimate of the return to education is unbiased only if measured schooling is exogenous. Endogeneity arising from measurement error in S is generally thought to bias the estimate of β towards zero, although this effect is believed to be small because the reliability of schooling data is typically quite high[12]. Secondly, endogeneity can arise because of omitted ability. That is, the returns coefficient, β, is biased (upwards) because chosen schooling levels are (positively) correlated with omitted ability, and ability is (positively) correlated with the wage rate. On the other hand, as we suggested earlier, Card (1999) and others[13] have argued that OLS estimates of β are biased downwards because individuals with high discount rates choose low levels of schooling, which have a higher marginal rate of return. Most of the recent studies reviewed in Card (1999) suggest that OLS estimates of β are indeed biased downwards. Card's review only includes one non-US study, however. Thus, it is of considerable interest to investigate the extent to which the conclusion is more general.

The twins methodology mentioned earlier has been used in recent studies in the US using the "Twinsburg" samples (Ashenfelter and Krueger (1994), Ashenfelter and Rouse (1998) and Rouse (1997)), while earlier work can be found in Berhman *et al* (1999). The recent US research seems to confirm that the return to education is close to the figures commonly found in IV studies. Elsewhere, studies exist for Sweden (Issacson, (1999)) and Australia (Miller *et al*, (1995)). Only one study exists for the UK.

Here we also try to address the problem of the potential endogeneity of schooling by presenting instrumental-variable estimates of the rate of return. It has been suggested in recent work (for example, Weiss, 1999) that marriage is subject to assortative mating. A common level of schooling is also more likely to lead to common experiences, and possibly common interests. Pencavel (1998) points out that, in US census data, husbands and wives have been becoming more similar in their

[12] This is true even though schooling is self-reported.

[13] Notably Lang (1993).

schooling backgrounds. In 1990 there were 8.62 times as many couples with schooling differences of no more than one category than those with schooling differences of more than one category. We have data on spouse's education (as reported by the reference person) in ten of the 28 countries and we have data on father's education (again, as reported by the reference person) in nine countries (eight of which are also in the group with data on spouse's education), and on mother's education in eight countries.

Table 4 presents IV estimates of β using spouse's education to instrument for S, together with the corresponding OLS estimates (for the subset of people recording spouse's education). Table 5 presents IV estimates using father's education as the instrument, along with the corresponding OLS estimates. IV estimates using mother's education as the instrument are given in Table 6[14].

Table 4 IV Estimates of Returns to Years of Education using Spouse's Education: ISSP Data

		Males		Females	
Country		Coeff	*Std error*	Coeff	*Std error*
USA	IV	0.084	*0.009*	0.116	*0.015*
	OLS	0.068	*0.005*	0.106	*0.009*
West Germany	IV	0.042	*0.008*	0.069	*0.012*
	OLS	0.038	*0.004*	0.056	*0.009*
Australia	IV	0.055	*0.011*	0.086	*0.021*
	OLS	0.055	*0.005*	0.060	*0.008*
Netherlands	IV	0.048	*0.014*	0.053	*0.017*
	OLS	0.034	*0.006*	0.047	*0.011*
Poland	IV	0.073	*0.009*	0.102	*0.014*
	OLS	0.071	*0.006*	0.106	*0.008*
East Germany	IV	0.033	*0.010*	0.054	*0.019*
	OLS	0.029	*0.005*	0.039	*0.007*
Italy	IV	0.075	*0.010*	0.113	*0.012*
	OLS	0.040	*0.004*	0.064	*0.006*
Ireland	IV	0.088	*0.014*	0.132	*0.023*
	OLS	0.063	*0.009*	0.109	*0.013*
Hungary	IV	0.081	*0.015*	0.103	*0.022*
	OLS	0.058	*0.011*	0.070	*0.011*
Czechoslovakia	IV	0.043	*0.024*	0.046	*0.014*
	OLS	0.036	*0.011*	0.033	*0.008*
Weighted Avg	IV	0.064	*0.011*	0.093	*0.017*
	OLS	0.053	*0.005*	0.076	*0.008*

Source: Trostel, Walker and Wooley (2001).

Note: Robust standard errors are in italics. The estimating equations include year dummies, union status, marital status, age and age squared.

[14] IV estimates using both parent's education as instruments yield results exactly as expected, that is, somewhere between those shown in Tables 5 and 6.

Table 5 *IV Estimates of Return to Years of Schooling using Father's Education: ISSP Data*

Country	Males		Females	
USA	0.106	*0.016*	0.136	*0.021*
West Germany	0.042	*0.009*	0.056	*0.015*
Australia	0.055	*0.011*	0.088	*0.033*
Austria	0.053	*0.012*	0.078	*0.017*
Poland	0.078	*0.013*	0.143	*0.018*
East Germany	0.048	*0.019*	0.043	*0.028*
Ireland	0.123	*0.025*	0.158	*0.043*
Hungary	0.099	*0.027*	0.072	*0.023*
Czechoslovakia	0.065	*0.024*	0.051	*0.014*
Weighted Average	0.072	*0.014*	0.103	*0.022*

Source: Trostel, Walker and Wooley (2001).

Note: Robust standard errors are in italics. The estimating equations include year dummies, union status, marital status, age and age squared.

Table 6 *IV Estimates of Return to Years of Schooling using Mother's Education : ISSP Data*

Country	Males		Females	
USA	0.128	*0.018*	0.125	*0.019*
West Germany	0.029	*0.009*	0.042	*0.014*
Australia	0.114	*0.031*	0.129	*0.050*
Austria	0.075	*0.016*	0.074	*0.019*
Poland	0.074	*0.020*	0.161	*0.024*
East Germany	0.038	*0.023*	0.049	*0.023*
Ireland	0.130	*0.028*	0.148	*0.032*
Hungary	0.086	*0.030*	0.075	*0.019*
Weighted Average	0.072	*0.014*	0.103	*0.022*

Source: Trostel, Walker and Wooley (2001).

Note: Robust standard errors are in italics. The estimating equations include year dummies, union status, marital status, age and age squared.

Thus, we exploit the very strong correlation between spouses' education and the lack of correlation between the wage of one spouse and the education of the other. Consistent with most previous evidence, the IV estimates are substantially greater than the OLS estimates; that is, OLS estimates appear to be biased downward significantly. Using spouse's education to instrument for observed schooling yields estimates of the rate of return that, on average, are higher than the corresponding OLS estimates. Despite our reservations that family background, such as parental education, provides rather weak instruments we find that using education of the father and education of the mother to instrument for schooling yields broadly the same conclusion.

The IV estimates shown in Tables 4-6 are consistent with the previous literature: namely, the OLS estimates of the return to schooling are biased downward substantially. On average, the IV estimates using spouse's education are over 20 percent higher than the corresponding OLS estimates. In addition, because of the very strong correlation between reported schooling and our instruments, our IV estimates are considerably more precise than most previous IV estimates of the return to education. Thus, the previous finding that IV estimates of the return to schooling are considerably higher than OLS estimates appear not to be unique to the US and the UK.

3.4 Existing UK Evidence

There is surprisingly little UK research into the private returns to education. Here we give a brief overview of the literature that directly focuses on this issue in tabular form[15]. Table 7 outlines the period, data set and variables used in the selected studies. The dependent variable is the logarithm of earnings throughout. Harmon and Walker (1995) use pooled Family Expenditure Survey (FES) data for 1978 to 1986. Miles (1997) uses FES data for individual years. Harmon and Walker (1999) use pooled General Household Survey (GHS) data for 1974 to 1994 where each year contains approximately 20,000 observations, of which about half have employee/wage data. Bell (1996), Dearden (1998) and Harmon and Walker (2000) uses the National Child Development Survey (NCDS) data, which is a continuing longitudinal survey of people who were born between 3 and 9 March 1958. There were initially over 18,000 in the study, of which 5,000 have been lost due to attrition and an even greater number feature crucial missing values. Hildreth (1997) and Ermisch and Francesconi (2000) use British Household Panel Study (BHPS) data for the years 1991 and 1995. The BHPS is a panel survey of 5,500 households (over 10,000 individuals) interviewed annually. Brown and Sessions (1998) use pooled British Social Attitudes (BSA) survey data. Beginning in 1983 there are around 3,000 adults in each annual survey.

[15] Various other studies whose focus is not on education are not listed here but include estimates of education returns.

Table 7 Summary of Previous Specifications

	Bell (1996)	**Blundell, *et al* (1997)**	**Dearden (1998)**	**Dearden *et al* (2000)**	**Ermisch & Francesconi (2000)**
Year	1991	1991	1991	Not stated	1991-1995
Data Set	NCDS	NCDS	NCDS	LFS	BHPS
Data Type	Cohort study	Cohort	Cohort study	Pooled X section	Panel
Education Yrs	✓				
Qualifications		✓	✓	✓	✓
Age				✓	✓
Age^2				✓	✓
Experience					
$Experience^2$					
Region		✓	✓	✓	
Family		✓	✓	✓	✓
School Type		✓	✓		
Employer	✓	✓	✓	✓	
Trade Union	✓	✓	✓		
Sex	✓	✓	✓	✓	✓
Year					
Ability	✓		✓		
Occupation	✓	✓			
IV/H2S		✓***	✓		✓
Work Selection					✓

	Harmon & Walker (1995)	**Harmon & Walker (1999)**	**Harmon & Walker (2000)**	**Hildreth (1997)**
Year	1978-86	1974-94	1991	1991, 1995
Data Set	FES	GHS	NCDS	BHPS
Data Type	Pooled X Section	Pooled X-Section	Cohort study	Panel
Education Yrs	✓	✓	✓	✓
Qualifications				✓
Age	✓	✓		✓
Age^2	✓	✓		✓
Experience				✓
$Experience^2$				✓
Region	✓	✓	✓	✓
Family			✓	
School Type			✓	✓
Employer				✓
Trade Union				✓
Sex	M only	M only	M only	✓
Year	✓	✓		
Occupation				✓
IV/H2S		✓	✓	
Work Selection	✓			✓**

* Corrected for selectivity into employment and self-employment
** Also corrects for selectivity by union status.
*** Also corrects for selection of sample with A levels.

Bell (1996), Brown and Sessions (1998), Dearden (1998), Ermisch and Francesconi (1997), Hildreth (1997) and Harmon and Walker (1995, 1999, 2000) use the log real gross hourly wage using the Retail Price Index as the deflator where necessary. Miles (1997) uses the "normal" level of net weekly household earnings.

Harmon and Walker (1995) and Harmon and Walker (1999) use years of full time schooling imputed from the reported school leaving age. Miles (1997) uses the actual school leaving age. Harmon and Walker (1999) also introduce school leaving age dummies to allow for non-linearity in the effect of schooling on wages. Harmon and Walker (2000) use years of post-16 schooling. Bell (1996), Dearden (1998), Brown and Sessions (1998) and Hildreth (1997) use both years of schooling and qualifications, again to allow for non-linearity in returns. Ermisch and Francesconi (2000) use O level, A level and "higher" dummies and interactions between these and age.

Hildreth (1997) uses both age and actual experience. Brown and Sessions (1998) use potential experience and potential experience squared. Ermisch and Francesconi (1997) use age and age squared, together with their interactions with qualifications. Due to the potential endogeneity of experience Harmon and Walker (1995) and Harmon and Walker (1999) use age and age squared. Miles (1997) uses age and age squared interacted with employment status to allow for the life cycle profile of earnings.

Dearden (1998) uses ordinary least squares (OLS) to obtain estimations of the standard Mincerian earnings function. She then uses the instrumental variable (IV) approach on a specification that includes years of schooling as the explanatory variable, and then in a model that uses qualifications she employs a selection model where schooling is treated as an ordered probit to overcome the fact that it is not a continuous variable. This is a Heckman two-step procedure where a correction term (the inverse Mill's ratio) is obtained in the first stage reduced form equation, and then included as a regressor in the earnings function. Separate estimates are obtained for males and females.

The work by Dearden, McIntosh, Myck and Vignoles (2000) used several LFS cross sections to estimate the impact of educational qualifications, and non-vocational qualification, on wage rates – although the main focus of that work was to investigate the impact of basic numeracy and literacy skills using the National Child Development Survey (NCDS) and the International Adult Literacy Survey (IALS).

Harmon and Walker (1995) and Harmon and Walker (1999) estimate using OLS, and then IV (and also use the Heckman two step model to allow for the fact that schooling is not a continuous variable). Harmon and Walker (1999) also use dummy variables for different levels of schooling to allow for non-linearity in the effect on wages. Harmon and Walker (2000) use OLS and IV on the endogenous (post-16) component of education. Ermisch and Francesconi (2000) follow a similar approach to Harmon and Walker (1999) with corrections being included for labour force participation and education selections. The reliance on only qualifications does allow for some non-linearity. Brown and Sessions (1998) also use two-step methods. The first of which is a multinomial logit regression to provide a correction term for selection into unemployment, self-employment or general employment. The second step is OLS with the relevant correction term included. They endeavour to distinguish between signals of ability and the actual increase in human capital, under the presumption that self-employed individuals attach no weight to the signal since they know their own ability. Hildreth (1997) provide OLS estimates after selection into unions has been corrected for and include gender decompositions. This paper also attempts to explain changes in the union wage differential. Miles (1997) uses standard OLS in his paper.

The finding that more educated people have higher earnings seems a strong and robust feature of the OLS literature. Despite the differences in specification and time periods covered the estimated size of the effect of years of education in the OLS studies lies within a relatively narrow range – all but two studies for men lie within the 4.1 to 6.1% range. However, there is less agreement about the effect of qualifications across studies.

There is also some agreement that (the few UK) IV estimates, as in the US and in our international review earlier, are higher than OLS. It seems likely that the reason why the Harmon and Walker IV results are so much higher is that they rely on education reforms as the source of instruments, while the work by Dearden relies on family background variables. Card (2001) also notes this in his review of US IV research and favours education reforms since it seems likely that family background variables are a source of weaker instruments because they are likely to be correlated with wages when the children grow up as well as with the level of education while they are young.

Table 8 *Estimated returns to years of schooling*

	Males		Females	
	OLS	**IV**	**OLS**	**IV**
Dearden (1998)	4.8%	5.5%	8.3%	9.3%
Harmon and Walker (1995)	6.1%	15.2%		
Harmon and Walker (1997a)	4.1%	14.0%		
Hildreth (1997)*	5.0%		5.0%	
Miles (1997)	≈3%		≈3%	
Brown and Sessions (1998)	10.8%			
Bell (1996*)	4.6%		4.6%	
Harmon and Walker (1997b)	5.1%	9.9%		

Table 9 *Estimated Returns To Qualifications*

		Males		Females	
		OLS	**IV**	**OLS**	**IV**
Dearden (1998)	A level	37.6%	41.7%	37.2%	43.9%
	Degree	50.1%	56.2%	63.6%	73.8%
Hildreth (1997***)	A level	30.9%		30.9%	
	Degree	68.25		68.2%	
Brown and Sessions (1998)	A level	34.3%			
	Degree	71.2%			
Ermisch and Francensconi	A level	9.6%	9.6%	43.5%	
(2000**)	Degree	26.4%	36.4%	71.3%	
Dearden *et al* (2000)	A level	16.8%		18.5%	
	Degree	27.7%		25.4%	
Blundell et al (1997) ****	Degree	20.8%	17.1%	39.1%	36.8%
Bell (1996)	A level	25.9%		25.9%	
	Degree	45.2%		45.2%	

* Males and Females in the same equation.

** Education not found to be endogenous for females.

*** Qualification returns calculated as the sum of the returns to years of schooling and qualification dummies.

**** The omitted category here is A levels so that these estimates correspond to the difference between the degree and A-level estimates in the other studies.

This begs the question of whether this means that a randomly selected person who invests in more education will earn higher wages. If the earnings-schooling relation is a causal one then the answer is yes: education then increases productivity and higher productivity increases wages. If the relationship is a spurious one then the answer is no: more able people are more productive and are therefore paid a higher wage. Indeed such people acquire more education so as to signal their high ability. It is therefore hard to say whether a simple relation between earnings and education can be interpreted as a return to education for a randomly selected person. To support such an interpretation, one must convincingly control for factors such as ability and family background that might both affect the choice of education and the wage. The instrumental variable approach takes the view that it is enough to purge the schooling variable of its correlation with the unobservable determinant of schooling by regressing schooling on the instrumental variable(s) and then including the predicted value from this regression in the wage equation.

Finally, only Bonjour *et al* (2000) provides twins estimates for the UK – they obtains results close to typical OLS results for cross section data. However the data they use is a highly selected sample of women involved in assisting medical research. This is an area where the UK clearly lacks good data to support the application of this methodology.

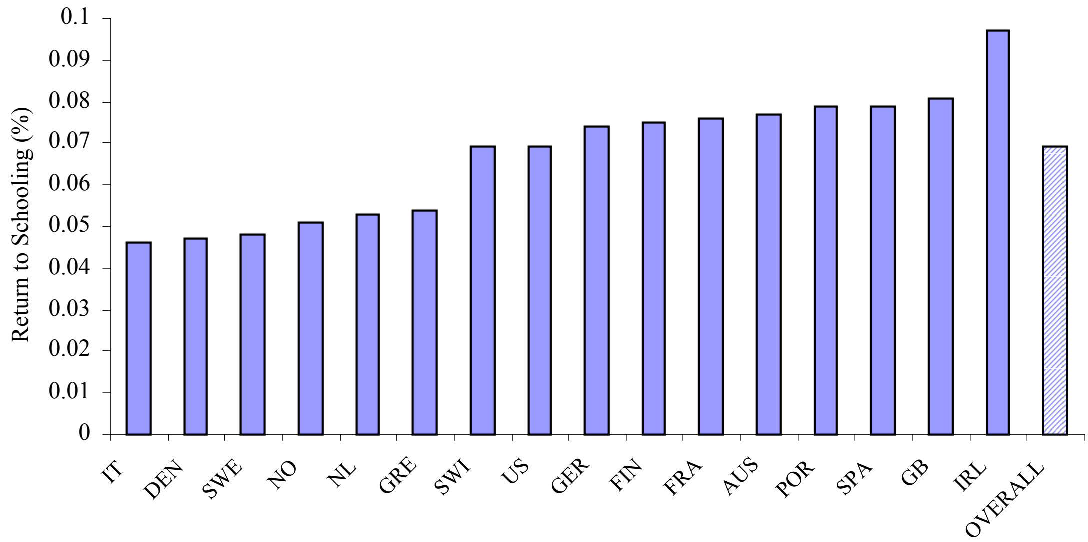
3.5 A Summary of Existing Evidence

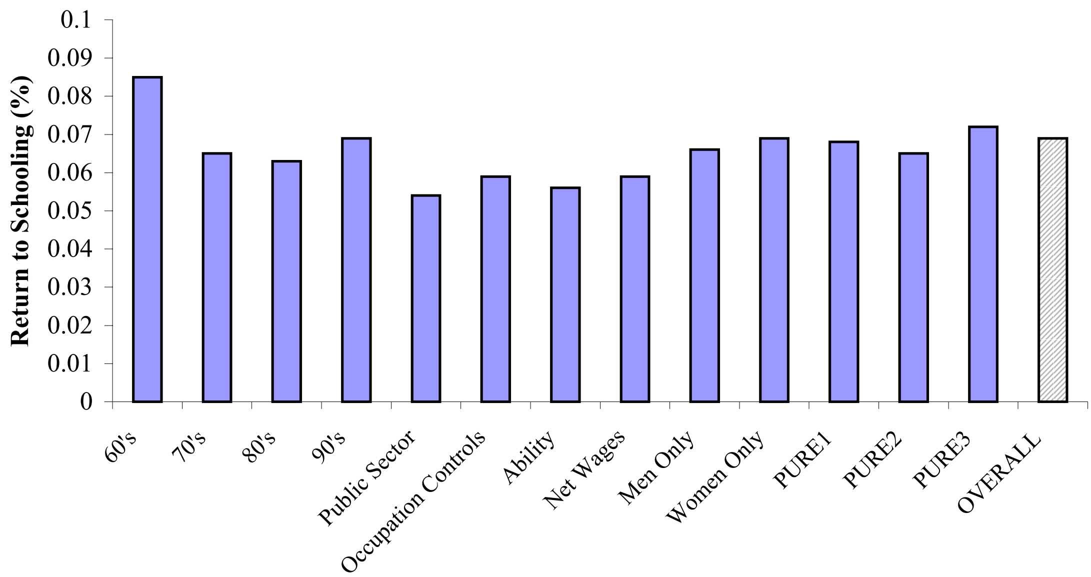
To summarize the various issues discussed above we use the methods common in meta-analysis to provide some structure to our survey of returns to schooling and to provide a framework to determine whether our inferences are sensitive to specification choices.

In Figure 1 we present the findings of a simple meta-analysis based on the collected OLS estimated rates of return to schooling from the PURE project (see Harmon, Walker and Westergard-Nielsen (2001)) supplemented by a number of findings for the US. Well over 1000 estimates were generated across the PURE project on three main types of estimated return to schooling - existing published work (labelled PURE1), existing unpublished work (PURE2), and new estimates produced for the PURE project (PURE3). Each block refers to a different sample of studies (for example only studies based on US originated studies).

A number of points emerge from the figure. There is a remarkable similarity in the estimated return to schooling for a number of possible cuts of the data with an average return of around 6.5% across the majority of countries and model specifications. There are number of notable exceptions. That Nordic countries generally have lower returns to schooling is confirmed while at the other extreme the returns for the UK and Ireland are indeed higher than average at 8% and 10% respectively. In addition estimated returns from studies of public sector workers, and from studies where net (of tax) wages are only available average about 5%[16]. Surprisingly, studies that control for ability seem not to produce estimated returns that are very much lower than those that do not. Estimates produced using samples from the 1960's also seem to have produced higher than average returns.

Figure1 *Returns to Schooling – A Meta Analysis*





[16] Note that we would expect the net returns to be lower than gross because tax systems are typically progressive.

4. New Evidence from the LFS Data

This section produces estimates of simple specifications that are designed to be broadly comparable with earlier research. We aim to give the flavour of the relationship between education and wages in the LFS data[17]. In this section we present estimates that derive from the simplest possible specification which has log wages linear in years of education. This specification is commonly used in US research where it is justified by the alleged homogeneous nature of the US education system[18]. In fact, the main vehicle for this kind of analysis in the US is the Current Population Survey (CPS, similar to the UK LFS) where qualifications were not recorded until the early 1990's so US research has, until recently, been forced to make this simplifying assumption. Despite its restrictive nature there are few tests of its legitimacy and we present some evidence below. In this section we also allow for the relationship between wages and education to shift over time. Separate results for men and women are presented.

The main difficulty with LFS arises from the censored nature of the data on age at which the individual left full-time education. The LFS question refers to *continuous* education so that interrupted spells of education may be censored. In particular, a "gap" year between secondary and higher education could result in the spell of higher education not being reported. We experimented with a variety of ways of controlling for this problem and found no substantial differences in results so this problem with LFS seems to be empirically unimportant. The results reported below use the qualifications data to place a lower bound on the years of education: thus, for example, if age left education is coded as under 20 but the individual records having a degree we created a dummy variable to capture this peculiarity. We also created a dummy variable for those who appear to have taken a long time to complete their degrees (allowing for variations in length arising from the subject taken). In fact, the prevalence of these exceptions was not large and it made little difference to our estimates of the coefficients of interest.

[17] Throughout we condition on having positive hours of work and earnings, age 25-59 inclusive, and we exclude Scotland and Northern Ireland.

[18] The linearity in S assumption also facilitates the use of more sophisticated estimation methods such as instrumental variables.

The results for the years of education specification are show in Table 10 for women and 11 for men. The dependent variable is the log of hourly wages so the coefficients can be interpreted as % effects on wages. The *t*-statistics show whether the coefficients are significantly different from zero (if the corresponding *t* is greater than 2 then the coefficient is statistically significantly different from zero at the conventional 95% confidence level). For women age earnings profiles are flatter than for men – reflecting the fact that age is a less good proxy for experience for them. For women there is a marriage penalty while for men there is a large premium for being married. For women there is a small premium for being a cohabiter relative to being married, while for men there is a small penalty for this. For women the union differential is close to double that for men while the penalty for being non-white is smaller. The penalty for having poor health is around 10-12% for both. Non-vocational qualifications were grouped into NVQ 1 and 2 and NVQ 3 and 4 according to the definitions in Dearden *et* al (2000)[19] and the small group of NVQ5 was omitted because these are essentially post (academic) degree professional qualifications[20]. A number of coefficients are not reported to keep the tables brief: regional effects are pronounced and predictable with, for example, a large premium for the South East; and the premium for having had a gap-year is large and clearly picks up some unobservable difference (the premium is about 25% for men in the simplest specification) arising, for example, from individuals who take such gaps being more confident, being fluent in languages, or having other skills that are positively rewarded in the labour market.

Experiments with more sophisticated specifications, including models that allowed for a variety of interactions between education and characteristics, while showing some significant differences in returns across groups, did not change the broad findings in these tables[21].

A number of experiments were conducted to attempt to identify nonlinearities in the relationship between log wage and years of education by including a separate dummy variable for each year of education. The censored nature of our measure of

[19] See their Table 4.2 on page 24.

[20] We found that there was no significant difference between NVQ1 and 2 when they were entered separately. The same was true for NVQ 3 and 4.

Table 10 OLS Returns to Education Years - Women

	All	1993	1994	1995	1996	1997	1998	1999	2000
Education Years	0.077	0.075	0.077	0.072	0.074	0.075	0.075	0.075	0.076
	119.2	*44*	*44.1*	*43.6*	*45.6*	*45.1*	*45.2*	*45.7*	*22.1*
Age	0.006	0.006	0.006	0.006	0.006	0.006	0.0006	0.005	0.004
	36.8	*13.2*	*15.2*	*14.2*	*14.2*	*13.5*	*12.8*	*12.5*	*4.61*
Age squared	-0.0007	-0.0007	-0.0006	-0.0007	-0.0007	-0.0006	-0.0007	-0.0007	-0.0006
	39.9	*14.3*	*13.3*	*15.3*	*16.1*	*13.3*	*14.6*	*14.6*	*6.6*
Married	0.132	0.121	0.119	0.127	0.133	0.141	0.142	0.144	0.115
	36.5	*12.2*	*12.3*	*13.2*	*13.6*	*14.7*	*14.6*	*14.9*	*5.43*
Cohabiter	0.068	0.039	0.063	0.072	0.058	0.076	0.071	0.072	0.055
	12.2	*2.52*	*4.32*	*5.01*	*4.11*	*5.43*	*5.15*	*5.66*	*2.02*
Non white	-0.117	-0.061	-0.127	-0.139	-0.051	-0.166	-0.137	-0.111	-0.124
	8.88	*1.28*	*3.02*	*3.65*	*1.15*	*5.22*	*4.1*	*3.44*	*1.7*
Union	0.077	0.065	0.092	0.087	0.09	0.082	0.066	0.064	0.067
	26.5	*6.86*	*12.3*	*11.6*	*11.7*	*10.1*	*8.57*	*7.99*	*3.76*
Health problem	-0.12	-0.118	-0.14	-0.124	-0.104	-0.111	-0.143	-0.122	-0.096
	21.9	*6.04*	*8.31*	*7.31*	*7.7*	*8.99*	*9.78*	*8.8*	*3.31*
NVQ 1-2	0.181	0.172	0.182	0.192	0.201	0.154	0.158	0.169	0.162
	19.1	*4.56*	*4.33*	*4.67*	*4.22*	*4.5*	*5.09*	*5.01*	*4.78*
NVQ 3-4	0.332	0.268	0.387	0.333	0.354	0.368	0.391	0.382	0.291
	21.1	*3.21*	*4.11*	*5.22*	*5.01*	*5.33*	*6.01*	*5.32*	*4.04*
R squared	0.3	0.29	0.3	0.28	0.29	0.29	0.28	0.29	0.27
N	76498	9817	10358	10827	10901	10870	10758	10497	2470

Note: Specification also includes year dummies (all column), region dummies, dummy variables to control for discrepancies between qualification and years of education.

Table 11 OLS Returns to Education Years - Men

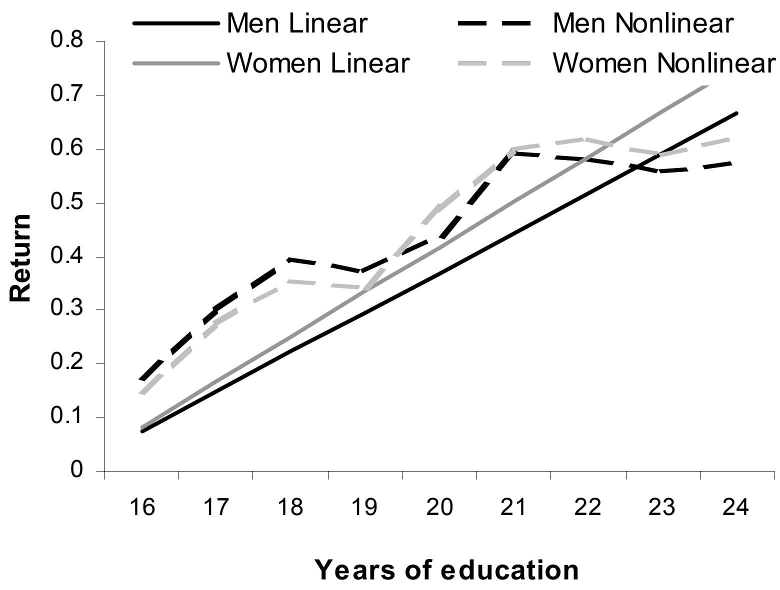
	All	1993	1994	1995	1996	1997	1998	1999	2000
Education Years	0.084	0.094	0.089	0.09	0.081	0.088	0.082	0.081	0.083
	128.2	*51.2*	*51*	*51.1*	*44.3*	*49.1*	*48.8*	*48.4*	*23.6*
Age	0.002	0.002	0.002	0.002	0.002	0.003	0.001	0.002	0.001
	13.4	*5.22*	*5.1*	*5.11*	*4.21*	*6.49*	*3.83*	*4.82*	*1.4*
Age squared	-0.0002	-0.0002	-0.0002	-0.0002	-0.0001	-0.0002	-0.0001	-0.0003	-0.0002
	12.5	*4.4*	*4.29*	*5.39*	*3.97*	*4.81*	*3.82*	*6.19*	*1.85*
Married	-0.023	-0.045	-0.033	-0.022	-0.034	-0.02	-0.011	-0.007	-0.003
	77.32	*4.9*	*3.89*	*2.78*	*3.9*	*2.46*	*1.96*	*0.99*	*0.43*
Cohabiter	0.033	0.018	0.033	0.035	0.015	0.042	0.031	0.048	0.049
	6.6	*1.2*	*2.39*	*2.61*	*1.06*	*3.33*	*2.38*	*3.75*	*1.86*
Non white	-0.045	0.005	-0.081	-0.057	-0.036	-0.064	-0.019	-0.053	0.047
	3.91	*0.01*	*2.61*	*1.75*	*1.17*	*2.25*	*0.62*	*1.88*	*0.69*
Union	0.188	0.149	0.21	0.188	0.2	0.198	0.185	0.174	0.185
	66.1	*15*	*27.2*	*24.5*	*25.9*	*22.3*	*23.7*	*22.1*	*11.2*
Health problem	-0.098	-0.104	-0.097	-0.127	-0.09	-0.099	-0.114	-0.069	-0.11
	19.2	*5.68*	*6.01*	*7.91*	*6.93*	*8.86*	*8.48*	*5.49*	*4.08*
NVQ 1-2	0.112	0.132	0.122	0.128	0.133	0.116	0.108	0.101	0.093
	17.2	*3.59*	*2.49*	*3.29*	*4.21*	*3.27*	*3.21*	*3.1*	*3.2*
NVQ 3-4	0.354	0.326	0.31	0.298	0.367	0.342	0.372	0.332	0.328
	19.2	7.21	7.02	6.82	6.39	7.21	8.04	6.52	6.35
R squared	0.37	0.39	0.39	0.39	0.36	0.36	0.35	0.35	0.36
N	80172	10216	10804	11128	11461	11458	11384	11064	2657

Note: Specification also includes year dummies (all column), region dummies, dummy variables to control for discrepancies between qualification and years of education.

[21] For example, we find a higher rate of return for non-white individuals. We were unable to control for parental social class since this is not recorded in the LFS data.

years of education because of the gap year problem makes this more difficult here but the restriction of linearity (while including the dummy variables that control for long degree programmes, gap year, and early achievers) was soundly rejected by the data for women and marginally rejected for the data for men. For example, we found that relaxing the specifications in Tables 10 and 11 to allow for each year of education to have an independent effect on wages we found returns as shown in Figure 2[22]. The returns, given by the slope of the lines in the figure, are clearly higher at low levels of education and lower at high levels.

Figure 2 *OLS Returns to Education Years – Men and Women*

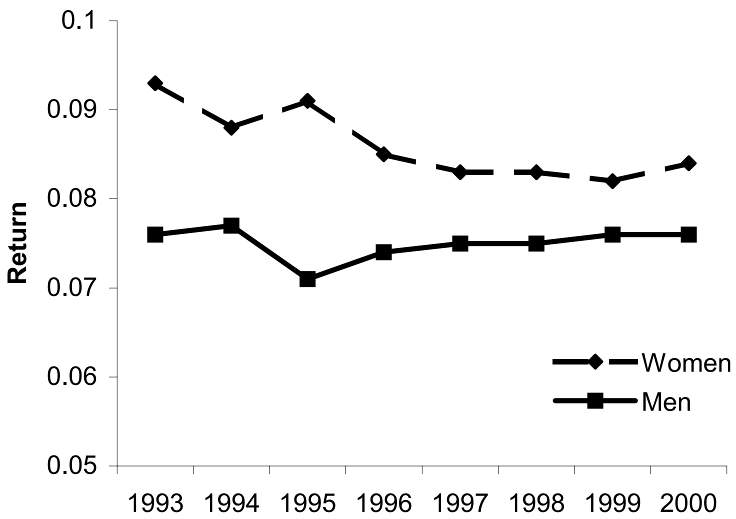


Some care needs to be exercised in interpreting Figure 2 . It seems that the return to just one year of education beyond 18 is small compared to the return to 3 years of schooling beyond 18. Similarly the return beyond 21 seems small. However, it needs to be borne in mind that the numbers who leave education at 19 and 20, and beyond 21, are small relative to the numbers stopping at 18 and 21. Moreover, it seems very likely that they are different. For example, they may have lower ability, motivation or perseverance. In Trostel and Walker (2001) we record in more details our attempts to model "sheepskin" effects. We found that there are large effects associated with the years of education when qualifications are obtained, using the US CPS data. This work exploits the collection of both years of education and qualifications in the year when CPS changed from one to the other. However, we

[22] In the case of both men and women we trimmed the data by dropping observations with age completed education below 15 and above 24. This amounted to less than 4% of each sample.

found that there were significant sheepskin effects in hours of work as well as in wages and this suggested to us that these effects are not, as commonly assumed, due to ability (which should leave hours of work unchanged) but to unobserved differences in effort. This finding undermines the interpretation of sheepskin effects: that the correlation between wages and education reflects unobserved ability. Thus, it seems likely that the returns between 16 and 18, and between 18 and 21 give the best guide to what the average return would be for a typical individual. While Figure 2 suggests that the average annual return between 16 and 18 is larger than that between 18 and 21, the difference is small and neither would be statistically different from the average returns estimated by OLS in a simple linear specification.

Figure 3 OLS Returns to Education Years – Men and Women



Thus, the stylised facts from the data are reflected in Figure 3 which plots the values estimated for each year:

- The estimated rate of return to education is around 9% for women and 8% for men and highly statistically significant.
- The returns for men change little over time but there is a slight fall in the return for women (a (just) statistically significant change of about –0.9% - i.e. by about 10% of its original value).

5. The Returns to Qualifications

The LFS data is large and this facilitates separate analyses for particular groups of individuals. In particular, this section of the report contains microeconometric estimates of the effects of a degree on wages that allows for different degree subjects to have differential effects. Separate results for men and women are presented. The large size of the LFS also allows us to examine the stability of the estimates over time and in particular, to examine whether there is any time trend in the returns to education or to particular qualifications[23]. The results for women are given in Table 12 below. The estimated effects of the characteristics remain unchanged (relatively flat age earnings profiles, a substantial union differential etc.) but the table does allow us to check the simplifying assumption that wages are linear in years of education. It is immediately clear from these estimates that there are substantial effects associated with qualifications (which persist even if education is included in the specification in addition to qualifications).

For women, attaining GCSE (or O level or CSE(1) standard) yields a premium of approximately 8% (relative to no qualifications), while 2 or more A levels yields a further 17% (0.254-0.078), and an undergraduate degree yields a further 19% (0.443-0.254). Since A-levels take two-years to acquire this estimated A-level effect broadly reflects the estimates in Table 12 of an annual average return of approximately 8%. The significant result for GCSE relative to no qualifications, and between 2+ A-levels and just one A-level is probably due to some unobserved differences in ability, background, or motivation since, in the data, they are associated with approximately the same level of years of education. The effects of degree relative to (2) A-levels suggest an annualised return of around 6%. This smaller return supports the results portrayed in Figure 2.

Again the results for the education variables are statistically very well determined, although a more general specification that allows for the coefficients to vary over time in an unrestricted way cannot be rejected. However, a specification that interacts the education variables with a time trend suggests that there are no

[23] A further difficulty with the LFS data is that there have been changes to the qualifications information over time. The major change was in 1996 – before which individuals were invited to list three qualifications, but after which they could list all 23 (although the maximum recorded in the data is 11).

systematic trends in the returns to particular qualifications. This confirms the results in the table below – there is significant year-by-year variation but no systematic trends. Figure 4 picks out the trends in some of the main variables of interest: GCSE, 2+ A levels and degree which confirms this impression[24].

Table 12 *OLS Returns to Qualifications - Women*

	All	1993	1994	1995	1996	1997	1998	1999	2000
GCSE	0.078	0.112	0.073	0.088	0.076	0.078	0.065	0.063	0.1
	23.5	*11.8*	*8.13*	*9.97*	*8.34*	*9.03*	*7.39*	*7.02*	*5.44*
1 A level	0.141	0.151	0.14	0.132	0.116	0.158	0.156	0.126	0.233
	16.1	*5.92*	*6.2*	*5.52*	*4.92*	*7.06*	*6.39*	*5.28*	*5.01*
2+ A levels	0.254	0.246	0.243	0.28	0.225	0.23	0.281	0.256	0.323
	36.5	*11.65*	*13.1*	*14.9*	*11.6*	*12.7*	*15.4*	*14.5*	*9.21*
Bachelor degree	0.443	0.481	0.451	0.458	0.434	0.437	0.416	0.439	0.482
	98.5	*38.5*	*37.1*	*38.3*	*35.2*	*36.7*	*34.5*	*36.4*	*19.6*
Master degree	0.547	0.565	0.54	0.592	0.553	0.549	0.552	0.493	0.694
	47.9	*15.9*	*16.3*	*17.1*	*17.8*	*18.7*	*19.1*	*18.5*	*11.5*
Age	0	0.0001	0	-0.0002	-0.0004	0.0009	-0.0004	0.0002	-0.0008
	0	*0.17*	*0*	*0.51*	*1.08*	*2.27*	*0.89*	*0.5*	*0.91*
Age squared	-0.0002	-0.0002	-0.0001	-0.0002	-0.0001	-0.0002	-0.0001	-0.0003	-0.0001
	10.7	*3.95*	*2.89*	*5.28*	*3.5*	*4.19*	*2.73*	*5.95*	*1.51*
Married	-0.019	-0.037	-0.035	-0.015	-0.029	-0.016	-0.008	-0.01	0.009
	6.12	*3.96*	*4.09*	*1.77*	*3.3*	*1.95*	*0.95*	*1.19*	*0.51*
Cohabiter	0.031	0.019	0.029	0.034	0.011	0.041	0.029	0.044	0.065
	6.12	*1.18*	*2.02*	*2.41*	*0.81*	*3.21*	*2.17*	*3.42*	*2.41*
Non white	-0.016	0.031	-0.047	-0.02	-0.016	-0.013	-0.017	-0.035	0.099
	1.34	*0.76*	*1.48*	*0.6*	*0.52*	*0.45*	*0.57*	*1.23*	*1.64*
Union	0.219	0.178	0.236	0.222	0.232	0.223	0.219	0.207	0.222
	73.7	*18.3*	*31.1*	*30*	*29.4*	*28.9*	*27.8*	*25.8*	*13.5*
Health problem	-0.104	-0.112	-0.099	-0.13	-0.088	-0.107	-0.117	-0.085	-0.113
	19.8	*5.89*	*5.99*	*7.8*	*6.58*	*9.35*	*8.55*	*6.58*	*4.19*
R squared	*0.33*	*0.34*	0.35	0.34	0.32	0.33	0.32	0.32	0.34
N	80172	10216	10804	11128	11461	11458	11384	11064	2657

Note: Specifications also includes year dummies (all column), region dummies, dummy variables for NVQ 1 and 2, and NVQ 3 and 4, dummy variables to control for discrepancies between qualification and years of education. t-values are in italics.

Table 13 looks in more details at the effects of a degree. Here we use the subsample of individuals who have at least two A-levels (on the grounds that these could have attended University had they wished), and decompose the degree holder by the broad subject of their degree. There is substantial variation across degrees with Education[25] and Arts subjects having no significant effect relative to 2+ A Levels. Languages and Economics/Business, Architecture, Law have returns of around 20%, while Health commands 24%, Maths 19%, Engineering 11% and Science 7%. Again

[24] In each case, adding years of education to these specifications makes no difference to the results here and the education years coefficient is small, positive, but statistically insignificant. The implication of this is that years of education beyond those required to achieve the individual's highest qualification contributes little to human capital and hence wages.

Figure 4 OLS Returns to Qualifications Relative to None – Women

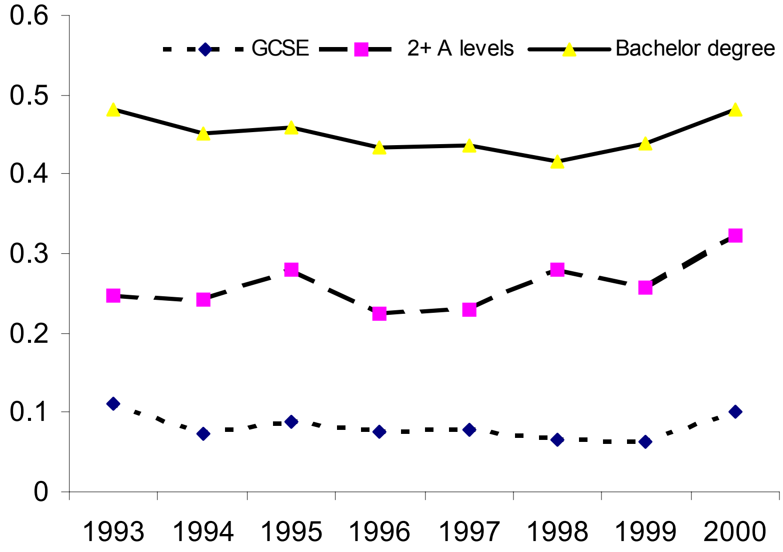


Table 13 OLS Returns to Degree by Subject - Women

	All	1993	1994	1995	1996	1997	1998	1999
Language	0.197	0.234	0.248	0.212	0.207	0.226	0.117	0.178
	16.9	6.92	7.59	6.84	6.96	7.45	3.58	5.58
Health	0.235	0.265	0.262	0.183	0.261	0.258	0.187	0.271
	10.7	3.8	3.84	2.44	3.48	6.74	3.82	5.5
Nursing	0.119	0.183	0.025	0.111	0.109	0.085	0.193	0.113
	5	2.03	0.3	1.72	1.87	1.44	2.91	1.99
Science	0.067	0.069	-0.009	0.058	0.082	0.113	0.122	0.067
	4.32	1.58	0.24	1.44	1.93	2.59	2.9	1.63
Maths	0.188	0.105	0.21	0.242	0.136	0.137	0.295	0.146
	8.4	1.74	3.17	4.27	2.15	2.49	4.86	2.29
Engineering	0.105	0.148	0.192	-0.096	0.114	0.04	0.083	0.131
	2.92	1.32	2.14	0.92	1.2	0.41	0.8	1.5
Architecture	0.212	0.217	0.11	0.152	0.246	0.145	0.357	0.273
	4.69	1.88	0.85	0.09	2.03	1.19	2.89	2.49
Economics	0.198	0.173	0.116	0.127	0.256	0.161	0.273	0.239
	12.9	4.04	2.57	3.21	5.89	3.96	6.27	5.95
Law	0.241	0.197	0.192	0.186	0.248	0.185	0.416	0.257
	10.1	2.75	2.77	3.06	3.68	3.19	6.45	4.34
Education	0.017	0.048	0.01	0.028	0.021	-0.021	0.027	0.011
	1.24	1.19	0.26	0.76	0.59	0.59	0.72	0.29
Social	0.042	0.008	-0.039	-0.016	0.034	0.08	0.1	0.099
	2.78	0.18	0.88	0.38	0.79	2.1	2.52	2.52
Arts	-0.024	0.004	-0.031	-0.039	-0.02	-0.023	-0.019	-0.023
	1.57	0.08	0.69	0.93	0.49	0.58	0.46	0.57
Combined	0.051	0.072	0.016	-0.015	0.078	0.032	0.073	0.062
	4.42	2.18	0.49	0.49	2.67	1.09	2.3	1.96
R squared	0.2	0.22	0.25	0.23	0.21	0.22	0.21	0.21
N	14823	1669	1776	1915	2095	2194	2259	2355

Note: Specifications also includes year dummies (all column), region dummies, dummy variables for NVQ 1 and 2, and NVQ 3 and 4, dummy variables to control for discrepancies between qualification and years of education.

[25] These are individuals who have taken an undergraduate degree in Education, not a PGCE which is treated here as a postgraduate qualification.

these results are not very stable over time, partly because some of these groups are rather small (which explains their low *t* values), but there are no systematic trends in any subject as can be seen in Figure 5. Nor is there any tendency for the returns to each subject to converge – which probably reflects the fact that UK higher education uses the A-level score to ration the supply rather than adjust the supply to match demand.

Figure 5 OLS Returns to Degrees – Women

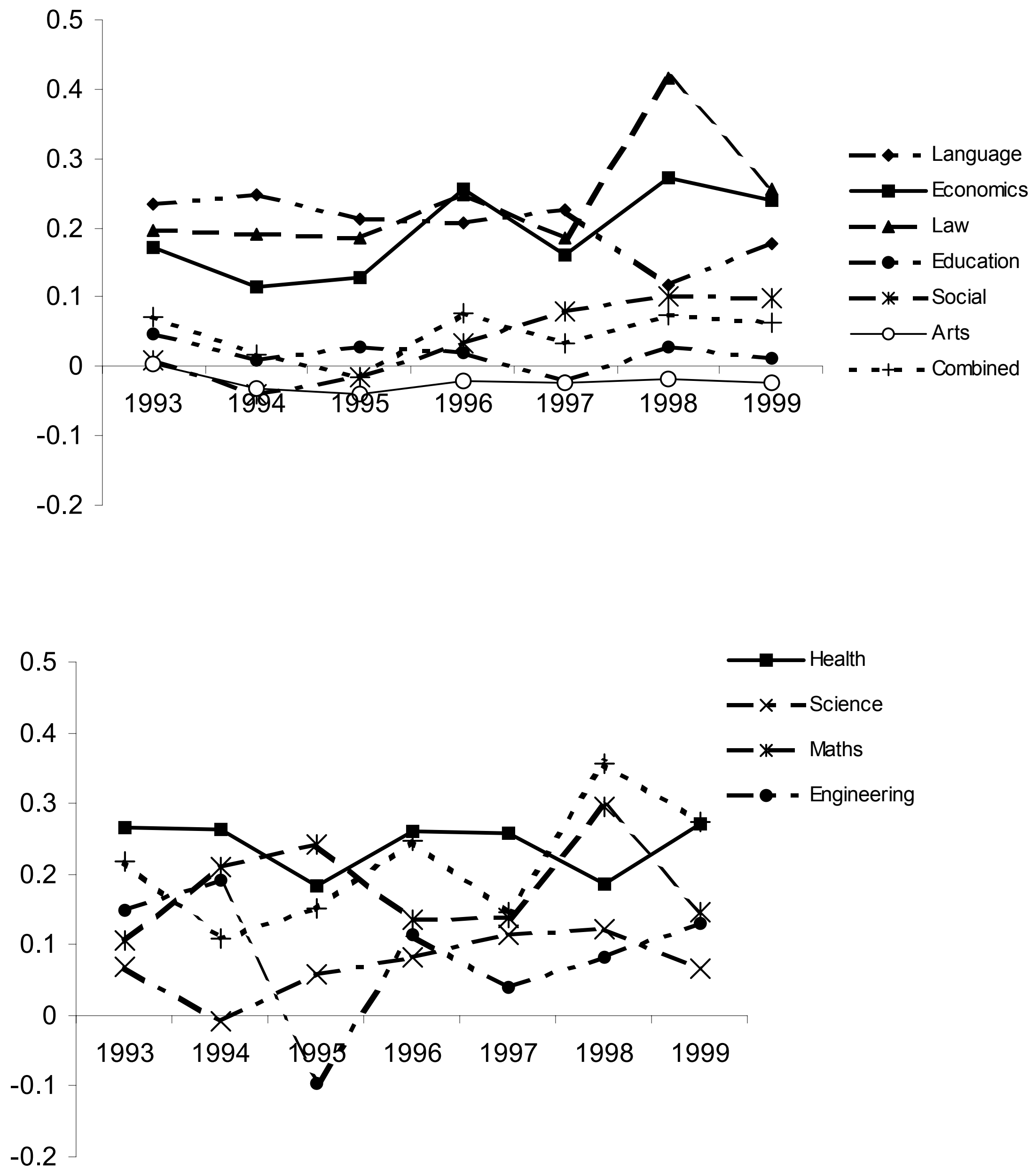


Table 14 shows the same basic results for men. The effects of characteristics are the same as the earlier results suggested. The returns to GCSE are approximately 10%; 2+ A-levels add a further 23%, and a degree another 15%[26]. The pooled results are all highly statistically significant while the year-by-year results show that the returns exhibit no systematic time trends. The returns to the principal qualifications – GCSE, A-Levels and Degree – are very similar to those for women with the A-Level effect marginally, but not statistically significantly, larger for men. There are no significant year on year differences in the main coefficients of interest – as shown in Figure 6.

Table 14 OLS Returns to Qualifications – Men

	All	1993	1994	1995	1996	1997	1998	1999	2000
GCSE	0.104	0.119	0.155	0.122	0.096	0.101	0.074	0.077	0.088
	24.4	*9.65*	*13.5*	*10.7*	*8.65*	*8.82*	*6.38*	*6.65*	*3.71*
1 A level	0.238	0.277	0.274	0.271	0.235	0.209	0.235	0.162	0.245
	22.9	*9.97*	*10.4*	*10.3*	*8.35*	*7.77*	*7.69*	*5.64*	*3.99*
2+ A levels	0.339	0.364	0.348	0.366	0.319	0.342	0.323	0.323	0.327
	48.5	*18.5*	*18.4*	*19.3*	*17.9*	*18.5*	*17.3*	*16.6*	*8.18*
Bachelor degree	0.488	0.487	0.507	0.46	0.473	0.492	0.502	0.493	0.509
	104.4	*38.1*	*40.8*	*37.4*	*38.4*	*38.6*	*39.3*	*38.5*	*18.2*
Master degree	0.535	0.549	0.525	0.504	0.512	0.52	0.547	0.558	0.583
	60.4	*21.9*	*19.9*	*21.3*	*21.1*	*22.4*	*23.5*	*25.3*	*12.1*
Age	0.004	0.004	0.005	0.004	0.005	0.004	0.004	0.004	0.003
	27.8	*9.75*	*11.7*	*10.8*	*11.6*	*10.3*	*9.58*	*9.39*	*3.3*
Age squared	-0.0007	-0.0007	-0.0006	-0.0007	-0.0007	-0.0006	-0.0007	-0.0007	-0.0006
	41	*15*	*14.4*	*16.4*	*17.1*	*13.9*	*14.7*	*15.2*	*6.65*
Married	0.128	0.118	0.115	0.119	0.129	0.136	0.138	0.14	0.108
	35.4	*11.7*	*11.9*	*12.5*	*13.4*	*14.3*	*14.2*	*14.5*	*5.14*
Cohabiter	0.063	0.047	0.06	0.064	0.058	0.074	0.062	0.072	0.056
	12.1	*2.96*	*4.14*	*4.44*	*4.15*	*5.39*	*4.61*	*5.4*	*1.97*
Non white	-0.083	-0.046	-0.078	-0.121	-0.029	-0.116	-0.088	-0.076	-0.111
	6.3	*0.99*	*1.92*	*3.32*	*0.79*	*3.78*	*2.65*	*2.27*	*1.55*
Union	0.082	0.061	0.098	0.089	0.096	0.086	0.073	0.067	0.083
	27.8	*6.57*	*13*	*11.8*	*12.5*	*10.7*	*9.12*	*8.27*	*4.73*
Health problem	-0.118	-0.118	-0.142	-0.118	-0.094	-0.104	-0.146	-0.125	-0.115
	21.7	*6.04*	*8.44*	*6.95*	*7.12*	*8.46*	*10*	*8.97*	*3.96*
R squared	0.29	0.3	0.3	0.28	0.29	0.28	0.28	0.29	0.26
N	76498	9817	10358	10827	10901	10870	10758	10497	2470

Note: Specifications also includes year dummies (all column), region dummies, dummy variables for NVQ 1 and 2, and NVQ 3 and 4, dummy variables to control for discrepancies between qualification and years of education.

[26] The results are largely unchanged by the inclusion of years of education – although this variable tends to have a negative coefficient that is on the margin of significance suggesting that education years over and above those necessary to attain the qualifications are negatively correlated with wages. The inference from this might be that a longer time to complete shows a lack of ability of motivation.

Figure 6 OLS Returns to Qualifications – Men

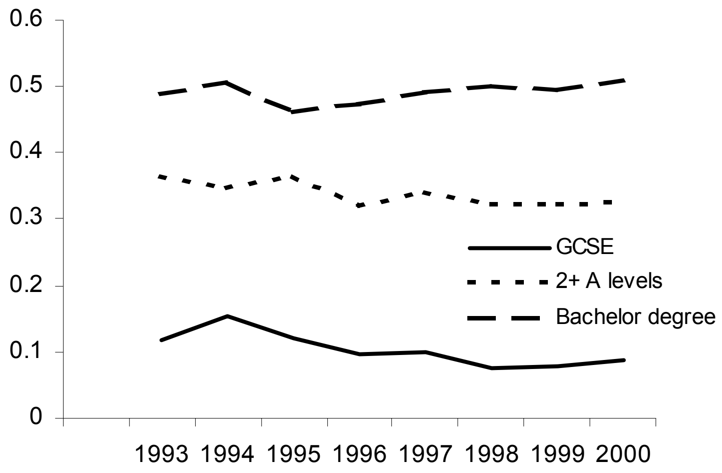


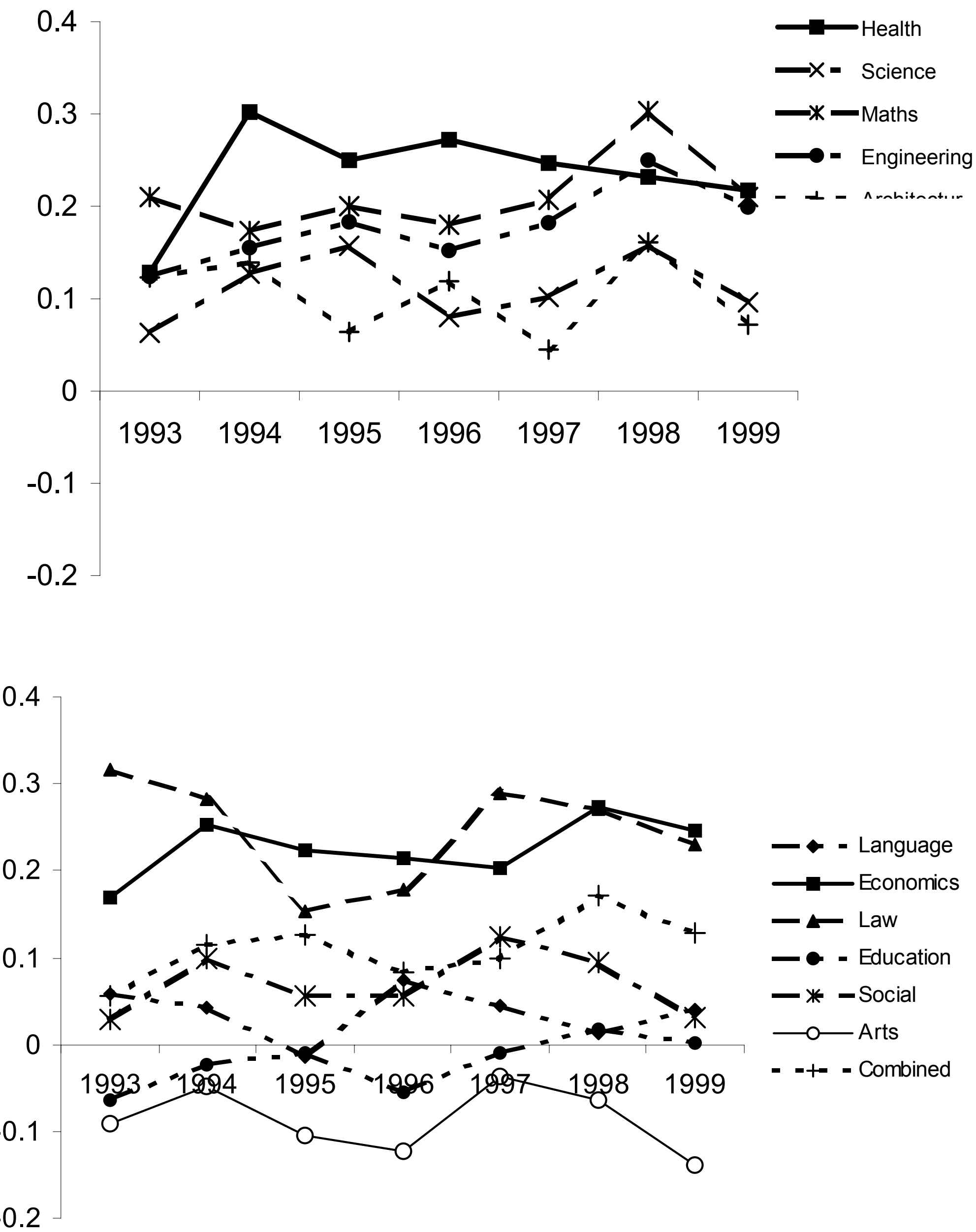
Table 15 OLS Returns to Qualifications – Men

	All	1993	1994	1995	1996	1997	1998	1999
Language	0.197	0.234	0.248	0.212	0.207	0.226	0.117	0.178
	16.9	6.92	7.59	6.84	6.96	7.45	3.58	5.58
Health	0.235	0.265	0.262	0.183	0.261	0.258	0.187	0.271
	10.7	3.8	3.84	2.44	3.48	6.74	3.82	5.5
Nursing	0.119	0.183	0.025	0.111	0.109	0.085	0.193	0.113
	5	2.03	0.3	1.72	1.87	1.44	2.91	1.99
Science	0.067	0.069	-0.009	0.058	0.082	0.113	0.122	0.067
	4.32	1.58	0.24	1.44	1.93	2.59	2.9	1.63
Maths	0.188	0.105	0.21	0.242	0.136	0.137	0.295	0.146
	8.4	1.74	3.17	4.27	2.15	2.49	4.86	2.29
Engineering	0.105	0.148	0.192	-0.096	0.114	0.04	0.083	0.131
	2.92	1.32	2.14	0.92	1.2	0.41	0.8	1.5
Architecture	0.212	0.217	0.11	0.152	0.246	0.145	0.357	0.273
	4.69	1.88	0.85	0.09	2.03	1.19	2.89	2.49
Economics	0.198	0.173	0.116	0.127	0.256	0.161	0.273	0.239
	12.9	4.04	2.57	3.21	5.89	3.96	6.27	5.95
Law	0.241	0.197	0.192	0.186	0.248	0.185	0.416	0.257
	10.1	2.75	2.77	3.06	3.68	3.19	6.45	4.34
Education	0.017	0.048	0.01	0.028	0.021	-0.021	0.027	0.011
	1.24	1.19	0.26	0.76	0.59	0.59	0.72	0.29
Social	0.042	0.008	-0.039	-0.016	0.034	0.08	0.1	0.099
	2.78	0.18	0.88	0.38	0.79	2.1	2.52	2.52
Arts	-0.024	0.004	-0.031	-0.039	-0.02	-0.023	-0.019	-0.023
	1.57	0.08	0.69	0.93	0.49	0.58	0.46	0.57
Combined	0.051	0.072	0.016	-0.015	0.078	0.032	0.073	0.062
	4.42	2.18	0.49	0.49	2.67	1.09	2.3	1.96
R squared	0.2	0.22	0.25	0.23	0.21	0.22	0.21	0.21
N	14823	1669	1776	1915	2095	2194	2259	2355

Note: Specifications also includes year dummies (all column), region dummies, dummy variables for NVQ 1 and 2, and NVQ 3 and 4, dummy variables to control for discrepancies between qualification and years of education.

The breakdown by degree subject for men is shown in Table 15 which again selects those with at least 2+ A levels (we drop (the first quarter of) 2000 because the cell sizes are very small). The large effects of Economics, Law and Health are a feature of the male estimates as well as the female, but the effects of a Language degree are very small and an Arts degree has a statistically significant negative effect. However, the returns to Science and Engineering degrees are higher than for women at 11% and 18% respectively and Maths is slightly higher at 22%. Again Figure 7 confirms that there are no systematic trends in returns by subject nor is there any tendency for them to converge. In fact, the results, at least for large groups, are remarkably stable over time.

Figure 7 OLS Returns to Degree by Subject – Men



It is worth stressing that the estimates in Tables 13 and 15 assume that the choice of degree subject is itself exogenous. It seems unlikely that this is the case since there are quite different entrance standards for different degree subjects so the results here are clearly contaminated by unobservable differences in ability. That is, part of the return to law and medicine is doubtless due to the high average ability of students who take such courses, while science and engineering might be downwards biased because A-level requirements are less stringent. It is also the case that the subject categories are quite broad in some cases and this may conceal some further differences. For example, economics includes management and business studies. Finally some of the groups are quite small. For example, there was no obvious way of grouping architecture students with other groups.

6. The Variance in the Return to Education

Finally, in addition to estimating the *mean* effect of education on wages we also estimate the *variance* in returns around this mean. There are two complementary ways in which we pursue this. In the first method estimation is by "quantile regression" methods which estimate the effect of education on wages at different parts of the wage distribution. The wage distribution reflects not only education but also other, unobservable, skills including ability and social skills. Those at the bottom of the wage distribution are liable to have little education but also a lesser endowment of unobservable skills. Thus it is interesting to ask whether the effects of education are independent of these unobservable skills or whether it compensates for them or complements them. If the effect is independent of unobservable skills then we should find the effect of education is the same throughout the wage distribution; if education compensates for low skill then we should find a larger effect at the bottom of the wage distribution than at the top, and vice versa for complementarity.

We are particularly concerned to show the extent to which, if it is true that the average (unobserved) ability of graduates has fallen over time, then we might see this reflected in the shape of how the returns to education vary by deciles of the wage distribution. If, over time, the expansion of education participation has drawn more and more from the lower end of the distribution of unobserved skills then we might expect to see the return to education at the low end of the distribution fall relative to the top.

The second method estimates a "random coefficients" model. Instead of assuming that the effect of education is the same for all individuals this model assumes that the effect differs (randomly) across individuals. The model estimates the mean effect of education and the variance around this mean. Again, by estimating the models for each separate year it is possible to see if the variance is getting larger over time. The modelling controls for observable differences in returns across individuals.

6.1 Quantile Regression Results

For the reasons stated above, it is possible that the returns to schooling may be different for individuals in the upper part of the wage distribution compared with individuals in the lower portion of the wage distribution. One of the properties of OLS estimation is that the regression line contains, or passes through, the mean of the sample. An alternative methodology to OLS is known as quantile regression (QR) which, based on the entire sample available, allows us to estimate the return to education at any arbitrary quantiles (or even, with a large enough dataset, at any centile) of the wage distribution[27]. The idea behind QR is to look at the returns at one part of the distribution, say the bottom quartile, so as to facilitate a comparison with returns at another part, say the top quartile. The comparison then allows us to infer the extent to which education exacerbates or reduces underlying inequality in wages due to other, perhaps unobservable, factors.

Figure 8 presents, in its final column, the average OLS return to schooling (from FES data for 1980, 1985, 1990 and 1995) together with the returns to schooling in different deciles of the wage distribution[28]. The OLS figures show that over these four half-decades the returns to schooling, on average, broadly increased, between 1980 and 1985 with little change between 1990 and 1995 (confirming our LFS results earlier). There is some indication in this figure that the returns to schooling are higher for those at the top of the wage distribution compared to those at the bottom (although the profiles are flat across a wide range of the wage distribution). Finally, the figure suggests that between 1980 and 1985 the returns at the bottom rose relative to the top, and there is some suggestion, comparing the 1980's with the 1990's, that the returns have risen at the top of the distribution.

One factor behind the distribution of wages is the distribution of inherent ability so that lower ability individuals dominate in the bottom half of the distribution and higher ability in the top half. Thus, the figure suggests that education has a bigger impact on the more able than the less able and this suggests some complementarity between ability and education which seems to have become larger over the 1980's.

Table 16 is based on the work of the international PURE[29] project. In most countries and for most years it would seem that the returns are higher at the top of the distribution than at the bottom. This supports the idea that there is complementarity

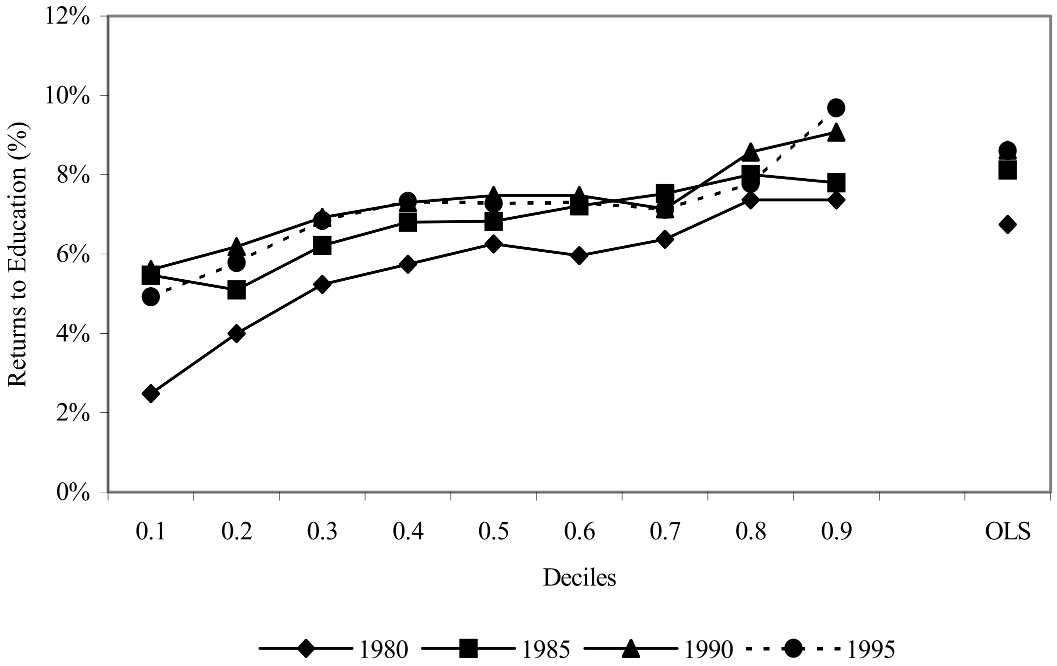
[27] See Buchinsky (1994) and Harmon *et al* (2001) for more details of QR.

[28] These results show the estimated education coefficient for successive cross sections of data in a parsimonious specification that controls just for age, and region. Nevertheless the figures for 90 and 95 are very close to our LFS estimates for men – which were also close to 8%.

[29] Thanks to Pedro Pereira and Pedro Silva Martins for providing this information.

between education and ability. Moreover, the table suggests that, comparing the results for the early 80's with that for the early 90's, this is either getting stronger or, at least, no weaker over time.

Figure 8 Quantile Regressions



Source: FES data for men. See Harmon, Walker and Westergard-Nielsen (2001).

Table 16 Quantile Regressions: Various Countries

	Year	1st decile	9th decile	OLS	Year	1st decile	9th decile.	OLS
Austria	1981	9.2	12.6	10.5	1993	7.2	12.8	9.7
Denmark	1980	4.7	5.3	4.6	1995	6.3	7.1	6.6
Finland	1987	7.3	10.3	9.5	1993	6.8	10.1	8.9
France	1977	5.6	9.8	7.5	1993	5.9	9.3	7.6
Germany	1984	9.4	8.4		1995	8.5	7.5	
Greece	1974	6.5	5.4	5.8	1994	7.5	5.6	6.5
Italy	1980	3.9	4.6	4.3	1995	6.7	7.1	6.4
Ireland	1987	10.1	10.4	10.2	1994	7.8	10.4	8.9
Netherlands	1979	6.5	9.2	8.6	1996	5.3	8.3	7.0
Norway	1983	5.3	6.3	5.7	1995	5.5	7.5	6.0
Portugal	1982	8.7	12.4	11.0	1995	6.7	15.6	12.6
Spain	1990	6.4	8.3	7.2	1995	6.7	9.1	8.6
Sweden	1981	3.2	6.6	4.7	1991	2.4	6.2	4.1
Switzerland	1992	8.2	10.7	9.6	1998	6.3	10.2	9.0
UK	**1980**	**2.5**	**7.4**	**6.7**	**1995**	**4.9**	**9.7**	**8.6**

Source: Harmon, Walker and Westergard-Nielsen (2001).

The weakness of this research reported above is that it is based on small samples since each year of data is limited in size. Thus, below, we apply the methodology to the much larger UK LFS datasets.

The results for men and women in the simple, linear in years of schooling specification[30] are presented in Tables 17 and 19 below. The first point to note is that in 1993 the returns to education for men were broadly the same for the 1st, 3rd, median, 7th and 9th deciles at about 7%. Even though the coefficients are well determined, we could not reject the assumption that the returns were constant across the distribution. For women, the return in the bottom decile, in 1993, was significantly smaller (by almost 2%) than the rest of the distribution where it was around 10%. Over time there was modest drift upwards in the returns for men in the top decile but the rest of the distribution remained flat. There is no evidence that the returns at the bottom of the distribution have fallen. For women, in Table 17, the returns on average fall slightly (as we showed earlier), but the fall seems concentrated in the bottom half of the distribution. Thus, there is weak evidence here that the expansion in education participation has resulted to a reduction in returns at the very bottom of the distribution for women.

In Tables 18 and 20 we see whether there is something similar in the returns to a degree relative to 2+ A-levels. These coefficients come from estimating the earlier specification, with a full set of dummy variables for academic qualifications (plus all of the other control variables). Because these estimates are based on a much smaller sample they are inevitably less stable and one should not read too much into year-to-year changes. For men in 1993 it seems like the returns are substantially higher at the bottom of the distribution (19%) compared to the top (just 6%). This is consistent with the idea that the less able get a higher return to a degree. But, over time, this discrepancy shows some signs of falling because the returns at the top seem to rise – until in 2000 the returns are close to 20% across the distribution. For women, the returns in 1993 are higher than for men, and also much higher at the bottom than the top, but here the rise in the returns at the top that also occurs is mirrored by a fall in returns at the bottom. Thus, while in 1993 there seemed to be much higher returns for both men and women at the bottom of the distribution, by the

[30] But still controlling for gap year, long degree, region, union status etc.

Table 17 Quantile Regressions: Return to Year of Education - Women

Decile	All	1993	1994	1995	1996	1997	1998	1999	2000
0.1	0.069	0.082	0.075	0.079	0.069	0.064	0.066	0.066	0.061
	67	*25.9*	*29.2*	*27.3*	*27.5*	*22.6*	*25.8*	*28.3*	*9.12*
0.3	0.086	0.096	0.088	0.09	0.081	0.082	0.081	0.081	0.081
	122	*47.2*	*43.4*	*44.5*	*41.9*	*39.3*	*37*	*40*	*23.1*
0.5	0.091	0.102	0.094	0.097	0.088	0.087	0.086	0.087	0.087
	128	*42.6*	*44.8*	*42.9*	*43.9*	*42.1*	*41.4*	*45.7*	*25.4*
0.7	0.093	0.102	0.093	0.095	0.095	0.088	0.088	0.088	0.095
	113	*43.8*	*45*	*34.8*	*43.9*	*43.2*	*45.8*	*36.4*	*22.4*
0.9	0.092	0.091	0.09	0.093	0.096	0.09	0.092	0.091	0.102
	84.4	*29.9*	*24.8*	*28.3*	*29.4*	*31.3*	*26.3*	*28.4*	*14.8*

Notes: Figures are coefficients on years of education variable in samples of all workers. *t* values in italic

Table 18 Quantile Regressions: Returns to Degree vs 2+ A Levels – Women

	All	1993	1994	1995	1996	1997	1998	1999	2000
0.1	0.271	0.337	0.363	0.209	0.295	0.327	0.172	0.272	0.206
	18.2	*8.61*	*7.45*	*5.96*	*7.62*	*8.6*	*4.69*	*7.12*	*2.47*
0.3	0.302	0.368	0.341	0.228	0.315	0.348	0.234	0.298	0.219
	31.2	*14.2*	*14.1*	*8.49*	*10.2*	*12.6*	*7.82*	*12.4*	*3.3*
0.5	0.27	0.337	0.28	0.265	0.277	0.258	0.23	0.25	0.262
	28.1	*12.4*	*11.9*	*9.89*	*13.9*	*9.41*	*10.2*	*10.9*	*3.6*
0.7	0.239	0.286	0.229	0.234	0.225	0.208	0.212	0.248	0.175
	22.8	*8.63*	*9.36*	*8.81*	*7.67*	*8.68*	*8.75*	*8.21*	*4.43*
0.9	0.208	0.204	0.187	0.247	0.183	0.232	0.155	0.255	0.201
	16.1	*5.77*	*4.54*	*6.54*	*5.01*	*6.73*	*3.84*	*6.92*	*3.96*

Notes: Figures are coefficients on degree dummy variable in samples with 2+ A levels. *t* values in italic.

Table 19 Quantile Regressions: Return to Year of Education - Men

	All	1993	1994	1995	1996	1997	1998	1999	2000
0.1	0.066	0.071	0.069	0.0655	0.064	0.068	0.062	0.064	0.077
	62.8	*22.6*	*22.3*	*26.4*	*21.9*	*28.6*	*22.2*	*22.3*	*12.3*
0.3	0.074	0.078	0.077	0.074	0.072	0.074	0.074	0.07	0.075
	98.5	*31.3*	*36.5*	*31.8*	*29.3*	*36*	*34.2*	*35.3*	*16.3*
0.5	0.076	0.076	0.079	0.074	0.074	0.076	0.076	0.075	0.076
	97.6	*39.6*	*40.2*	*31.2*	*38.9*	*39.9*	*39.9*	*35.1*	*16.8*
0.7	0.077	0.077	0.078	0.072	0.074	0.079	0.077	0.08	0.076
	101.4	*38.1*	*38.9*	*34.3*	*35.1*	*36.5*	*35.8*	*36.5*	*14.5*
0.9	0.077	0.072	0.076	0.071	0.079	0.078	0.082	0.085	0.086
	79.3	*25.6*	*31.1*	*27.5*	*27.1*	*30*	*32.5*	*27.1*	*14.2*

Notes: Figures are coefficients on years of education variable in sample of all workers. *t* values in italic.

Table 20 Quantile Regressions: Returns to Degree vs 2+ A Levels – Men

	All	1993	1994	1995	1996	1997	1998	1999	2000
0.1	0.217	0.191	0.217	0.188	0.208	0.203	0.272	0.154	0.18
	15.2	*5.3*	*4.97*	*4.35*	*6.93*	*4.68*	*6.02*	*4.14*	*1.48*
0.3	0.195	0.186	0.184	0.175	0.207	0.194	0.205	0.208	0.224
	21.1	*7.16*	*7.47*	*8*	*10.4*	*10.4*	*9.96*	*7.5*	*4.09*
0.5	0.175	0.159	0.168	NC	0.182	0.189	0.189	0.197	0.23
	19.8	*6.68*	*7.83*		*7.59*	*6.73*	*7.4*	*7.84*	*3.43*
0.7	0.129	0.116	0.136	0.059	0.124	0.129	0.132	0.128	0.212
	15.1	*4.29*	*7.13*	*2.22*	*5.6*	*5.76*	*4.93*	*4.11*	*5.51*
0.9	0.079	0.06	0.089	0.035	0.102	0.049	0.089	0.127	0.188
	6	*1.89*	*2.24*	*1.03*	*3.57*	*1.1*	*1.89*	*3.27*	*2.81*

Notes: Figures are coefficients on degree dummy variable in samples with 2+ A levels. *t* values in italic, NC = failed to converge.

of the decade this seems to have equalised somewhat so that returns to a degree are now broadly constant across the distribution.

6.2 Random Coefficient Results

In estimating the standard model of human capital accumulation it is usual to assume that the return to schooling is constant across individuals. However, investments in human capital are inherently risky for two reasons: first, education is usually discrete and the wages associated with each finite unit are not observable by the individual prior to committing to that unit; and, secondly, the individual does not know, in advance, whether he will "succeed" in that unit. If the returns to education depend, at least in part, on the credentials that the individual attains then the possibility of failing to achieve the standard required to attain a credential implies some risk.

In the work below 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 the return to schooling as a random coefficient. Thus we estimate both the mean return and the variance around this mean.

One motivation for our work is that there has been a rapid expansion in participation in post-compulsory education. Earlier work has sought to estimate the mean return to education over time to investigate if this expansion in the supply of skilled workers has resulted in the return to skill falling. The weight of evidence suggests that the increase in supply has broadly matched the increase in demand so that there has been no tendency for the mean return to fall[31]. However, if the increase in supply has been brought about by dipping further into the ability distribution, and if the returns to education arise from signalling innate ability, then we should observe a rise in the variance in returns as more and more low ability individuals acquire the signal. The same would also be true if innate ability and human capital were complementary.

[31] For the US see, in particular, Blackburn and Neumark (1993), Murnane *et al.* (1995) and Cawley *et al.* (1997). Despite the large expansion in post-compulsory education that has occurred in both the UK and US, and many others, there seem to be no estimates that show a *statistically significant* decline in returns. See Goux and Maurin (1994) and Jarousse (1988) for a review across a number of countries.

We specify the basic Mincer-type earnings function as $Y_i = (\beta + \varepsilon_i) S_i + \mathbf{X}_i'\boldsymbol{\gamma} + u_i$ where Y is the log wage, $\mathbf{X}$ is a vector of explanatory variables, including a constant term and a quadratic in age to proxy for experience, and S is years of schooling, and u captures unobservable determinants of wages. We explicitly allow the individual specific coefficient on schooling to be a random parameter so that β is the mean return. This is clearly equivalent to $Y_i = \beta S_i + \mathbf{X}_i'\boldsymbol{\gamma} + e_i$ where the variance in the new error term, e, is proportional to schooling. Thus, this is a specific example of the general heteroscedastic model. Note that the estimation method distinguishes between the "dispersion" in the returns around β from the sampling error of β which we refer to in the tables as the "standard error" of β.

Our analysis presumes that schooling is exogenous. While there is some evidence[32] for the US and the UK that instrumental variables result in larger estimates than does least squares our concern here is with the dispersion of returns. Table 21 presents results from OLS and the random coefficients (RC) models for men and women in the pooled LFS 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 nine successive years), marital status (married or cohabiting versus divorced, widowed, separated and never married), ethnic background (white versus non-white), and union membership (member versus non-member). In addition to the direct control for year of schooling in this specification we also include interactions of schooling with the other covariates to allow the return to schooling to vary by observable characteristics. Our data does not really allow us to explore the endogeneity of schooling since it contains no instruments for schooling and we reserve this extension for future work[33].

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: for

[32] In addition to the review above, see the studies reviewed in Card (1999) and the meta-analysis in Ashenfelter, Harmon and Oosterbeek (1999).

[33] The LFS data contains wages only since 1993 and the post 1993 data contains very few individuals who completed their schooling before 1947 when the school leaving age was raised from 14 to 15. While there are a large proportion of observations whose education post-dates the second school leaving age increase (in 1974) Harmon and Walker (1995) finds that it is the first reform that has the biggest effect on schooling.

example the return to nonwhites is substantially smaller, while the returns to union members is (contrary to our expectations) substantially larger. If we average across all individuals then the estimated (mean) return is 8.2% for men and 8.9% for women. These results change little when we use RC to estimate the return - the returns for women rises only slightly. Our estimate of the dispersion in the return to schooling is about 4% for men and 3.3% for women. That is 95% of men have returns in the +/- 8% interval around the mean, while the dispersion for women is lower with 95% of women within +/- 6.6% of the estimated mean. Thus the dispersion is large, even though we have allowed for differences by observable characteristics.

Table 21 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*
S	4.03	*0.23*	4.13	*0.45*	7.43	*0.24*	8.62	*0.62*
S * Age /100	0.22	*0.03*	0.22	*0.07*	-0.01	*0.04*	0.04	*0.07*
S * $Age^2/100^2$	-0.43	*0.19*	-0.27	*0.39*	-0.96	*0.20*	-1.45	*0.37*
S * Married/100	0.53	*0.16*	1.24	*0.39*	0.50	*0.16*	0.37	*0.30*
S * Cohabiter/100	0.02	*0.24*	-0.09	*0.48*	-0.04	*0.25*	-0.16	*0.43*
S * Nonwhite/100	-2.45	*0.27*	-3.82	*0.59*	-3.88	*0.30*	-4.07	*0.66*
S * Ill health100	-0.31	*0.26*	0.31	*0.54*	-0.32	*0.27*	-0.86	*0.52*
S * Union/100	1.90	*0.11*	1.95	*0.22*	1.01	*0.12*	0.58	*0.37*
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*
Married	0.13	*0.01*	0.11	*0.01*	-0.05	*0.01*	-0.04	*0.01*
Cohabiter	0.08	*0.01*	0.09	*0.02*	0.02	*0.01*	0.03	*0.02*
Non white	-0.18	*0.02*	-0.15	*0.03*	0.02	*0.02*	0.01	*0.03*
Ill 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*
Dispersion parameter	-		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. The default individual is single, non-union, white and with no health problems.

We plot the estimated mean return and its dispersion (which we label "Risk" – because one can think of this as the uncertainty that individuals face when making human capital investment decisions) in Figures 9 and 10. 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.

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 and ends up just slightly lower at the end of the period relative to the start. For women the returns behave quite differently with a downturn in the return to schooling, albeit insignificant (in contrast to the earlier OLS results which were significantly lower), in the later period in both OLS and RC. In contrast the dispersion parameter has fallen slightly over time from around 4% and around 3%.

Figure 9: Year-on-Year Estimates of the Return to Schooling for Men: OLS and RC

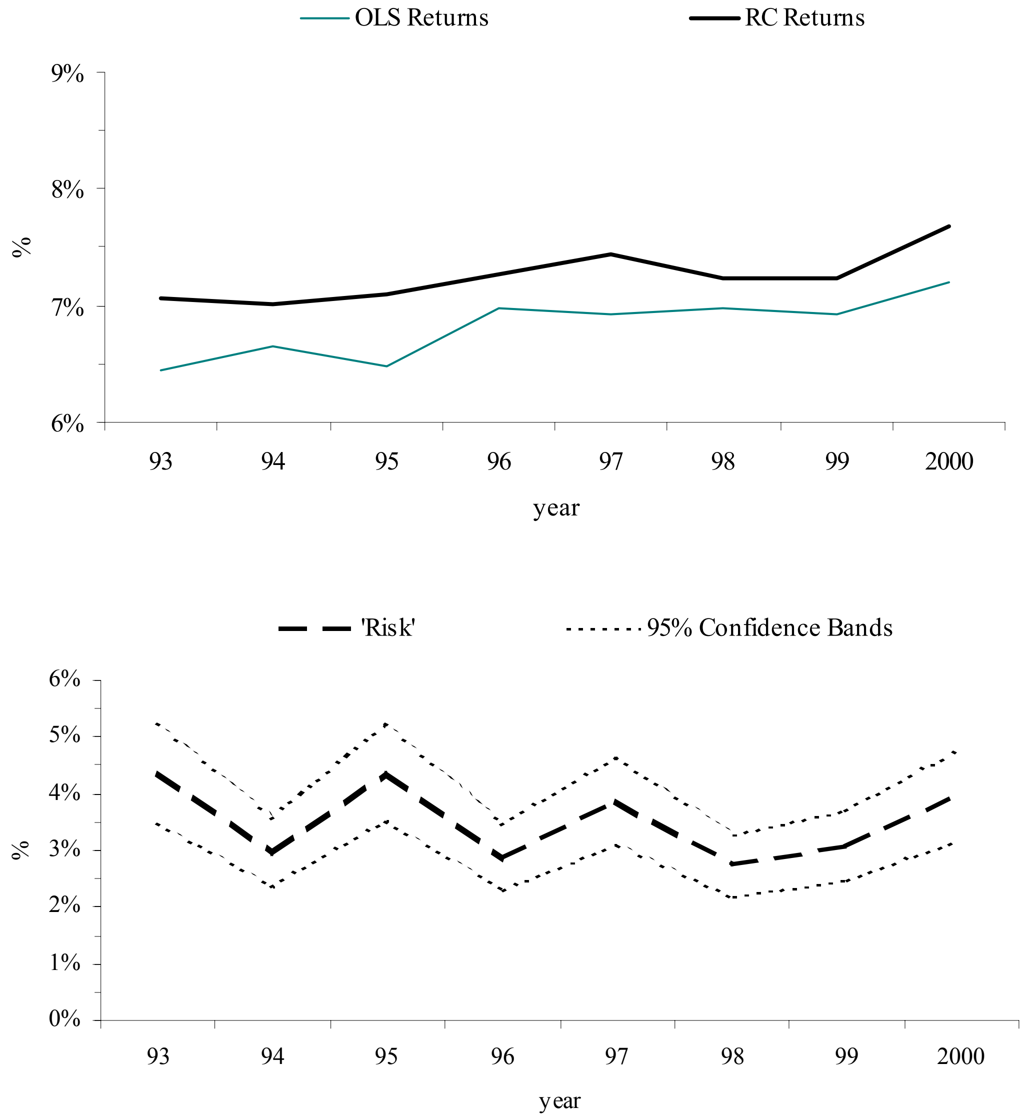
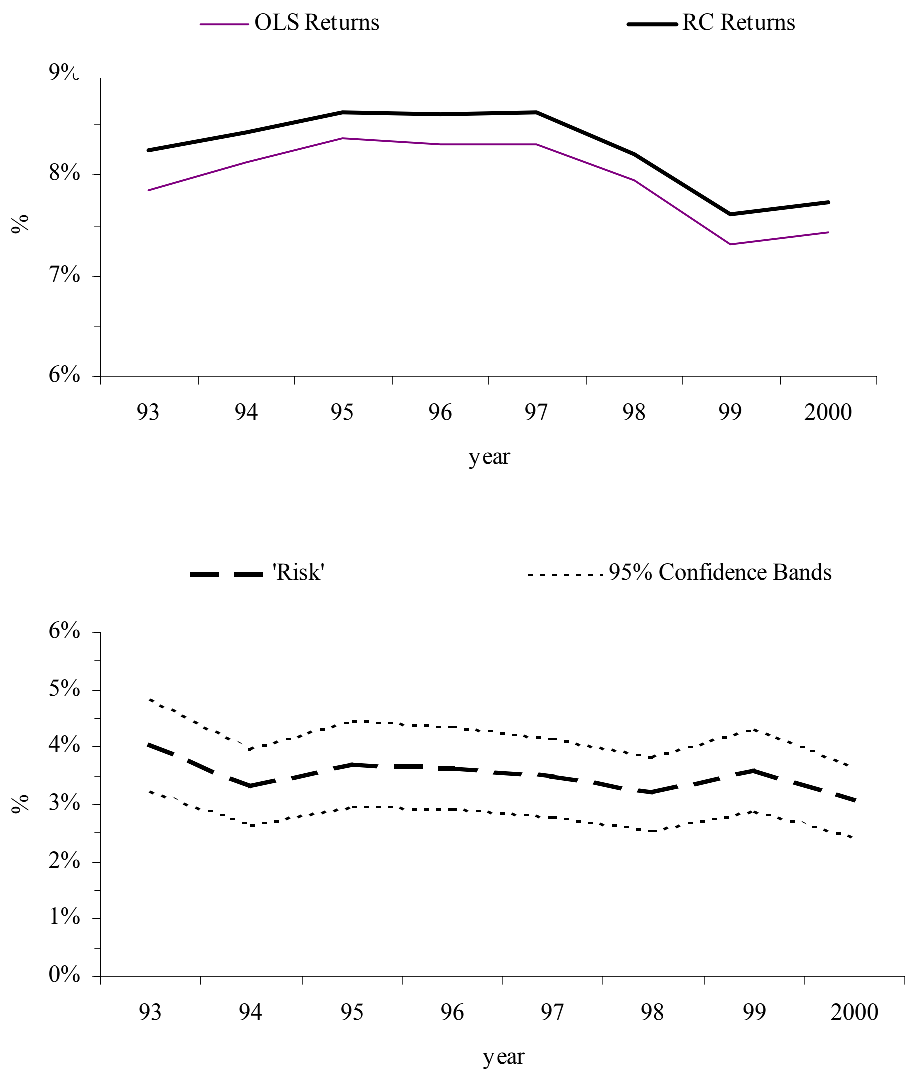


Figure 10: Year-on-Year Estimates of the Return to Schooling for Women: OLS and RC



This analysis is motivated by the concern that the expansion of participation in post-compulsory schooling might result not just in a reduction in the mean return to schooling but, in a model where the return differs across individuals, it may also lead to a longer tail of low return individuals. That is, the expansion of education may have led to a reduction in the average unobserved ability of educated individuals. The expansion in education participation that occurred in the mid-to-late 1960's was undoubtedly fuelled by much greater participation by the children of lower class parents, but there has been some concern expressed that the 1990's expansion in post-compulsory education has been drawn disproportionately from middle class households. Since the additional participation is drawn from a group, which already has high participation, then we might therefore expect that these would be lower ability than had been the case in earlier years. Whereas if the expansion in post-compulsory education comes about through policies that relax credit constraints then there would be every reason to hope that this would increase the average ability of the pool of educated workers. Our estimates suggest that, thus far, the dispersion in returns, due to unobservable effects, has not risen over time and has possibly fallen for women.

We can use these dispersion estimates to examine how many individuals have estimated returns below some critical value. Despite the stability in the dispersion parameter there is some tendency towards greater overall dispersion because of the effect of observable characteristics which have changed over time in a way that adds to the overall variance in returns. In particular our results in Table 21 suggest that, for both men and women, returns are lower for single, for non-white, for non-union, and for those that self report that ill health affects their work. In fact all of these factors have become more prevalent over the course of the 1990's: for example, the proportion married has fallen from 73% to 66% amongst workers in LFS across the period and union membership has fallen from 39% to 32% for men and 32% to 28% for women. Thus changes in the observable characteristics of workers over time has been responsible for the growth in dispersion in returns – unobservable factors appear not to have changed significantly.

Thus, in Figure 11 and 12 we present the whole distribution of returns and can show the proportion of individuals who have returns less than 6% for random coefficient estimates, allowing for differences in both observed and unobserved reasons, and which allow the mean return to vary across time. There is a slightly larger number of men who have returns less than 6% (because their dispersion parameter is larger). This figure rises from around 11% in 93/4 to 26% in 99/00, largely due to the changes in the observable characteristics. For women there is a smaller increase in the variance of returns as can be seen in Figure 12. The proportion with returns less than 6% in 93/4 is only 5% that rises to 18% in 99/00, again largely because of changes in observable characteristics.

Figure 11 Distribution of Returns - Men

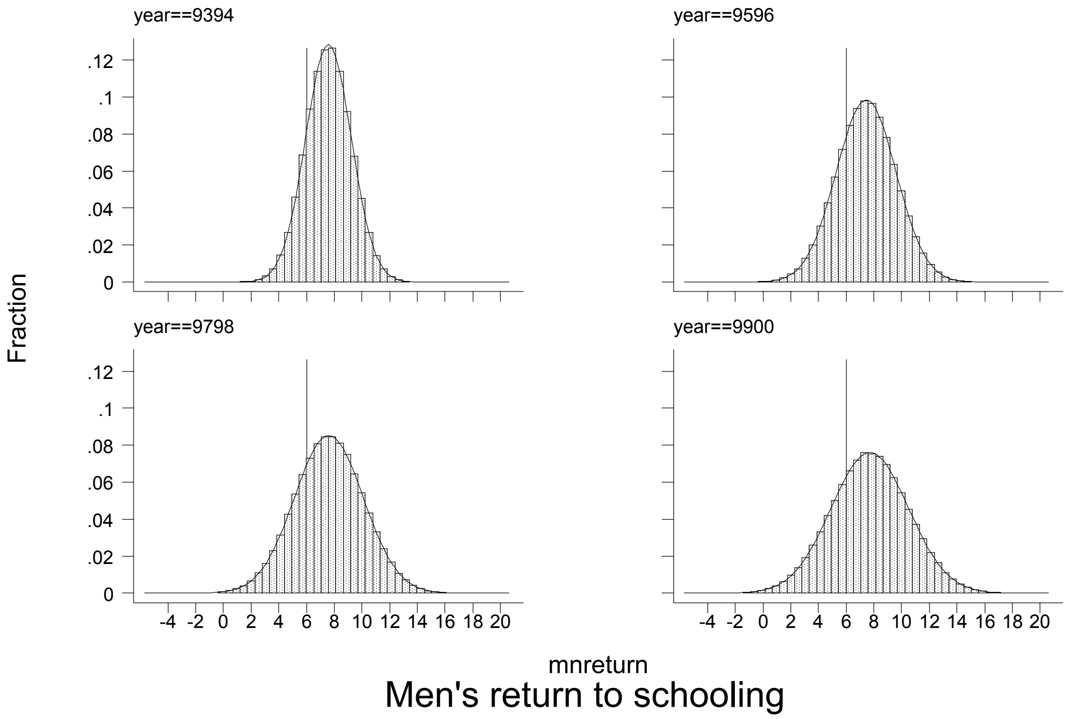
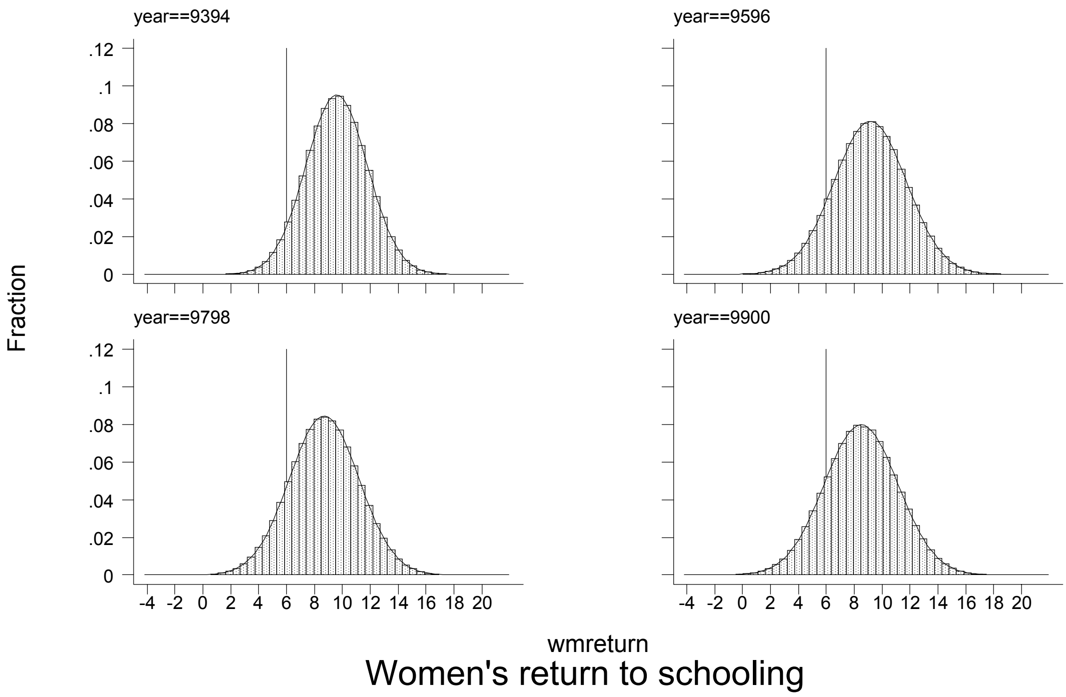


Figure 12 Distribution of Returns - Women



7. Conclusion

This report has looked at the evidence that can be gleaned from the UK LFS data on the returns to education. The results suggest that the returns to education are high and reasonably stable over time. In particular, the returns to the main academic route through GCSE, A-levels to a university degree are large and stable for both men and women. If anything, the returns to qualifications look higher for men than women.

The story on how the variance in returns has changed over time is more complicated. The QR results seem to show that the returns to women were broadly constant across most of the distribution back in 1993 but a little lower at the bottom and, even over a short span of time, there has been some change in this. Returns as a whole seem to have fallen a little but the fall has been larger at the bottom of the distribution. The RC results for women are broadly consistent with this: the average return falls a little, and there is a large dispersion in returns due to unobservables but this does not appear to get any larger over time. Rather, the interaction effects on the observable variables gets larger and this accounts for the growth in the tail of low return individuals. Thus, the expansion in education participation has not made the situation any worse.

The QR results for men suggest that only at the top of the wage distribution has there been much of an increase in returns – something that is unlikely to be explained by the recent growth in participation. In the RC modelling the variation for men for unobserved reasons shows no time trend, but is itself volatile. Again the main factor behind the growth in the overall dispersion is the changes in the coefficients on interactions between education and observable characteristics that has occurred over time.

Finally, it is worth bearing in mind that the LFS results highlighted in this report treat education as exogenous. While it is difficult to deal with this problem effectively in LFS data it should be noted that the majority of previous research, including work for the UK, suggests that doing so would result in higher estimated returns. It is likely to be the case that these higher results measure, not the return to the average person, but the return to the individual most affected by the reform used to instrument education in the specific study. Since most reforms have aimed at increasing education participation amongst those most likely to leave school at an early age, it seems likely that the IV results in the literature are relevant to low education individuals. However, whether such findings of higher (marginal) returns to education would apply in the random coefficients or quantile regression methods remains to be seen – there are few studies yet available, none for the UK, and this should be a priority for future research.

References

Ashenfelter, Orley and Alan Kreuger (1994), "Estimates of the Economic Return to Schooling for a New Sample of Twins", *American Economic Review* (84): 1157-73.

Ashenfelter, Orley and Cecelia Rouse (1998), "Income, Schooling, and Ability: Evidence from a New Sample of Identical Twins", *Quarterly Journal of Economics* 113:253-284.

Ashenfelter Orley and David Zimmerman (1997), " Estimates of the Return to Schooling from Sibling Data: Fathers, Sons and Brothers." *Review of Economics and Statistics* 79:1-9.

Ashenfelter, O., Harmon, C. and Oosterbeek, H. "Estimating the Economic Return to Schooling - A Meta Analysis", *Labor Economics*, 1999, 6, 453-70.

Becker, Gary (1964), *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*, (Columbia University Press, New York).

Bell, David (1997), "The Performance of Immigrants in the UK: Evidence from the GHS", *Economic Journal*, 107.

Berhman, T, Rosenzweig, M, and P Taubman, (1994), "Endowments and the Allocation of Schooling in the Family and in the Marriage Market: The Twins Experiment," *Journal of Political Economy* , 102, 1131-1174.

Blackburn, M.L. and Neumark, D. "Omitted-Ability Bias and the Increase in the Return to Schooling." *Journal of Labor Economics*, 1993, 11, 521-544.

Blackburn, M. and David Neumark (1995), "Are OLS Estimates of the Return to Schooling Biased Downward?" Review of Economics and Statistics, 77(2), May: 217-30.

Blundell, Richard, Dearden, L., Goodman, A., and H. Reed, (1988), "Higher Education, Employment ad Earings", IFS Report.

Bonjour, Dorothe; Jonathan Haskel and Denise Hawkes (2000), "Estimating Returns to Education Using a New Sample of Twins." Mimeo, Queen Mary and Westfield College.

Bound, J., Jaeger, D.A. and Baker, R. "Problems with Instrumental Variables Estimation when the Correlation Between the Instruments and the Endogenous Explanatory Variable is Weak." *Journal of the American Statistical Association*, 1995, 90(430), 443-50.

Bound, John and Gary Solon (1998), "Double Trouble: O the Value of Twins Based Estimation of the Returns to Schooling", NBER Working paper 6721.

Brown, Sarah and John Sessions (1998) "Education, Employment Status and Earnings: A Comparative Test of the Strong Screening Hypothesis", *Scottish Journal of Political Economy* 45(5):586-91.

Buchinsky, M. (1998) 'The Dynamics of Changes in the Female Wage Distribution in the USA: A Quantile Regression Approach ', *Journal of Econometrics*, 13, 1-30.

Butcher, Kristin and Anne Case (1994), "The Effect of Sibling Composition on Women's Education and Earnings", *Quarterly Journal of Economics* 109:531-563.

Cameron, Stephen and James Heckman (1998), "Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males.", *Journal of Political Economy,* 106(2), April: 262-333.

Card, David (1995), "Earnings, Schooling, and Ability Revisited", in Solomon Polacheck, editor, *Research in Labor Economics* Vol.14 (JAI Press, Greenwich Connecticut) 23-48.

Card, D. "The Causal Effect of Education on Earnings." In O. Ashenfelter and D. Card, editors, *Handbook of Labor Economics*, Volume 3b, North Holland, 1999.

Card , David (2001), "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems**", *Econometrica,*** 69.

Cawley, J., Heckman, J. and Vytlacil, E. "Cognitive Ability and the Rising Returns to Education." *National Bureau of Economic Research*, WP6388, 1995.

Dearden, Lorraine (1995) "Education, Training and Earnings" phd Thesis, University College Dublin.

Dearden, Lorraine (1998) "Ability, Families, Education and Earnings in Britain", Institute for Fiscal Studies Working Paper no.W98/14.

Dearden, Lorraine, Stephen Machin and Howard Reed (1997), "Intergenerational Mobility in Britain." Economic Journal;107(440), January: 47-66.

Dearden, Lorraine, Steve McIntosh, Michal Myck and Anna Vignoles (2000). "The Returns to Academic, Vocational and Basic Skills in Britain." DfEE/Skills Task Force Research Report 192.

Dearden, L. "Ability, Families, Education and Earnings in Britain", Institute for Fiscal Studies Working Paper 98/14, 1998.

Ermisch, John and Marco Francesconi (2000), "Educational Choice, Families and Young People's Earnings" *Journal of Human Resources*, 35, 143-176.

Goux, D. and Maurin, E. "Esimations de Fonctions de Gains sur Données de Panel: Endogénéité du Capital Humain et Effets de la Sélection." *Economie et Prévision*, 1994, 116, 1-22.

Griliches, Zvi (1977), "Estimating the Returns to Schooling: Some Econometric Problems." *Econometrica*, 45(1): 1-22.

Griliches, Zvi (1979), "Sibling Models and Data in Economics : Beginnings of a Survey." *Journal of Political Economy*, 87: S37-S64.

Harmon, Colm and Ian Walker (1995), "Estimates of the Economic Return to Schooling for the United Kingdom", *American Economic Review* 85: 1278-1286.

Harmon, Colm and Ian Walker (1999) " The Marginal and Average Return to Schooling in the UK", *European Economic Review* 43(4-6): 879-87.

Harmon, Colm and Ian Walker, (2000) "Returns to the Quantity and Quality of Education: Evidence for Men in England and Wales." *Economica* 67: 19-35.

Harmon, Colm, Westergaard-Nielsen, Niels and Ian Walker (2001), *Education and Earnings across Europe*, Edward Elgar.

Herrnstein, R and C. Murray (1994), *The Bell Curve*. New York: Free Press.

Hildreth, Andrew (1998) “Wages, Work and Unemployment” *Applied Economics*: 30(11) 1351-47.

Hildreth, Andrew (1997) “What has happened to the Union wage Differential in Britain in the 1990s” The Institute for Labour Research, University of Essex discussion paper no. 97/07.

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

Lang, K. (1993), “Ability Bias, Discount Rate Bias and the Return to Education.” Mimeo, Boston University.

Ichino, Andrea and Rudolf Winter-Ebmer (2000), “The Long-Run Educational Cost of World War Two”. Mimeo, EUI Florence.

Isacsson, Gunnar (1999), “Estimates of the Return to Schooling in Sweden From a Large Sample Of Twins”, *Labour Economics*: 6(4) 471-489.

Jarousse, J.P. “Working Less to Earn More: An Application to the Analysis of Rigidity in Educational Choices.” *Economics of Education Review*, 1988, 7, 195-207.

Lam, D. and Schoeni, R.F. "Effects of Family Background on Earnings and Returns to Schooling Evidence from Brazil", *Journal of Political Economy*, 101, 1993, 710-740.

Lang, K., “Ability Bias, Discount Rate Bias and the Return to Education.” Mimeo, Boston University, 1993.

Layard, Richard and George Psacharopoulos (1979) “Human Capital and Earnings: British Evidence and a Critique”, *The Review of Economic Studies*: Vol. 46, No. 3.

Miles, David (1997), "A Household Level Study of the Determinants of Income and Consumption", *The Economic Journal*, 107, 1-25.

Miller Paul, Charles Mulvey and Nick Martin (1995), “ What do Twins Studies Reveal about the Economic Return to Education? A Comparison of Australian and US Findings”, *American Economic Review* 85: 586-599.

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

Murnane, R., Willett, J. and Levy, F. “The Growing Importance of Cognitive Skills in Wage Determination.” *Review of Economics and Statistics*, 1995, 77, 251-266.

Park, Jin Heum (1996), “ Measuring Education Over Time”, *Economics Letters* 60: 425-428.

Pencavel, J. "Assortative Mating by Schooling and the Work Behaviour of Wives and Husbands", *American Economic Review (Papers and Proceedings)*, 1998, 88, 326- 29.

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

Rouse, Cecilia E. (1997), "Further Estimates of the Economic Return to Schooling from a New Sample of Twins", Unpublished Discussion Paper, Princeton University Industrial Section.

Spence, Michael (1973) "Job Market Signalling" *Quarterly Journal Of Economics*:87 (3) 355-373.

Stewart, Mark (1983) "On Least Squares Estimation When the Dependent Variable is Grouped." *Review of Economic Studies*, Vol. 50 (4), pp. 737-753.

Trostel, Philip, Walker, Ian and Woolley, Paul (2001), "The Returns to Education in 28 Countries", Centre for Economics of Education Working Paper, LSE, forthcoming *Labour Economics*.

Weiss, Y. "The Formation and Dissolution of Families: Why Marry? Who Marries Whom? And what Happens upon Marriage and Divorce?" in Mark R. Rosenzweig and Oded Stark (eds.), *Handbook of Population and Family Economics*, 1999.

Willis, Robert (1986), "Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Functions", in Orley Ashenfelter and Richard Layard, editors, *Handbook of Labor Economics*, (North Holland, Amsterdam and New York).