Estimates of the Causal Effects of Education on Earnings over the Lifecycle with Cohort Effects and Endogenous Education*

Giuseppe Migali† and Ian Walker‡

November 8, 2017

Abstract

We estimate the wage premia associated with educational attainments focusing on the lifecycle pattern of earnings. We employ a model where educational attainment is discrete and ordered and log wages are determined by a simple function of work experience for each level of attainment. We distinguish between lifecycle and cohort effects by exploiting the fact that we have a short panel. We find that age earnings profiles lose the traditional bell shape and become flat when we allow for cohort differences. Females still have higher college premia compared to males. However, there are clear earnings inequalities between cohorts with smaller college premia for younger compared to older cohorts, for both males and females.

Keywords: Returns to education, Earnings inequality, Lifecycle effects, Cohort effects, Endogeneity

JEL Classification: I21, J31, C32

*The authors would like to thank the participants to the world meeting of the Society of Labor Economics, the Royal Economic Society conference, the International Workshop on Applied Economics of Education, the European Association of Labor Economics conference, and the Lancaster University Economics seminar for useful comments.

†corresponding author: g.migali@lancaster.ac.uk, Department of Economics, Lancaster University Management School, Bailrigg Lancaster LA1 4YX, UK. Dipartimento S.G.S.E.S., Universita' Magna Graecia, Catanzaro, Italy.

‡Lancaster University Management School and IZA.

1 Introduction

Estimating the returns to schooling is a major industry for applied economists. The methodological difficulty in estimating the causal effect of education is well known: bias (due to ability, school quality, non-cognitive skill) arising from the correlation between wages and the unobservable determinants of schooling contaminates least squares estimates, which can then only be interpreted as an upper bound. There is a large literature that attempts to address this problem using instrumental variable methods by exploiting potential exclusion restrictions. In this context, the simplicity of the workhorse empirical specification of the human capital earnings function is extremely convenient. This specification has log wages being determined by an additively separable and linear function of schooling and a quadratic function of some measure of experience (usually age). Moreover, there is just one variable, schooling, that is endogenous so the search for exclusion restrictions, that has so taxed the ingenuity of researchers in this area, need not continue beyond just one.

This paper allows for the relationship between log wages and schooling to be considerably more complex than the simple human capital earnings function suggests, and yet accommodates endogenous schooling. Our work complements that of Heckman et al. (2008) who compute the internal rate of return (IRR) to the investment in education for different levels of schooling. They start from a general non-parametric approach to the estimation of the determinants of log earnings but do not explicitly allow for endogenous schooling. In contrast to that work, we adopt a parametric model but allow for the selection associated with endogenous schooling. Leaving aside the issue of endogenous schooling, parametric models do have some advantages over non-parametric: they converge faster, they do not require the estimation of smoothing parameters, they are easy to interpret, and parametric estimates can (to the extent that the parameterisation is correct) be used to extrapolate out of sample. On the other hand, estimates of parametric models are conditional on the maintained functional form assumptions. Here we implement what we think of a useful compromise between generality and tractability.

Within the confines of studies that attempt to deal with endogenous schooling, there have been only a few attempts to depart from a simple linear and separable specification. For example, Willis and Rosen (1979) in their structural model treat schooling as a simple endogenous college education dummy variable; while Kenny et al. (1979) exploit the minimum schooling level change and use a Tobit specification for hours; and Harmon and Walker (1995) use an extension of the Heckman two-step approach where the latent variable for years of schooling is treated as an ordered Probit. To the best of our knowledge, all such studies that estimate the effect of endogenous education do so within a model where schooling is some univariate function of years of schooling and assumes that the effect of schooling does not vary across experience - that is, separability between schooling and experience is a maintained hypothesis. This limitation extends to twin studies (see, for example, Ashenfelter et al., 1998) where identification is invariably facilitated by estimating an assumed linear relationship between within twin pair earnings differentials and their schooling difference. Here linearity is crucial because there is typically rather small variance in the within difference in education levels. Moreover, differencing twin data invariably exacerbates the bias in the schooling coefficient arising from measurement errors which is usually dealt with using an instrumental variables approach based on cross-reported schooling.

Adopting the maintained hypothesis of linearity in schooling seems increasingly perverse since there is considerable evidence (see, Heckman et al., 1996; Jaeger and Page, 1996; Hungerford and Solon, 1987) against it. In particular, many studies suggest that the effect of schooling is not linear - that schooling itself is not univariate but rather is, at the very least, best thought of as a succession of levels of achievement that is not simply college vs no college or years of homogeneous schooling. Moreover, many studies show that age earnings profiles are certainly not parallel across education levels (e.g. Neal, 2002; Heckman et al., 2006).

In our selection model education is captured by achievement measures that are ordered: from the lower secondary school level of education associated with the minimum school leaving age (that the US literature thinks of as High School drop-outs), via High School graduation (around the age of 18), through an undergraduate college degree (around the age of 21), and up to postgraduate qualifications (although slightly less commonly so in the UK than the US, according to Lindley and Machin (2016)). We estimate the probability that individuals have a particular level of education by exploiting the fact that they are mutually exclusive and ordered and assume that unobservables are normally distributed. We then estimate age earnings profiles for each education achievement group separately, controlling for selection into each level of educational achievement. Therefore, we do not impose separability, nor do we impose the restriction that the schooling has a linear effect on log wages.

Even this simple departure from the usual separable linear framework comes, of course, at some cost. While the assumption that levels of education of achievement are ordered seems like a natural one, the assumption that the distribution of the residuals in the equation that determines academic progression through the education system is normally distributed is essentially arbitrary. It is only as justifiable as any other parametric assumption and is a long way from the fully non-parametric (or, even, semi-parametric) approach that might be possible to estimate.[1] Of course, adopting normality makes an important contribution towards identifying the parameters of interest in our selection model and sacrificing such a contribution would place a correspondingly greater burden on the validity of the exclusion restrictions. Furthermore, in the case where the treatment is discrete (as in our ordered probit case) the distributional assumptions allow the researcher, in principle, to identify the whole distribution of treatment effects. Thus, while the parametric restriction is strong, it buys the researcher a lot of information. In defence of normality one might argue that unobserved ability is the primary unobserved contribution to the explanation of progression through the education system. While such ability may not be the same as IQ it is likely to be highly correlated with it, and much of the literature suggests that this approximates

[1] See, for example, Blundell and Powell (2004), for a survey of semi-parametric selection models.

a Normal distribution. We view our own approach as a practical compromise that might become a new workhorse specification which could be implemented with many datasets.

Our main finding is a rejection of the traditional linear workhorse specification used in the literature to estimate the returns to schooling. We still find higher returns to education for females over males. Once we control for cohort differences we find that age earnings profiles that are usually thought to peak in late middle age become flat. In particular, younger cohorts have smaller college premia, for both males and females.

This paper is structured as follows: Section 2 describes the data, in Section 3 we explain the estimation method. Section 4 presents the results, and Section 5 concludes.

2 Data

Our aim is to estimate the wage premia associated with educational attainments and we focus on the lifecycle pattern of earnings at different educational attainment levels. Thus, it is important that we estimate the true lifecycle effects net of any cohort and/or calendar time effects. It is, of course, impossible to make such distinctions when using purely cross-section data without imposing some constraints (for example, using cohort controls that are defined over groups of birth years rather than continuously). The distinctive features of this work are: we separately identify lifecycle and cohort effects, by exploiting pooled cross sections and panel datasets to estimate models of earnings and its growth that do not impose the restriction that the age profiles (and the effects of other observables) are the same at each level of education; and we control for selectivity into each education level using a number of exclusion restrictions.

We use the UK Labour Force Survey (LFS), a very large and flexible dataset that is the UK equivalent of the US CPS data, and has been used extensively elsewhere to study the returns to education (see, for example, Walker and Zhu, 2008). The data we use is a sequence of pooled cross-sections over a period of 17 years which offers better prospects, but even here we should expect significant collinearity problems. We exploit the fact that each cross-section has a short panel element to it: individuals are interviewed over 5 quarters and earnings data is collected in the first and final waves - an interval of approximately one year. Estimation exploits the availability of both the cross-section and longitudinal data sets to the full. In particular, we estimate the lifecycle earnings profiles for each education group using the 1997-2014 pooled longitudinal data. Then, controlling for the estimated lifecycle pattern, we estimate the cohort and year effects, together with the impact of education, on the level of wages using the pooled cross sections that are also available from 1997 to 2014. We allow for non-random selection into each education level.[2]

Since we aim to extend the linear workhorse specification that has been used extensively with the labour Force Survey in the UK, and similar datasets elsewhere, we would like to be able to demonstrate the contribution of our extension using this familiar dataset. The LFS is a quarterly sample survey of households living at private addresses. Its purpose is to provide information on the UK labour market that can then be used to evaluate labour market and educational policies. The survey seeks information on respondents' personal circumstances and their labour market status during a specific reference period, normally a period of one week or four weeks (depending on the topic) immediately prior to the interview.

The survey has been conducted on a quarterly basis, with each sample household retained for five consecutive quarters, and a fifth of the sample replaced each quarter. This is known as Quarterly LFS (QLFS) and it was designed to produce cross-sectional data, such that in any one quarter, one wave will be receiving their first interview, one wave their second, and so on, with one wave receiving their fifth and final interview. Thus, there is an 80% overlap in the samples for each successive quarter. The UK LFS has existed since the mid 1970's but it is only since 1993 that data on gross earnings has been collected, and only since 1997 has earnings been recorded in both waves 1 and 5.

In recent years, it has been recognised that linking together data on each individual across

[2]Like the overwhelming majority of the literature we ignore the potential selection issue associated with using a sample of individuals with positive wages. Relatively little research has addressed the issue and our own data does not allow us to pursue this here but we have found little evidence that this is an important empirical issue in practice.

quarters would produce a rich source of longitudinal data, therefore five-quarter longitudinal datasets have also been produced for the same period, for example linking spring 1998 with spring 1999 and containing data from all five waves of the survey. This is known as Longitudinal LFS (LLFS), and because of the resources involved in production and the size of the resultant datasets, the longitudinal datasets include only a subset of the full LFS variable set. In our analysis we exploit both the QLFS and the LLFS. We consider the period 1997-2014 inclusive, and since we focus on the population of working age, the datasets have been restricted to women and men aged 18 to 65 at the first quarter.

Our procedure is the following. We first append all five-quarter LLFS and we obtain a total sample of 524,052 observations.[3] The proportion of employees is around 61%, the self-employed are around 9%, there is a small percentage of people (less than 1%) in government training programs, and the remaining people are inactive in the labour market. We restrict the sample to be employees only.[4] We drop individuals observed only in either the first or the fifth quarter, and employed individuals for whom the earnings variable is missing. We only consider people born between 1940 and 1991. The remaining sample size is 370,570 observations.[5]

We stack all QLFS datasets, from 1997 to 2014, which include around 125,000 individuals per quarter and in each quarter five waves of, on average, 25,036 individuals. If we restrict the sample to the same age range used in the LLFS (18-65) we get around 18,000 individuals per wave in each quarter, and the proportion of employees in each wave is around 61%, and the self-employed are 7.8%. Earnings are collected only in the first and fifth wave, we therefore keep only employees reporting a positive wage (around 98%) in the first wave and

[3]LFS is a panel of addresses not people. Movers are not followed so attrition between waves 1 and 5 accounts for the lower number of cases available for linking and higher attrition, the original five-quarter datasets always contain fewer observations than the QLFS datasets.

[4]The proportion of employees reporting positive earnings in both first and fifth wave is around 98%.

[5]Table A1 in the appendix shows the summary statistics for wage, age and education levels in the unrestricted QLFS sample. Considering that wage is not adjusted to the inflation and we are not restricting to employees only, Table A1 demonstrates that our sample selections do not affect the wage differences across education levels in the data.

the new sample is 530,028.[6]

For practical reasons, we append the resulting LLFS and QLFS datasets and apply further restrictions. The following groups were dropped from the analysis on the grounds that they may have experienced a different education system that is difficult to compare with England and Wales: residents of Northern Ireland (3% in QLFS, 2% in LLFS) and Scotland (9%, in both QLFS and LLFS); people born outside the UK (3% in LLFS, 3% in QLFS); people still in full-time education or never had education; and people that completed their education younger than 14 and older than 30 (0.3% in QLFS, there is no equivalent variable in LLFS). The final sample size is 423,037 in QLFS and 314,055 in LLFS.

The main variables of interest for our analysis are earnings, education and individual characteristics. We constructed our variables in the following ways. Average gross hourly pay[7] including paid overtime. Usual earnings are obtained using information asked directly of all employees and those on schemes, e.g. gross pay before deductions (self-assessed), expected gross earnings (self-assessed). The proportion of non-response to the earnings question is similar by education level and across each LFS data set. Therefore, there is no concern about non-random non-response.

We further restrict the total number of hours worked in the reference pay period to lie in the range 0...94 (losing less than 0.05%). The resulting hourly pay rate is transformed into a real wage rate by dividing by the Retail Price Index (All items) with September 2014 as the base period. The top and bottom 1% of the wage distribution by category of highest academic qualifications were trimmed to avoid outliers arising from measurement error in the wage rate influencing the results unduly. Finally, as usual, we consider the log of the wage rate rather than the wage rate itself.

Our analysis concentrates on education qualifications, rather than the age at which indi-

[6] This sample size corresponds to 1 wave per quarter for 70 quarters from 1997 to 2014. Month of birth is not available for subsequent cross-sections.

[7] This is a derived variable defined as the ratio of usual earnings to usual hours (in main job). The proportion of non-response to the earnings questions is very similar across education levels and we are therefore not concerned about non-random non-response.

viduals leave education. In England and Wales, compulsory education is from the age of 5 to 16 with 5 to 10 being spent in primary school, and 11 to 16 spent in lower secondary education. Students undertake national examinations, typically in five to ten subjects, known as the General Certificate of Secondary Education (GCSEs) at age 16.[8] After the age of 16 they can enter the labour market or continue into post compulsory upper secondary education. Students can choose between academic and vocational qualifications. The academic track consists of GCSEs at 16, followed by A-levels at 18 and university undergraduate degree usually from the age of 19 to 22, possibly followed by a postgraduate degree. There is a very clear ordered progression along this academic education track.

The vocational track is less easy to characterise - typically students would leave formal education at the age of 16 and engage in some occupational training perhaps on a part-time basis while at work, or attend some further education college on a full-time basis for approximately two years gaining vocational qualifications before entering work. Many vocational qualifications are specialised and taken by rather small group of individuals. Fortunately, there is a well-developed method of grouping equivalent qualifications into levels, known as National Vocational Qualifications (NVQ) equivalents. These are defined in Table 1 and divided into five NVQ levels: from NVQ1 (below GCSE qualifications) to NVQ5 (postgraduate level qualifications).[9] In general NVQ3 corresponds to high school graduates and below NVQ3 to high school drop-outs. Table 1 gives examples of the vocational qualifications and their associated NVQ levels, as well as the most common non-vocational ones. We follow established practice in how NVQs are defined with the exception that we pool together the NVQ4 and NVQ5 qualifications due to the small number of observations at NVQ5 level.[10] The interpretation of the estimate of NVQ4 together with 5 is that it is the return to an undergraduate degree including the option value of being able to take a postgraduate degree.

[8] We refer here to the education system after 1973. Prior to 1973 it was common for tracking to start earlier and there was a distinction between the examinations that vocational track pupils took. We convert these older qualifications into their modern equivalents using conventional criteria.

[9] See Makepeace et al. (2003) for the details of how this can be done.

[10] We group NVQ4 and NVQ5 and we refer to this as NVQ4 hereafter.

Our omitted category is no qualifications. Apprentices are recorded as trade apprenticeship (7% of our final sample) and included in NVQ3, although our results are not sensitive to this, or whether we drop them. We generated the educational variables separately and we observed the same proportion in both QLFS and LLFS. Here we report the descriptive statistics that refer to the QLFS sample which is larger and includes all of the individuals in the LLFS. The total sample size of 423,037 comprises 204,735 males and 218,302 females. Women outnumber men in the sample largely because of their higher rate of self-employment.

Summary statistics and the distribution of the earnings, given the NVQ levels, are provided in Table 2. Only those individuals earning a positive wage are included in the sample. We observe that the largest group of individuals (34.4% of males and 34.8% of females) have a NVQ4/5 (NVQ4 around 27% and 28% for males and females respectively, while NVQ5 constitute a further 6.9% and 6.7%). Those with NVQ1 are a very small proportion. The proportion with higher education has grown considerably in recent years following the huge expansion in HE during the 1990's. Note that we also include vocational equivalents in this category. We also notice that the proportion of males with NVQ3 is much higher than that for females. This is probably due to the inclusion of both vocational and academic qualifications, and the percentage of females taking vocational courses is very small compared to males.[11] Comparing wages at NVQ3 and NVQ2, we find that males with the higher qualification earn 12% more than those with NVQ2 and this percentage drops to 7% for females. The wage differential between NVQ4 (pooled with 5) and NVQ3, which broadly corresponds to the "college premium" in the US literature, is around 38% for males and 42% for females.

[11] For a more detailed picture of the problem see Walker and Zhu (2007) who show the NVQ distributions disaggregated by academic and vocational paths.

3 Estimation and Identification

3.1 Estimation method

We adopt a Heckman selection model, extended to allow for an ordered choice across several education levels. Of course, a selection model is more restrictive than IV since it assumes that the distribution of unobservables in both wages and schooling are jointly normally distributed. Our educational variable is ordered - so that NVQ4 corresponds to the highest qualification that can only be obtained if one has achieved NVQ3, and so on. Therefore, we can estimate the probability that individuals have an NVQ at any particular level and exploit the fact that they are mutually exclusive. The main advantage over an IV approach, is that by making an assumption about the distribution of the unobservable determinants of earnings we estimate the effect of qualifications on hourly wages across the whole distribution of the unobservables and, in particular, at the mean. Thus, the selection model method, unlike IV, yields estimates that are comparable to the much simpler least squares regression method.

Our modelling proceeds in two stages. In the first stage, we use a first difference model to estimate the growth in real wages across a year in the panel data. By exploiting the fact that our data is a short panel we are able to estimate lifecycle effects for each level of education separately.[12] Estimating in first differences, using only the panel element of the 1997-2014 data, allows us to model the lifecycle pattern of wages independently of any cohort trends providing those cohort effects are fixed effects in the data - that is, provided that cohort effects in the cross-section equation are additively separable from lifecycle (age) effects. We estimate this difference equation separately for each education level to avoid

[12] The use of a "short" difference in real log wages has obviously more measurement error than a long difference. However, since this wage growth variable is the dependent variable in the first difference model the measurement error should not bias the age coefficient. Furthermore, an advantage of using LFS short difference is that we know that additional work experience is equal to the change in age between waves. Most datasets do not have work experience in them and it is proxied by age. This induces measurement error in the experience measure, and since this is used as an explanatory variable then it would induce bias. Thus it may be better to use a short panel (and risk higher standard errors) to estimate the effect of experience on wage growth, than use a long panel where the much greater measurement error in experience would induce bias.

imposing separability. Implicit in estimating in differences using OLS we presume that unobserved ability affects wages only through the level and not through an effect on growth. This is a common implicit assumption in the literature, but not an innocuous one because ability that is unobserved early in life may become revealed to employers later in life (see, for example, Altonji et al., 2001). However, our attempts to correct for a selection effect in the wage growth equation, in the same way as we approach selection in the wage levels equation, did not suggest that this was a statistically significant issue. Indeed, one might argue that, across the space of just a year it seems reasonable to impose the assumption that the selectivity correction variable has not changed and so is a fixed effect that gets differenced out from the wage growth equation.

Moreover, our model does not impose the restriction that the selectivity works in the same way at all NVQ levels. We do not impose the restriction that the age profiles (and the race effects) are the same at each level of NVQ. So, we allow for non-separability in the earnings function between the schooling effect and the age effect.

In the second stage, we impose the estimated age effect from the first stage on the estimation of the remaining parameters. We do this, by subtracting the age effect from both sides of the log wage level equation before estimation. In the levels equation there is no case for thinking that selection into NVQ level is not an issue, so this stage of modelling is a Heckman two-step procedure that demands the ordered probit selection equation be used to generate the selection correction terms.[13]

Our baseline model is a conventional Mincerian human capital earnings function except that it is non-separable in education level:

$$W_{isq} = \alpha_s + c_{is} + \delta_s Age_{iq} + \rho_s Age_{iq}^2 + \mu_s t_q + u_{isq} \tag{1}$$

where W is the log of the hourly wage rate, $i = 1 \dots N$ indicates individuals, $s = 0 \dots 4$ are the NVQ educational levels; $q = 1, 5$ LFS quarter and $t = 1997 \dots 2014$ indexes years. We

[13] We use the oheckman user-written command in STATA15. See Chiburis et al. (2007).

assume that cohort effects are captured by an additively separable cubic function of year of birth Yob, $c_{is} = c_{1s}Yob_i + c_{2s}Yob_i^2 + c_{3s}Yob_i^3$, and $E[u_{isq}] = 0$. Thus, equation (1) is written for each of five levels of education, and each is modelled as the sum of quadratic age effects, while cubic cohort effects, time effects, and an additive error term.

The estimation of (1) is difficult because individuals' age added to their birth year is identical to the survey year, so that there is an exact linear relationship between the age, cohort, and time effects. However, the interval period between the first and the fifth LFS wave is around 1-year, so we can take the first difference of (1) and remove any time invariant component of the error term, thus controlling for time-invariant unobserved individual heterogeneity. This, together with the pooled nature of the cross-section data allows us to identify age, cohort and calendar time effects. Simply differencing equation (1) we obtain

$$\Delta_q W_{is} = (\delta_s + \mu_s) + 2\rho_s Age_{iq} + \nu_{iq} \tag{2}$$

where the wage growth is linear in age, since the wage level is quadratic in age, for each education classification. Therefore, we are not imposing separability between age and schooling. Lifecycle effects on earnings are given by the constant term in (2), which is a cumulative effect of age and time, and ρ_s which corresponds to the effect of Age^2 in (2).

The identifying assumptions for (2) is that selection bias is driven only by fixed effects and that cohort effects are additively separable in (1). If these assumptions hold, we can obtain consistent estimates of the parameters ρ_s and $\delta_s + \mu_s$ from (2) for all s, but we cannot separately identify δ_s from μ_s unless we are prepared to make some assumption about the value of one of them. For example, one might be prepared to assume that the rate of productivity growth was, say, 0.02 per annum. We then impose these consistent estimates on the cross-sectional log earnings, and estimate a selection model exploiting our pooled data.

In our selection model we estimate an ordered probit as the first step

$$S_i^* = \omega + \psi \mathbf{Z}_i + v_i \tag{3}$$

$$S_i = \begin{cases} 0 & if \quad S_i^* \leq 0 \\ 1 & if \quad 0 < S_i^* \leq \eta_1 \\ 2 & if \quad \eta_1 < S_i^* \leq \eta_2 \\ 3 & if \quad \eta_2 < S_i^* \leq \eta_3 \\ 4 & if \quad \eta_3 < S_i^* \leq \eta_4 \end{cases}$$

We use the estimated ordered probit coefficients from equation (3), to generate the relevant Inverse Mills Ratios (IMRs) to capture the likelihood that an individual has a particular level of education. In the second step, we estimate log earnings for each NVQ level, having imposed the estimates of the lifecycle parameters from (2). That is, we estimate

$$W_{is} - (\widehat{\delta_s + \mu_s})t - \widehat{\rho}_s age_i^2 = a_s + \theta_{is} + \lambda_s \widehat{IMR}_{is} + \beta_s \mathbf{X}_i + \epsilon_{is}. \tag{4}$$

where $\theta_{is} = (c_{1s} - \delta_s)Yob_i + c_{2s}Yob_i^2 + c_{3s}Yob_i^3$, $\widehat{IMR}_{is}$ is the predicted IMR from the ordered probit, $\mathbf{X}$ is a vector of controls, and v_{is} and ϵ_{is} are bivariate Normal with $cov(v_{is}, \epsilon_{is}) \neq 0$.

We include the IMRs in the wage equations (4) to correct for the fact that individuals with a particular level of education will have a particular unobserved component to earnings. We include in $\mathbf{X}$ an ethnic variable, which is also comprised in the ordered probit, and we assume a cohort effect for each education classification. Thus, our final specification allows the intercept and the coefficients on the controls to vary by schooling qualification. We compute the standard errors for (4) by bootstrapping with 200 replications. Our method is effectively a split sample one that adopts the plausible endogeneity methodology in Conley et al. (2012).

As discussed by Heckman (1990) and Card (2012) identification in selection models (as for IV) has to be able to justify the inclusion of variables that affect education that do not also affect earnings directly - the so-called exclusion restrictions. In equation (3) the vector

Z includes at least some variables that are not contained in **X**. In the absence of such, the selection variables will be collinear with the independent **X** variables, and, while we could still use the two-step procedure and obtain the estimates of the IMR, their identification would come only from the distributional assumptions. It is well known that such estimates would be sensitive and would rely exclusively on the assumption that the IMR is a non-linear transformation of the same regressors as in the outcome equations (Heckman, 1979). In particular, the IMR is close to being linear in the absence of a regressor that is a very strong predictor of the dependent variable in the selection equation.

In a traditional two-step selection model with only two outcomes in the participation equation, a standard t-test on the estimate of the coefficient, λ, of the IMR is a valid test of the null hypothesis of no selection bias. In this traditional case we would expect a positive estimate of λ because we expect that more highly educated individuals might earn more because they have unobservable attributes, like ability and perseverance, that are positively rewarded in the labour market and which are positive correlated with education. In our model, with multiple treatments, the IMR represents the correlation between a particular level of education compared to all the others, therefore a significant coefficient can be interpreted as evidence of selectivity but the sign of this coefficient does not have as clear an interpretation as in the binary model.

3.2 Identification approach

Figure 1 shows the evolution of the distribution of highest qualification across cohorts - in groups of 5 years. Births in the 40's were characterised by low levels of qualification, especially for women. Successive cohorts showed higher levels of qualifications with females catching up with, and eventually overtaking, males.

The trends in Figure 1 are confounded by reforms that define our identification strategy which is based on the exclusion from the wage equation (4) of three variables that we think can reasonably be considered to be exogenous and affect wages only through education. The first is the raising of the school leaving age reform (RoSLA): those born before August 1958 faced a minimum school leaving age of 15 years, those born after that date were required to stay in school until at least 16 years of age. RoSLA combines two effects, the first is a quantitative increase in years of education and the second is the attainment of academic qualifications. Harmon and Walker (1995) use RoSLA both as an IV in the first stage linear probability model, and as an exclusion restriction in a nonlinear selection model. Similarly, Oreopoulos et al. (2006) estimate the LATE for secondary schooling exploiting RoSLA as an IV. They find large gains from compulsory schooling. These estimates are not very different from those of US and Canada, although the proportion of people affected by the change in compulsory schooling in the UK was much higher than that in the literature that uses North American data.

To account for selection into higher levels of education and distinguish the effect of qualification we use a second exclusion restriction, the Easter Leaving Rule (ELR), which is still based on year of birth and sets two possible leaving periods. This institutional rule was introduced in the school year 1963/64 and remained valid until the school year 1996/97. Students affected are those born between 1947 and 1980. Precisely: if a student is born between the 1st of September and the 31st of January then she could leave school at Easter of the year she turned 15 (or 16 after RoSLA). If she is born between the 1st of February and 31st of August then she is required to stay until the end of the summer term (last week of May). Since most of the NVQ2 exams take place during the summer term, students constrained to stay longer have higher likelihood to get some academic qualification, especially after RoSLA. Dickson and Smith (2011) using the LFS, exploit both RoSLA and ELR to estimate the effect of education on wage and employment outcomes. They find that most of the returns to RoSLA are due to higher qualifications, although there is a small additional return from the longer length of schooling. In addition, Del Bono and Galindo-Rueda (2006) using LFS data find that individuals leaving after the summer term also experience better labour market outcomes.

The third exclusion restriction is month of birth. It is well documented in the literature that there is an impact of date of birth on cognitive test scores, with the youngest children in each academic cohort year performing poorer, on average, than the older members of their cohort. According to recent studies (for example, Hoogerheide et al., 2007; Kleibergen, 2002) month of birth appears to be uncorrelated with other covariates, unconditionally. Unlike Buckles and Hungerman (2013) we cannot analyze the relationship between month of birth and family background with our data. However, we feel that the weight of evidence, in the UK at least, suggests that month of birth only has indirect effects on log wages - through the level of educational achievement.

Puhani and Weber (2008) use a sample of German children and investigate the impact of age at school entry on test scores at the end of primary school (age 10). They find that children who start school aged 7 rather than aged 6 have test scores that are 0.42 standard deviations higher at the end of primary school. Bedard and Dhuey (2006) use internationally comparable data for OECD countries to estimate the impact of relative age on test scores at ages 9 and 13. They find that children being one month older get higher test score at the age of 9 than at age 13. Ashworth and Heyndels (2007) consider the effects of month of birth in soccer education programs. They find systematic differences in players' performance depending on the months in which they are born. These differences could conceivably produce productivity and wage differences in adulthood. Crawford et al. (2007) is a recent example that notes the relationship between month of birth and educational attainment in the UK.[14] They show that children born later in the school year perform significantly worse in exams than those born earlier in the school year, even up to GCSE

[14]The English rule for admission says that children have to start school at the beginning of the term following their 5th birthday. There are three terms: start September, start January, start April. However, children start at the beginning of the academic year during which they will turn 5 in all Local Education Authorities. So almost all children start school in September whilst aged 4, in what is called the Reception (kindergarten) class. Then they will be aged 5 by the time the school year ends in August 31st and at the start of Year 1. If exceptionally children do not start until age 5, then they will start in Year 1 rather than Reception. And if they start Reception in January or April, the only adjustment is in how much time they spend in Reception. As this class is not so different from nursery school etc, this should not be an issue to use month of birth as exclusion restriction.

level (NVQ2 level). A child born in September will, on average, perform better in academic tests than a child born in the following August, simply because they start school (and sit the tests) up to a year younger.[15] This means that access to further and higher education, and hence future success in the labour market, is likely to be affected by month of birth. In subsequent work, Crawford et al. (2014) show that the majority of the effect of birth timing on outcomes is attributable to the age at which high stakes tests are taken. This begs the question as to whether parents might manipulate the timing of birth to ensure better outcomes. It has become commonplace to directly examine the density of the running variable in RD designs and that has been formalised in the test due to McCrary (2008), and Tables A1 and A2 in the Appendix provide reassurance of the timing of births. Indeed, we never find a significant jump in the distribution of month of birth for those born after September, except for women in the cohort 1951-55. However, there is no concern, since the direction of the jump is actually the opposite of what we would have expected in the case of correlation between month of birth and family background.

All three exclusion restrictions are in principle uncorrelated with the unobservable determinants of the earnings, therefore satisfy the condition of the random assignment to treatment, in terms of the Angrist et al. (1996) causal model. Formal tests of the validity and strength of our instruments suggest that we do not have a weak instrument problem. The LR test for the joint significance of the instruments, allowed us to reject the null. We also performed a two-sample Kolmogorov-Smirnov test for equality of distribution of month of birth and we did not find significant differences by RoSLA. This further suggests that any strategic behaviour by parents did not change when the school leaving age was increased.

In Figures 2 and 3 we show, for males and females respectively, the proportion of NVQ levels by month of birth for four selected 6-year birth cohorts chosen to capture the role of the policy changes that define our instruments: 1940-45 which was pre-ELR and pre-ROSLA, 1950-55 which was during ELR but pre-ROSLA, 1960-65 which was also during ELR but

[15] Crawford et al. (2007) show that September born children have on average 0.2 year more completed education than August born children.

post-ROSLA, and 1980-85 which was when the ELR was abandoned and was post-ROSLA. Note that, for diagrammatic purposes, we omit the 1955-59 cohort since the ROSLA reform occurred for the middle of that cohort, and we omit 1966-1979 births since there were no changes to affect those cohorts.

Cohort trends are also evident in Figures 2 and 3, although it is less clear cut than in Figure 1. Successive panels show marked rises in the overall extent to which individuals are qualified. For example, between the 1940-45 cohort and the 1950-55 cohort, between which ELR occurred, the proportion with no or basic qualifications fell by around 50%. Between the 1950-55 and 1960-65 cohort, between which ROSLA was introduced, NVQ2 rose at the expense of a further fall in those below this level. And between the 1960-65 cohort and the 1980-85 cohort, those with more than NVQ3 rose despite the demise of the ELR.

If we consider just those born five years either side of the ROSLA, to reduce the extent to which there are cohort trends, we find that the reform immediately reduced the probability of leaving school at the old minimum, age 15, from approximately 30% to close to zero and that the probability of leaving school, at the new minimum rose immediately from approximately 30% to close to 60%. The distribution above leaving at 16 remained approximately unchanged (see, also, Chevalier, 2004).

The month of birth effect is less distinct. For each cohort group and both genders there is a slight fall in those lowly qualified across the school year. We expect this effect to be stronger when there was streaming by ability test (called the 11+ examination) administered to all children in the final year or primary schooling at age 10/11. The scoring of this test was not age adjusted and the test occurred on the same day for all children, and approximately the top 20% were admitted to an academic school. There were strong expectations that such an academic track would involve remaining in school through to 18 and an expectation that the best of these students would go to university. This selection test, and the associated segregated schooling, disappeared gradually in the UK over the 1970's and so we might expect month of birth to be less pronounced for the latter two cohorts in Figures 2 and 3 -

and this seems to be the case.

4 Estimation Results

We restrict ourselves to a parsimonious specification that excludes variables such as children and whether married that are likely to be endogenous. However, we include in both the wage and education equations self-reported ethnicity (grouped simply into white and non-white).

In Table 3, we report the results of the first step of the second stage of our estimation approach, i.e the ordered probit selection equation. We include, in this step, a cohort effect represented by a cubic function of the year of birth to capture long run social changes, as distinct from the sharp effect of RoSLA. This social change turns out to be quite significant for educational attainment. The month of birth effect is captured by including a continuous month variable, where September is equal to one and corresponds to the oldest children in each class cohort. We find, as expected, that its sign is negative and highly significant indicating that the oldest children in each class do better.[16] Finally, we also include dummy variables to capture the effects of ELR - which we find also have significant effects on the probability of attaining qualifications.

We show, in Table 4, the corresponding marginal effects of RoSLA, ELR, and month of birth evaluated at each level of education. The effect of RoSLA, measured as the difference between the coefficients after and before, is positive from NVQ0 to NVQ2, and it has its highest values at NVQ0 and NVQ2. The intention of the government and the consequent effect of the policy was to increase the participation at the lower secondary levels of education, and it is here where we find the strongest effect of the RoSLA. Note that if we had used RoSLA as an IV then the estimated effect of the NVQ would be weighted towards those at the

[16]To assess the stability of our exclusion restrictions we estimated the selection model applying one restriction at a time, and we found that the results are substantially unchanged when using only RoSLA , ELR or month of birth. We have also performed pairwise comparison of different education levels by estimating separate probit models. The results show a stronger effect of RoSLA and month of birth at lower levels of education, and a stronger effect of ELR at higher NVQ levels.

bottom of the education distribution - e.g. for those who wanted to leave at 15 and the policy forces to remain in school until 16. In contrast, in our selection model, we estimate the effect on the whole population. The effect of the ELR for the those born between February-August is clearly higher at NVQ4 compared to September-January and to No ELR. Looking at the month of birth effect, we notice that the coefficient is significantly positive and decreasing from NVQ0 to NVQ3, as expected, while it is significantly negative at NVQ4: the youngest children in their school year cohort have a higher probability of achieving only lower levels of education, while the older children in each class have a higher probability of attaining higher levels. Similar results are found for females in Table 5.

In Table 6 we report the results of the estimation of equation (1), for each education level separately and without cohort effects. The age effects are well determined - well enough determined to allow us to reject the imposition of a common age effect across all education levels in the last column, where we present estimates of the separable model by including each of the qualification dummies. This imposes the constraint that the effect of qualifications is merely to shift the age earnings profile in a parallel fashion. A test that the age effects are common across qualification levels is rejected, so we are confident that the conventional separable model can be rejected in favour of our non-separable one.[17]

In Table 7, we present the estimates of the second step of the selection model where again the dependent variable is the level of the log wage and we allow for time effects but not for cohort effects. This is the conventional approach to the estimation of the selection model except that we estimate one equation for each qualification level and correct for selectivity into that qualification level. That is, Table 7 does not impose separability but it does assume that there are no cohort effects. The important finding is that the coefficient of the IMR, λ, is always significant for females, and significant for NVQ0, NVQ3 and NVQ4 for males. This means that our exclusion restrictions are detecting, and correcting for, the presence of selection bias.

[17] We also tested the model using cubic age effects. However, this effect is mostly insignificant and an F test of the significance of the cubic in all NVQ levels rejects this extension.

Table 8 presents the estimates of the wage growth equation (2). The time difference between our wage observations is five quarters.[18] We have therefore multiplied the coefficient of age in this equation by 5/2 to generate an annual age effect. Again we allow for non-separability and find that we are able to reject the null of separability.

We then impose the estimated age effects in Table 9 by adjusting the dependent variable by subtracting the estimated age effect from Table 8. That is, the dependent variable in Table 9 is that shown in equation (4), which we estimate separately by qualification level, correcting for selection into qualification. We again find that selection is jointly significant, for both males and females. In this case, we can include additive cohort effects (not reported) since the age effects are imposed from the first step estimates. We find that these cohort effects are always jointly significant.[19]

In Table 10 we compare the average predicted wages obtained from the two estimated selection models. In the top panel of Table 10 we report the average predicted wages (from the estimates in Table 7), which include age effects but no cohort effects. In the bottom panel we show the predicted wages (from the estimates in Table 9) which do not include age effects, since the log wages have been explicitly corrected for them, but allow for cohort effects. We notice that the predicted wages are higher in the conventional model, and their fall in the new model is of the same magnitude for both males and females. The college premium (NVQ4 minus NVQ3) is line with the UK literature: females get a higher premium of around 43%, while for males the premium is around 38% in the conventional model, and the premia are substantially unchanged in the new model. The returns to NVQ3 versus NVQ2 are much higher for males than females, in both models. This is consistent with our raw data, since we have more males with vocational qualifications than females.

As we stated above, in the selection model we are estimating the effect of education on earnings across the whole distribution of unobservables - in particular, at the mean. In fact,

[18]Figure A3 in the Appendix shows how the wage growth varies by nvq levels and by sex. It is interesting to observe that more educated females have higher growth rates than males.

[19]We tested these effects both by and across NVQ levels.

if we add back to the fitted wages in the new model the age effects estimated in equation (2), we obtain returns to education for the two selection models that are practically the same. However, the fact that, at the mean, the predicted earnings are similar does not imply the absence of differences across the entire distribution. Indeed, to highlight these differences we compute the age-earnings profiles.

Figure 4 shows the profiles obtained from the raw pooled data, that is from the OLS estimation using quadratic age with discrete schooling groups and no cohort effects (see Table 6). These profiles are identical, and we do not report them, to those obtained from the estimation of the conventional selection model with age effects only (see Table 7). In Figure 4, we observe the well-known convex shape of the profiles, where the peak for males is at age 45 with a college premium of around 40%, whilst for females the peak is at 46 years old with a college premium higher than 40%. It is evident that the age-earnings profile for NVQ4 is higher than NVQ3 at all ages, and steeper than NVQ3 at early ages for both males and females. Notice that we find that women with low qualification levels have flatter lifecycle wage profiles.

Figure 5 shows the profiles obtained from the estimation in Table 9 when controlling for life cycle effects but assuming no cohort differences. We observe increasing profiles, which implies strong age effects throughout lifecycle for all educational levels. This clearly contrasts with Figure 4, because now we have profiles where age effects are immune from cohort effects.

Finally, in Figure 6 we show our last set of profiles which combine all our extensions to the simple workhorse specification used in the literature. We consider discrete groups of educational qualifications, we control for lifecycle effects and we separately allow for cohort differences. We find two clear results: the age earnings profiles are now flat and younger cohorts have lower returns to education compared to older cohorts, for both males and females. However, although the college premiums are decreasing in magnitude they are still evident for younger cohorts; on the other hand, the returns to high school, for the same cohorts, appear to be very marginal. [20]

[20] Selection into employment could be a possible cause of the differences across lifecycles between men

We investigate further these differences, by focusing in Figure 7 on the returns to education for high school (NVQ3 minus NVQ2) and higher education (NVQ4 minus NVQ3), for the age groups in the overlapping cohorts showed in Figure 6.

For example, a 44-year-old graduate male (female) from the cohort 1950-55 has a college premium of around 34% (39%) while at the same age a graduate male (female) from the cohort 1960-65 has a premium of 26% (36%). Whereas a 36-year-old graduate male (female) from the cohort 1960-65 has a college premium of around 25% (35%) while at the same age a graduate male (female) from the cohort 1970-75 has a premium of 12% (24%). Looking at the gender differences, females from any cohort have higher college premiums than males; whereas high school returns are much bigger for males than females, because there are more males with vocational qualifications. Overall the lower college premiums for younger cohorts may be due to the higher education expansion in the last decades in the UK, which has increased the supply of college graduates; and recently younger graduates may have suffered more the effects of the Great Recession compared to older cohorts.

5 Conclusion

This paper has proposed and implemented a simple methodology, to estimate the returns to education, that is sufficiently tractable that it could be used with many datasets, and yet provides a significant generalisation of the usual additively separable and linear human capital earnings function. We separately estimated lifecycle and cohort effects, and identification was achieved through exploiting two education reforms and month of birth. All have significant effects on educational attainment. Our results amount to a strong rejection of

and women. For example, in a lifecycle model people with a lower taste for leisure, and therefore work longer hours, will invest in more education. So education is endogenous and the non-workers will have least education, and would have low wages if they were to work. There are other possible reasons for observing different lifetime patterns between males and females. Female participation in the labor market may be discontinuous (women devote time to the household sector) and this may lower their human capital returns. Women also may tend to choose occupations that maximize their lifetime earnings where their skills do not depreciate during the years spent in the household sector.

the simple workhorse specification that is commonly used in this literature. Compared to the conventional models, the returns to schooling are almost unchanged for both genders, and we still find that females have higher college premia than males. Age earnings profiles have the traditional bell shape if we do not control for cohort effects, whereas they become flat when we allow for cohort differences. We also observe substantial earnings inequality between younger and older cohorts. Younger cohorts have lower returns to higher education, for both males and females.

References

Altonji, J. G., Pierret, C. R., et al. (2001). Employer learning and statistical discrimination. *The Quarterly Journal of Economics*, 116(1):313–350.

Angrist, J. D., Imbens, G. W., and Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American statistical Association*, 91(434):444–455.

Ashenfelter, O., Rouse, C., et al. (1998). Income, schooling, and ability: Evidence from a new sample of identical twins. *The Quarterly Journal of Economics*, 113(1):253–284.

Ashworth, J. and Heyndels, B. (2007). Selection bias and peer effects in team sports the effect of age grouping on earnings of german soccer players. *Journal of Sports Economics*, 8(4):355–377.

Bedard, K. and Dhuey, E. (2006). The persistence of early childhood maturity: International evidence of long-run age effects. *The Quarterly Journal of Economics*, pages 1437–1472.

Blundell, R. W. and Powell, J. L. (2004). Endogeneity in semiparametric binary response models. *The Review of Economic Studies*, 71(3):655–679.

Buckles, K. S. and Hungerman, D. M. (2013). Season of birth and later outcomes: Old questions, new answers. *Review of Economics and Statistics*, 95(3):711–724.

Card, D. (2012). Earnings, schooling, and ability revisited. *35th Anniversary Retrospective*, 35:111.

Chevalier, A. (2004). Parental education and child's education: a natural experiment.

Chiburis, R., Lokshin, M., et al. (2007). Maximum likelihood and two-step estimation of an ordered-probit selection model. *Stata Journal*, 7(2):167–182.

Conley, T. G., Hansen, C. B., and Rossi, P. E. (2012). Plausibly exogenous. *Review of Economics and Statistics*, 94(1):260–272.

Crawford, C., Dearden, L., and Greaves, E. (2014). The drivers of month-of-birth differences in children's cognitive and non-cognitive skills. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 177(4):829–860.

Crawford, C., Dearden, L., and Meghir, C. (2007). *When you are born matters: the impact of date of birth on child cognitive outcomes in England*. Centre for the Economics of Education, London School of Economics and Political Science.

Del Bono, E. and Galindo-Rueda, F. (2006). The long term impacts of compulsory schooling: Evidence from a natural experiment in school leaving dates. Technical report, ISER Working Paper Series.

Dickson, M. and Smith, S. (2011). What determines the return to education: an extra year or a hurdle cleared? *Economics of education review*, 30(6):1167–1176.

Harmon, C. and Walker, I. (1995). Estimates of the economic return to schooling for the united kingdom. *The American Economic Review*, 85(5):1278–1286.

Heckman, J. (1990). Varieties of selection bias. *The American Economic Review*, 80(2):313–318.

Heckman, J., Layne-Farrar, A., and Todd, P. (1996). Human capital pricing equations with an application to estimating the effect of schooling quality on earnings. *The Review of Economics and Statistics*, pages 562–610.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica: Journal of the Econometric Society*, pages 153–161.

Heckman, J. J., Lochner, L. J., and Todd, P. E. (2006). Earnings functions, rates of return and treatment effects: The mincer equation and beyond. *Handbook of the Economics of Education*, 1:307–458.

Heckman, J. J., Lochner, L. J., and Todd, P. E. (2008). Earnings functions and rates of return. *Journal of Human Capital*, 2(1):1–31.

Hoogerheide, L., Kleibergen, F., and van Dijk, H. K. (2007). Natural conjugate priors for the instrumental variables regression model applied to the angrist–krueger data. *Journal of Econometrics*, 138(1):63–103.

Hungerford, T. and Solon, G. (1987). Sheepskin effects in the returns to education. *The review of economics and statistics*, pages 175–177.

Jaeger, D. A. and Page, M. E. (1996). Degrees matter: New evidence on sheepskin effects in the returns to education. *The review of economics and statistics*, pages 733–740.

Kenny, L. W., Lee, L.-F., Maddala, G., and Trost, R. P. (1979). Returns to college education: An investigation of self-selection bias based on the project talent data. *International Economic Review*, pages 775–789.

Kleibergen, F. (2002). Pivotal statistics for testing structural parameters in instrumental variables regression. *Econometrica*, 70(5):1781–1803.

Lindley, J. and Machin, S. (2016). The rising postgraduate wage premium. *Economica*, 83(330):281–306.

Makepeace, G., Dolton, P., Woods, L., Joshi, H., and Galinda-Rueda, F. (2003). Changing britain, changing lives. three generations at the turn of the century.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of econometrics*, 142(2):698–714.

Neal, D. (2002). The measured black-white wage gap among women is too small. Technical report, National Bureau of Economic Research.

Oreopoulos, P., Page, M. E., and Stevens, A. H. (2006). The intergenerational effects of compulsory schooling. *Journal of Labor Economics*, 24(4):729–760.

Puhani, P. A. and Weber, A. M. (2008). *Does the early bird catch the worm?*, pages 105–132. Physica-Verlag HD, Heidelberg.

Walker, I. and Zhu, Y. (2007). The labour market effects of qualifications. *Futureskills Scotland Research Series, http://www. scotland. gov. uk/Resource/Doc/919/0065442. pdf.*

Walker, I. and Zhu, Y. (2008). The college wage premium and the expansion of higher education in the uk. *The Scandinavian Journal of Economics*, 110(4):695–709.

Willis, R. J. and Rosen, S. (1979). Education and self-selection. *Journal of Political Economy*, 87(5, Part 2):S7–S36.

Table 1: NVQ Equivalent Qualifications

NVQ level	Examples of academic qualifications	Examples of vocational qualifications
5	PhD Masters degree	PGCE (Teaching) Non-masters postgraduate quals
4	Undergraduate degree	Other teaching quals, HE below degree, RSA Higher, HNC/HND, Nursing
3	2+ A-levels 3+ SCE Highers AS-level or equivalent (4+)	GNVQ - advanced, RSA advanced dip/cert, OND, ONC, BTEC, City and Guilds Advanced Craft , Scot. Cert 6th year Studies, Higher national qualification or equivalent (Scotland) , ,(other) NVQ Level 3
2	1 A-level, 1-3 AS levels, 1-2 Highers 5+ GCSEs at A-C	GNVQ – intermediate, RSA diploma, City and Guilds - Craft , BTEC, SCOTVEC etc. first or general diploma , Scottish CSYS , Intermediate 2 Higher qualification (Scotland), (other) NVQ Level 2
1	1-4 GCSEs/SCE passes (but below Level 2)	GNVQ, GSVQ foundation level, CSE below grade1, GCSE below grade C, BTEC, SCOTVEC first or general certificate, SCOTVEC modules, RSA other, City & Guilds other , YT/YTP certificate , Key Skills, Basic Skills
Apprenticeships	n.a.	Modern and Traditional
Other	n.a.	Unspecified in the data

Table 2: Summary statistics

Males

	Mean	Std. Dev.	Min.	Max.	N
log hourly wage (Sept14 prices)	2.571	0.512	0.935	4.187	199509
RoSLA	0.724	0.447	0	1	204,735
age	41.364	9.577	18	65	204,735
year of birth	1963	11.616	1940	1990	204,735
month of birth dobm	6.403	3.411	1	12	204,735
non white	0.973	0.027	0	1	204,735

Females

	Mean	Std. Dev.	Min.	Max.	N
log hourly wage (Sept14 prices)	2.349	0.489	0.926	4.184	214,462
RoSLA	0.728	0.445	0	1	218,302
age	39.671	11.696	18	65	218,302
date of birth year	1965	11.996	1940	1990	218,302
date of birth month	6.394	3.415	1	12	218,302
non white	0.971	0.030	0	1	218,302

log hourly wage and NVQ percentage

Males

	Mean	Std. Dev.	Perc.	N
NVQ0	2.221	0.371	8.25	15,180
NVQ1	2.257	0.380	5.24	9,651
NVQ2	2.400	0.446	21.19	38,995
NVQ3	2.524	0.441	30.85	56,779
NVQ4	2.900	0.501	34.48	63,458

Females

	Mean	Std. Dev.	Perc.	N
NVQ0	1.995	0.329	10.21	20,620
NVQ1	2.076	0.347	6.04	12,198
NVQ2	2.204	0.394	29.63	59,852
NVQ3	2.275	0.414	19.25	38,881
NVQ4	2.701	0.468	34.87	70,434

Table 3: Ordered probit first step selection model

Dep var: NVQ levels		
	Males	Females
RoSLA	-0.12032***	-0.09978***
	(0.00921)	(0.00865)
ELR Sep-Jan	0.06361***	0.12227***
	(0.01064)	(0.01031)
ELR Feb-Aug	0.07407***	0.14125***
	(0.00995)	(0.00968)
month of birth	-0.00263***	-0.00348***
	(0.00059)	(0.00057)
year of birth	0.02127***	0.03269***
	(0.00198)	(0.00197)
(year of birth)^2	-0.00020***	0.00006
	(0.00008)	(0.00007)
(year of birth)^3	-0.00000**	-0.00001***
	(0.00000)	(0.00000)
nonwhite	0.19518***	0.26581***
	(0.01327)	(0.01223)
cut1	-1.14075***	-0.59129***
	(0.01136)	(0.01139)
cut2	-0.84974***	-0.26528***
	(0.01129)	(0.01137)
cut3	-0.15266***	0.65697***
	(0.01119)	(0.01140)
cut4	0.65647***	1.14289***
	(0.01119)	(0.01142)
N	328109	360534

Significance levels : $*$ 10% $**$ 5% $***$ 1%

std. err. in brackets.

Ho: yob=yob^2=yob^3=0, rej at 1%

LR $\chi^2(3)$=535.66 for males

LR $\chi^2(3)$=3365.56 females.

Table 4: First Step - ordered probit - marginal effects

Dep var: NVQ levels						
	Males					
	RoSLA		ELR			month
	Before	After	No	Sep-Jan	Feb-Aug	of birth
NVQ0	0.06899	0.08629	0.08892	0.07914	0.07761	0.00039
	(0.00087)	(0.00073)	(0.00141)	(0.00079)	(0.00056)	(0.00009)
NVQ1	0.04713	0.05486	0.05633	0.05211	0.05143	0.00017
	(0.00052)	(0.00045)	(0.0007)	(0.00048)	(0.00041)	(0.00004)
NVQ2	0.19148	0.20964	0.21364	0.2042	0.20262	0.0004
	(0.00122)	(0.00081)	(0.00143)	(0.00095)	(0.00075)	(0.00009)
NVQ3	0.31064	0.31211	0.31331	0.31351	0.31343	0.00001
	(0.00086)	(0.00082)	(0.00081)	(0.00081)	(0.00081)	(0.00000)
NVQ4	0.38176	0.3371	0.3278	0.35104	0.35491	-0.00097
	(0.0026)	(0.00127)	(0.00315)	(0.00177)	(0.00109)	(0.00022)

All estimates are significant at 1%

(for simplicity we omit the stars to indicate significance).

Std. err. in brackets.

Table 5: First Step - ordered probit - marginal effects - Females

Dep var: NVQ levels

	Females					
	RoSLA		ELR			month
	Before	After	No	Sep-Jan	Feb-Aug	of birth
NVQ0	0.09229	0.10896	0.12208	0.09975	0.09657	0.00059
	(0.00087)	(0.00086)	(0.00168)	(0.00085)	(0.00059)	(0.0001)
NVQ1	0.06226	0.06924	0.07582	0.0669	0.06553	0.00025
	(0.00057)	(0.00047)	(0.00078)	(0.00051)	(0.00043)	(0.00004)
NVQ2	0.29055	0.30479	0.31961	0.30363	0.30087	0.00049
	(0.00135)	(0.00077)	(0.00130)	(0.00094)	(0.00079)	(0.00008)
NVQ3	0.18311	0.18079	0.17917	0.18378	0.18428	-0.0001
	(0.00064)	(0.00067)	(0.00073)	(0.00065)	(0.00065)	(0.00002)
NVQ4	0.3718	0.33623	0.30333	0.34593	0.35275	-0.00124
	(0.00247)	(0.00106)	(0.00294)	(0.0016)	(0.00099)	(0.0002)

All estimates are significant at 1%

(for simplicity we omit the stars to indicate significance).

Std. err. in brackets.

Table 6: Basic OLS model with no cohort effects

	Males					
	NVQ0	NVQ1	NVQ2	NVQ3	NVQ4	NVQ tot
Dep var: log hourly earnings - QLFS						
age	0.04524***	0.06704***	0.07988***	0.08188***	0.11051***	0.08381***
	(0.00132)	(0.00162)	(0.00088)	(0.00075)	(0.00096)	(0.00045)
age^2	-0.00047***	-0.00074***	-0.00086***	-0.00091***	-0.00119***	-0.00091***
	(0.00002)	(0.00002)	(0.00001)	(0.00001)	(0.00001)	(0.00001)
NVQ1						0.09978***
						(0.00417)
NVQ2						0.24485***
						(0.00310)
NVQ3						0.33226***
						(0.00292)
NVQ4						0.69677***
						(0.00290)
constant	1.22677***	0.90171***	0.76020***	0.85428***	0.53550***	0.43504***
	(0.02695)	(0.02962)	(0.01644)	(0.01469)	(0.01906)	(0.00901)
N	25440	16655	65354	100724	111991	320164

	Females					
	NVQ0	NVQ1	NVQ2	NVQ3	NVQ4	NVQ tot
age	0.00838***	0.01639***	0.03478***	0.05073***	0.07249***	0.08381***
	(0.00124)	(0.00143)	(0.00071)	(0.00093)	(0.00092)	(0.00045)
age^2	-0.00006***	-0.00012***	-0.00037***	-0.00058***	-0.00079***	-0.00091***
	(0.00001)	(0.00002)	(0.00001)	(0.00001)	(0.00001)	(0.00001)
NVQ1						0.09978***
						(0.00417)
NVQ2						0.24485***
						(0.00310)
NVQ3						0.33226***
						(0.00292)
NVQ4						0.69677***
						(0.00290)
constant	1.73349***	1.62589***	1.45690***	1.28111***	1.18130***	0.43504***
	(0.02687)	(0.02864)	(0.01374)	(0.01709)	(0.01794)	(0.00901)
N	36387	23057	106656	65956	122835	320164

Significance levels : $*$ 10% $**$ 5% $***$ 1% Std. err. in brackets. Dependent Variable in 2014 prices.
Test on separability for males rej. Ho at 5%, F_{stat}=453.3; for females rej. Ho at 5%, F_{stat}=389.5
NVQ tot is a categorical variable for each NVQ level.

Table 7: Estimates Heckman model - no cohort effects

Dep var: log earnings

Males

	NVQ0	NVQ1	NVQ2	NVQ3	NVQ4
λ	-0.16874***	-0.09386	0.01438	-0.36992***	-0.56753***
	(0.04931)	-0.06502	-0.03511	-0.03554	-0.05635
non white	-0.10317***	-0.03995*	-0.02405*	-0.10762***	-0.10636***
	(0.02101)	-0.02418	-0.01394	-0.01577	-0.01425
age	0.04012***	0.06522***	0.07809***	0.07462***	0.10544***
	(0.00164)	-0.00225	-0.00114	-0.00126	-0.00144
age^2	-0.00041***	-0.00071***	-0.00084***	-0.00081***	-0.00110***
	-0.00002	-0.00003	-0.00002	-0.00002	-0.00002
constant	0.99839***	0.80377***	0.79363***	0.95523***	1.19035***
	-0.09016	-0.07734	-0.02499	-0.02235	-0.07347
N	199509				

Females

	NVQ0	NVQ1	NVQ2	NVQ3	NVQ4
λ	-0.34536***	-0.12989***	-0.08962***	-0.15321***	-0.11483***
	(0.02007)	(0.02413)	(0.01471)	(0.01882)	(0.02341)
non white	-0.04025	0.04425*	0.05787***	0.02057*	0.02060**
	(0.02773)	(0.02511)	(0.01035)	(0.01235)	(0.00966)
age	0.01293***	0.01551***	0.03473***	0.04885***	0.07400***
	(0.00146)	(0.00178)	(0.00087)	(0.00115)	(0.00119)
age^2	-0.00003**	-0.00007***	-0.00034***	-0.00051***	-0.00078***
	(0.00002)	(0.00002)	(0.00001)	(0.00002)	(0.00002)
constant	0.91329***	1.42723***	1.36408***	1.25830***	1.22388***
	(0.05933)	(0.05092)	(0.02012)	(0.01982)	(0.03233)
N	214462				

Significance levels : $*$ 10% $**$ 5% $***$ 1% Bootstrapped Std. err. (200 reps) in brackets.
Dependent Variable in 2014 prices.
Ho: $\lambda_0 = \cdots = \lambda_5 = 0$ $\chi^2_5 = 176.6$ for males; $\chi^2_5 = 508.80$ for females;

Table 8: First difference model from LLFS

Males						
	NVQ0	NVQ1	NVQ2	NVQ3	NVQ4	NVQ tot
Dep var: first difference log hourly earnings - LLFS						
5/2×age	-0.00084***	-0.00060***	-0.00096***	-0.00094***	-0.00106***	-0.00096***
	(0.00012)	(0.00015)	(0.00007)	(0.00006)	(0.00006)	(0.00003)
NVQ1						-0.00522
						(0.00401)
NVQ2						-0.00418
						(0.00302)
NVQ3						-0.00050
						(0.00282)
NVQ4						0.00532*
						(0.00280)
constant	0.12245***	0.10100***	0.12883***	0.13060***	0.14653***	0.13300***
	(0.01153)	(0.01165)	(0.00582)	(0.00489)	(0.00499)	(0.00399)
N	10097	6898	25888	43305	47789	133977

Females						
	NVQ0	NVQ1	NVQ2	NVQ3	NVQ4	NVQ tot
5/2×age	-0.00024*	-0.00034***	-0.00059***	-0.00070***	-0.00076***	-0.00062***
	(0.00012)	(0.00013)	(0.00006)	(0.00007)	(0.00006)	(0.00003)
NVQ1						0.01014***
						(0.00329)
NVQ2						0.00592**
						(0.00247)
NVQ3						0.00529**
						(0.00270)
NVQ4						0.01339***
						(0.00243)
constant	0.06163***	0.08384***	0.10119***	0.10925***	0.12257***	0.09818***
	(0.01195)	(0.01099)	(0.00498)	(0.00598)	(0.00504)	(0.00385)
N	15565	10674	46098	26575	51585	150497

Significance levels : $*$ 10% $**$ 5% $***$ 1% Std. err. in brackets. Dependent Variable in 2014 prices.
LLFS: Test on separability for males rej. Ho at 5%, F_{stat}=9.083; for females rej. Ho at 5%, F_{stat}=21.361
NVQ tot is a categorical variable for each NVQ level.

Table 9: Estimates Heckman model - with cohort effects

Dep var: log earnings corrected for lifecycle effects

Males

	NVQ0	NVQ1	NVQ2	NVQ3	NVQ4
λ	0.01207	0.08630	-0.00995	0.21145***	0.36755***
	(0.10411)	(0.11318)	(0.06704)	(0.05906)	(0.09100)
non white	-0.02144	0.02209	0.02114	0.02098	0.05106***
	(0.02944)	(0.03462)	(0.02223)	(0.01898)	(0.01716)
constant	2.73713***	2.45484***	2.83214***	2.89805***	2.89863***
	(0.16343)	(0.12431)	(0.03675)	(0.02289)	(0.12046)
N	163082				

Females

	NVQ0	NVQ1	NVQ2	NVQ3	NVQ4
λ	-0.03122	-0.32632***	-0.26709***	-0.26823***	-0.29991***
	(0.05125)	(0.07665)	(0.04073)	(0.05488)	(0.06416)
non white	-0.00345	-0.03479	0.01358	0.00099	-0.02929*
	(0.02612)	(0.03414)	(0.01705)	(0.02194)	(0.01643)
constant	1.93785***	2.13229***	2.59946***	2.93491***	3.58021***
	(0.05275)	(0.02467)	(0.01684)	(0.05855)	(0.11416)
N	178827				

Significance levels : $*$ 10% $**$ 5% $***$ 1% Bootstrapped Std. err. (200 reps) in brackets.
Dependent Variable in 2014 prices. All equations include cubic year of births that capture additive cohort effects. Age effects are imposed from Table 8, according to equation 4.
Ho: $\lambda_0 = \cdots = \lambda_5 = 0$ $\chi^2_5 = 34.36$ for males; $\chi^2_5 = 88.43$ for females;

Table 10: Predicted wages

Selection Model ^{a} *with age effects and no cohort effects*

	Males		Females	
	Mean	Std. dev	Mean	Std. dev
NVQ0	2.231	0.086	1.992	0.052
NVQ1	2.274	0.145	2.081	0.076
NVQ2	2.422	0.185	2.212	0.081
NVQ3	2.542	0.154	2.284	0.104
NVQ4	2.927	0.195	2.718	0.130

Selection Model ^{b} *with cohort effects and lifecycle correction*

	Males		Females	
	Mean	Std. dev	Mean	Std. dev
NVQ0	2.190	0.431	1.968	0.240
NVQ1	2.234	0.405	2.058	0.336
NVQ2	2.358	0.494	2.164	0.402
NVQ3	2.483	0.480	2.228	0.427
NVQ4	2.860	0.543	2.658	0.471

[a] Predicted wages from Table 7.

[b] Predicted wages from Table 9.

Note: If we add the age effects to the predictions in model b we obtain average wage predictions similar to model a.

Figure 1: Distribution of Highest Qualifications across Birth Cohorts

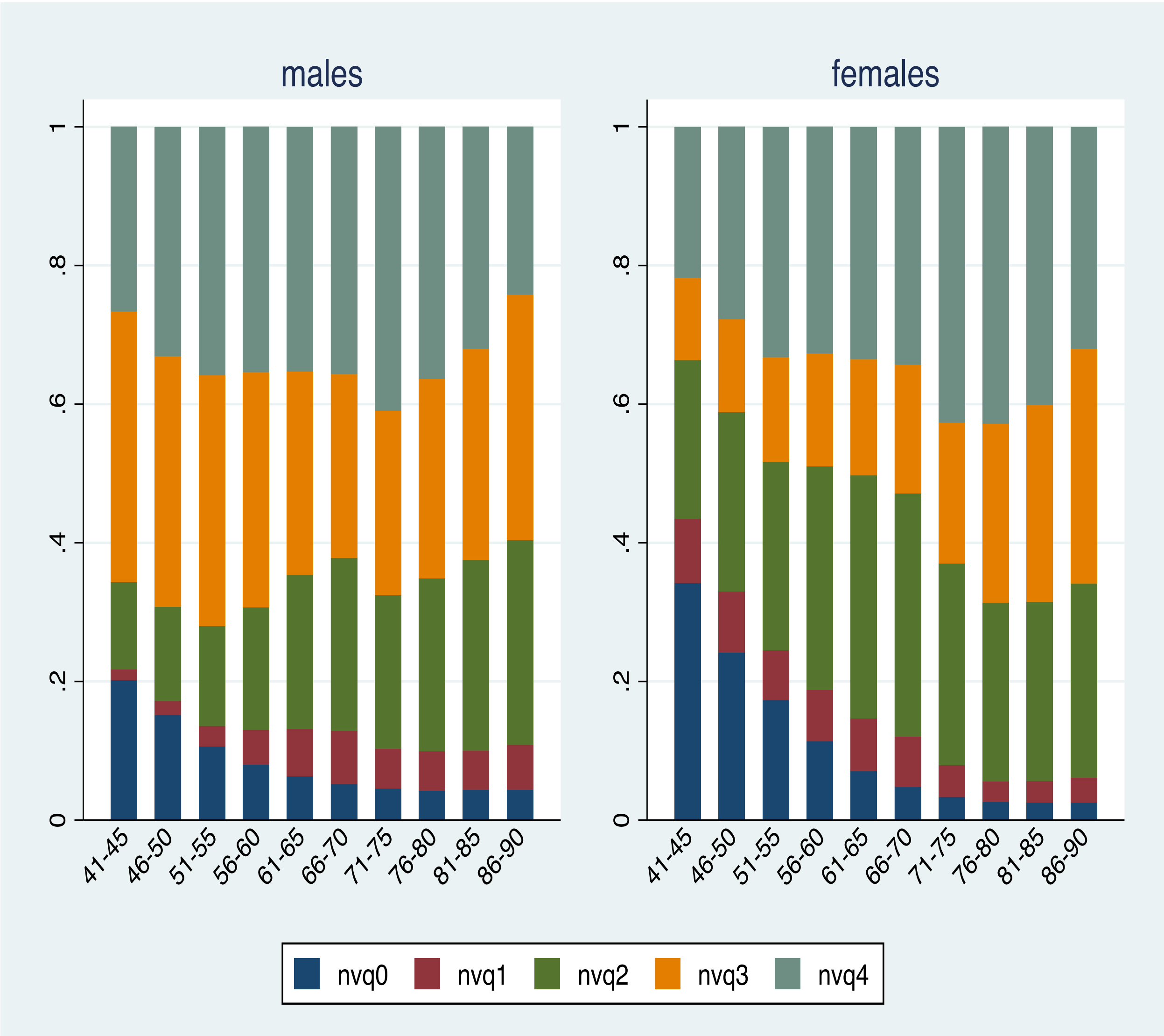


Figure 2: Effects of Month of Birth, RoSLA and ELR: Males

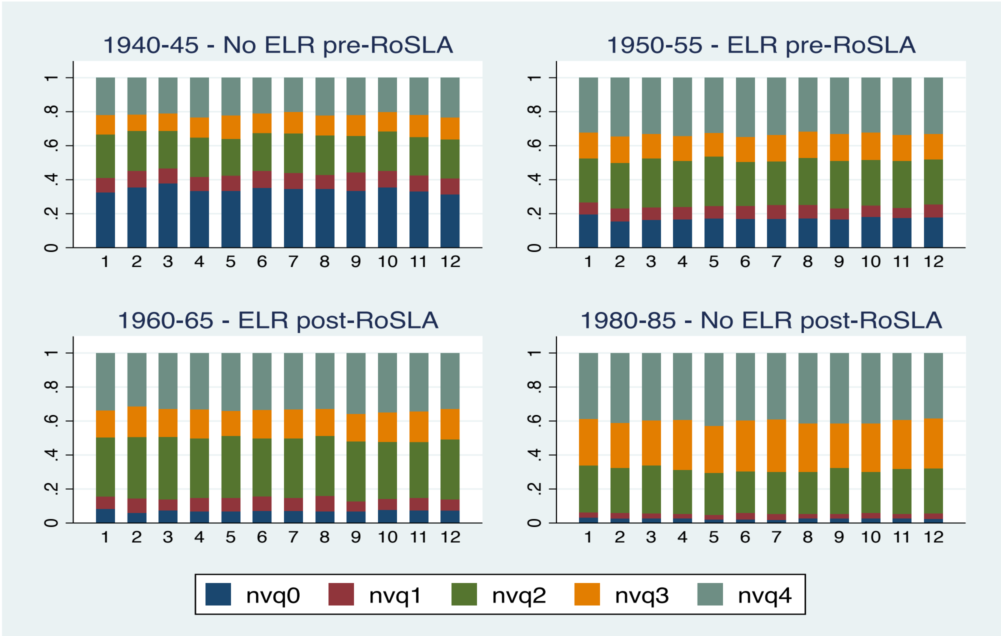


Figure 3: Effects of Month of Birth, RoSLA and ELR: Females

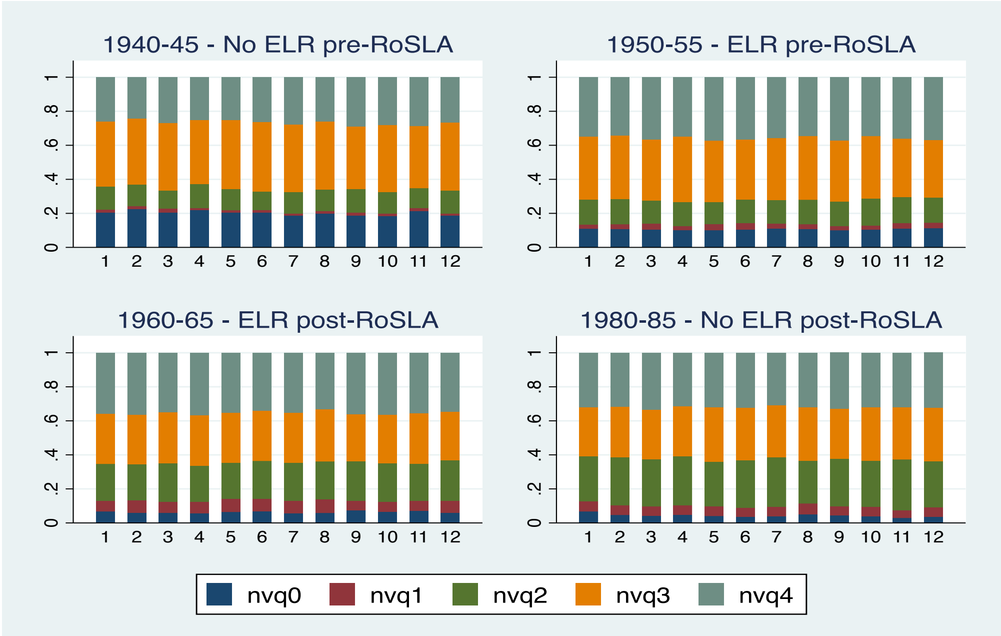
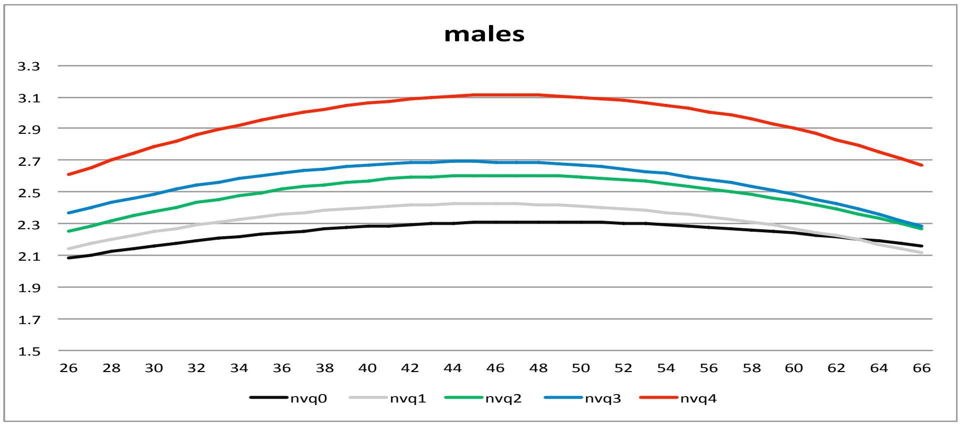


Figure 4: Raw LFS pooled data - OLS quadratics with discrete S groups and no cohort effects



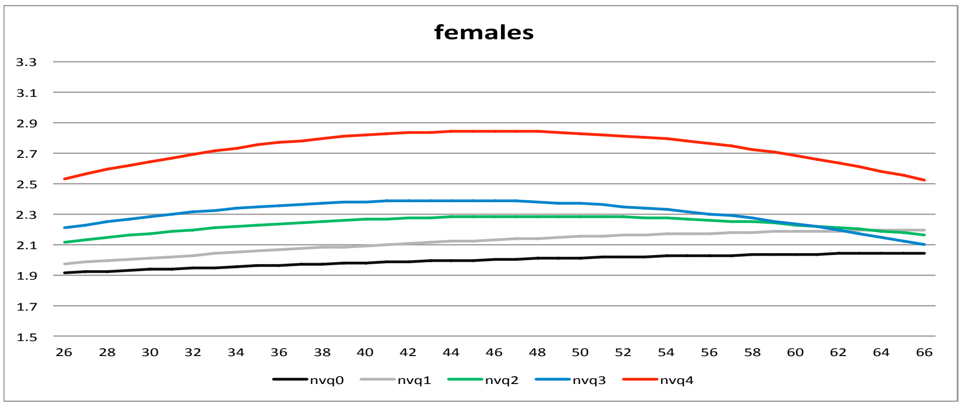


Figure 5: Lifecycle effects - Quadratic in age, assuming no cohort differences

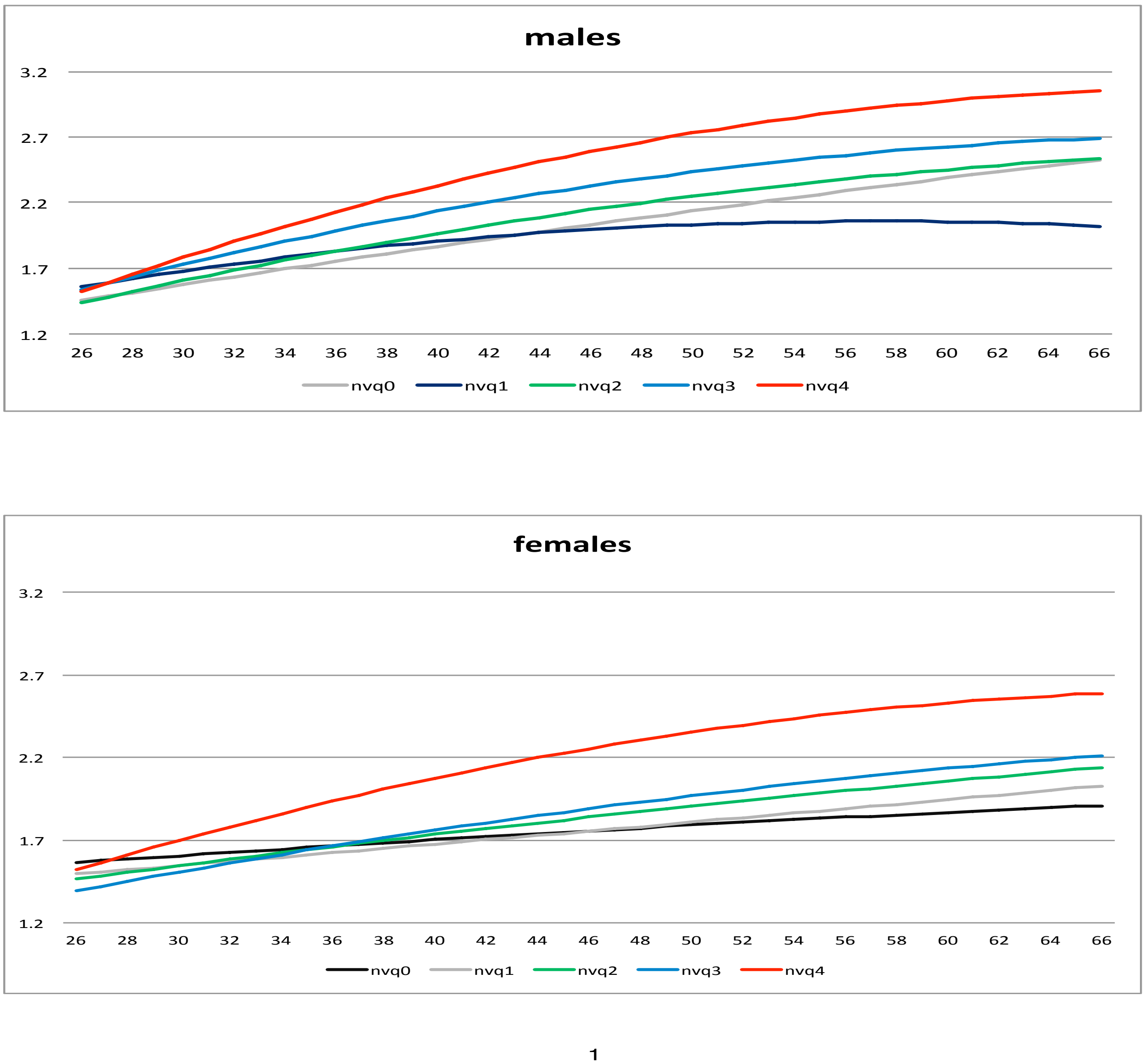


Figure 6: Lifecycle effects - Allowing for cohort differences

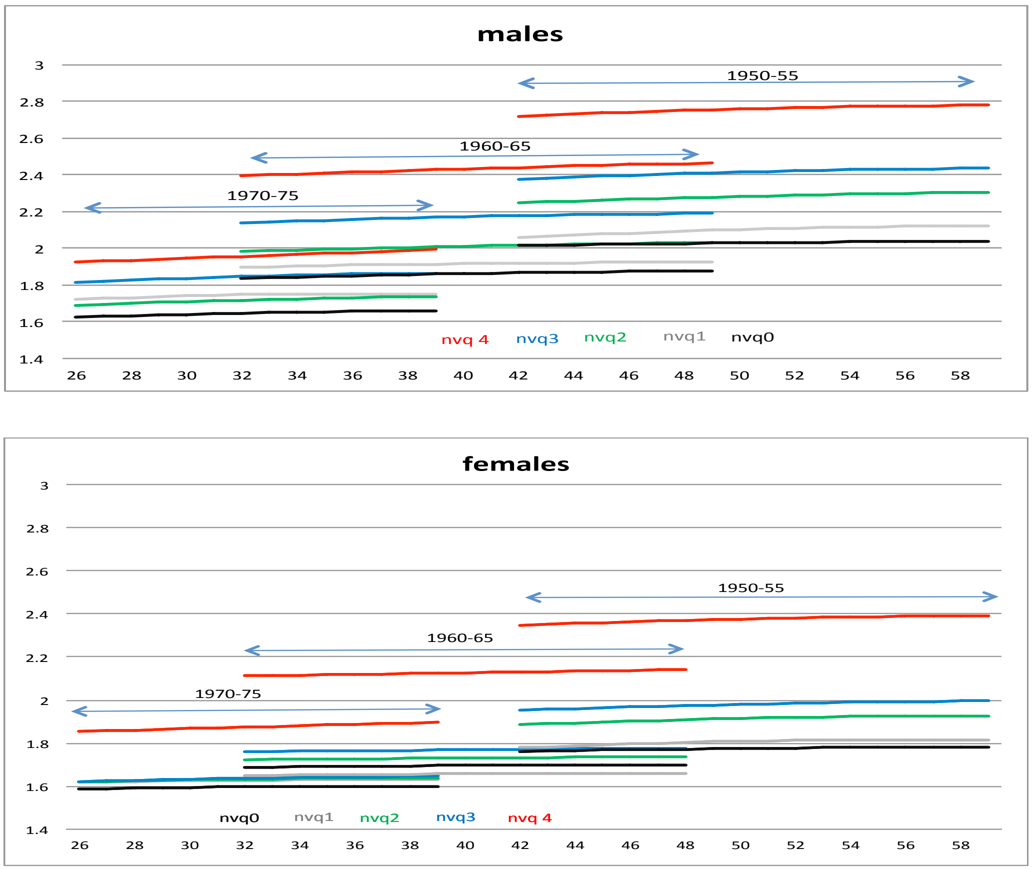


Figure 7: Returns to education by cohort - High school and Higher education

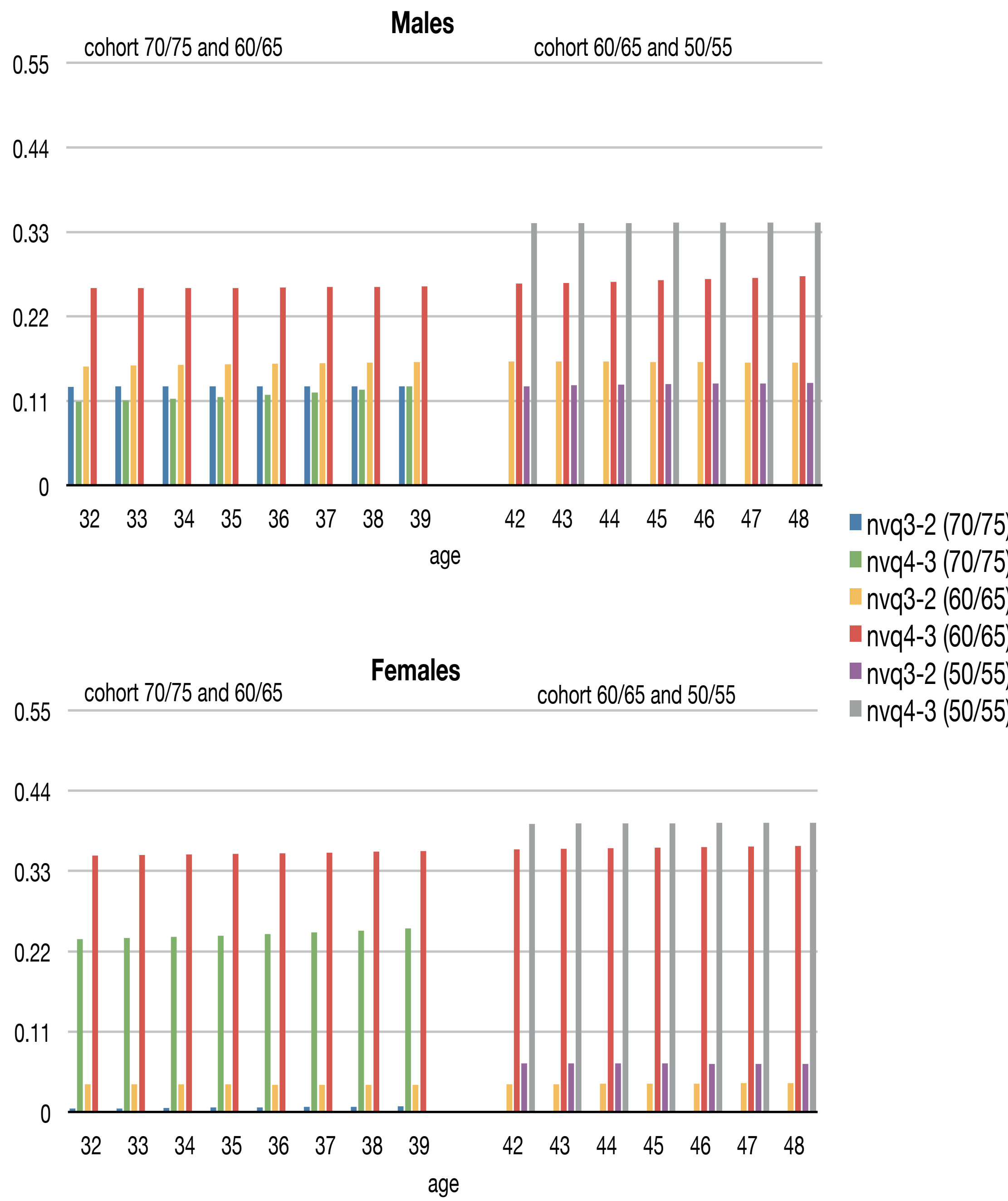


Table A1: Summary statistics unrestricted QLFS

Males	Mean	Std. Dev.	N
log hourly wage $_{unadjusted}$	2.295	0.616	466,384
age $18 - 65$	40.467	11.940	473,130

Females	Mean	Std. Dev.	N
log hourly wage $_{unadjusted}$	2.064	0.576	498,917
age $18 - 65$	40.281	11.667	504,911

log hourly wage and NVQ qualifications

Males	Mean	Std. Dev.	N
NVQ0	1.895	0.497	37,059
NVQ1	1.970	0.495	19,974
NVQ2	2.128	0.561	81,112
NVQ3	2.218	0.546	132,766
NVQ4	2.638	0.585	127,675

Females	Mean	Std. Dev.	Perc.	N
NVQ 0	1.656	0.429	49,252	
NVQ1	1.767	0.447	27,985	
NVQ2	1.912	0.479	128,549	
NVQ3	1.976	0.511	89,002	
NVQ4	2.427	0.544	137,375	

Figure A1: McCrary tests of month of birth, Males

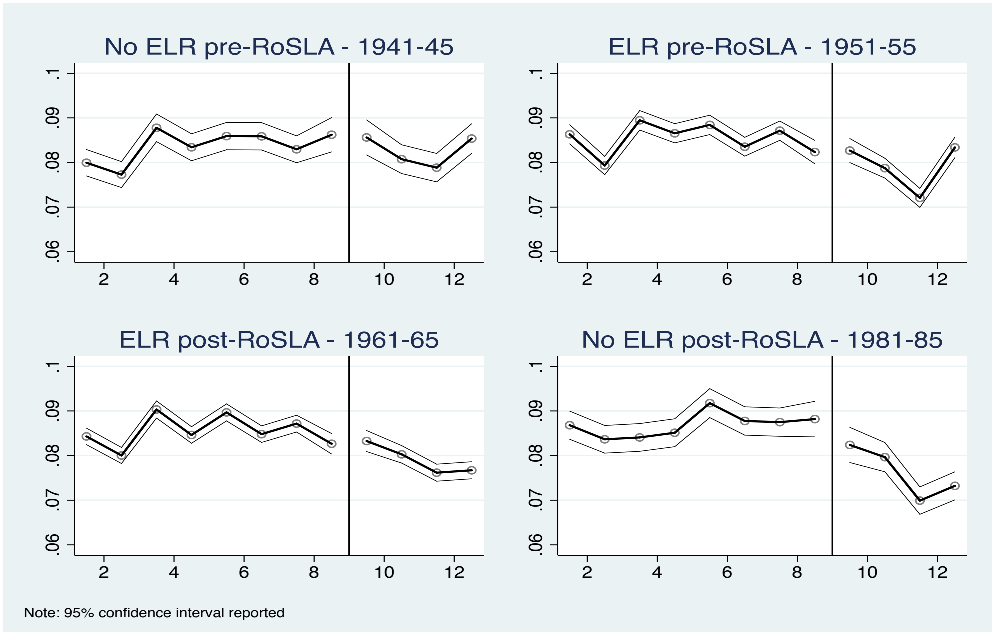


Figure A2: McCrary tests of month of birth, Females

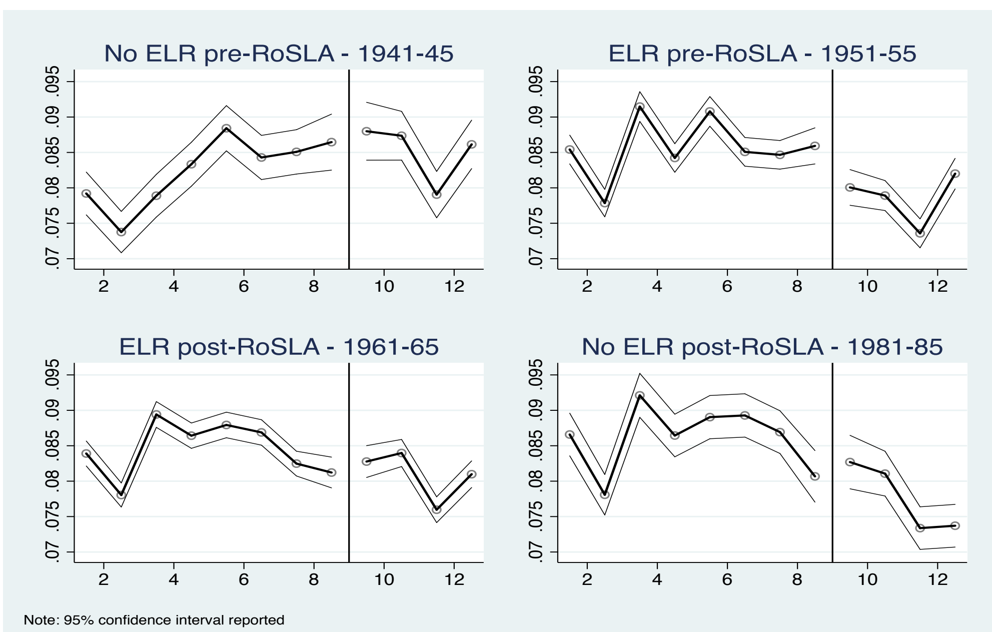


Figure A3: Log real wage differences by NVQ and sex