The Changing Distribution of Male Wages in the UK

Amanda Gosling
Steve Machin
Costas Meghir

THE INSTITUTE FOR FISCAL STUDIES
Working Paper Series No. W98/9

The Changing Distribution of Male Wages in the UK

Amanda Gosling[*], Stephen Machin[†] and Costas Meghir[‡]

September 1999

Abstract

Forthcoming: **Review of Economic Studies.**

This paper uses microeconomic data from the UK Family Expenditure Surveys (FES) and the General Household Surveys (GHS) to describe and explain changes in the distribution of male wages. Since the late 1970s wage inequality has risen very fast in the UK, and this rise is characterised both by increasing education and age differentials. We show that a large part of the changes in the UK can be summarised quite simply as increases in education differentials and a decline of growth of entry level wages which persist subsequently. This fact we interpret as cohort effects. We also show that, like in the US, an important aspect of rising wage inequality is increased within-group wage dispersion. Finally we use the GHS to evaluate the role of alternative education measures.

Key words: Wage dispersion; Returns to education; Cohort effects.
JEL classification numbers: D31, J31.
Address for Correspondence:
Institute for Fiscal Studies, 7 Ridgmount Street, LondonWC1E 7AE, UK

[*]Institute for Fiscal Studies, Department of Economics University of Essex and CEPR
[†]Institute for Fiscal Studies, Department of Economics, University College, London and Centre for Econonomic Performance, LSE
[‡]Institute for Fiscal Studies and Department of Economics, University College, London

1. Introduction[1]

Wage dispersion in the UK has risen sharply since the late 1970s. The scale of this increase is shown in Figure 1.1 where we plot indexed real hourly wages of male workers for the 10th, median and 90th percentiles of the hourly wage distribution from 1966 to 1995. Following relatively small changes from the mid-1960s to late 1970s, this shows a rapidly widening gap between high paid and low paid workers through the 1980s and 1990s. The relative size of the recent UK increase in wage inequality is large both in terms of the UK's own historical experience and in terms of comparison with other countries.[2] Of course, there is no *a priori* reason why the structure of wages should remain stable over time. Technological innovations, changes to the distribution of education and in the structure of labour and product markets are likely to alter the demand for and supply of different skill attributes. This process can sometimes lead to permanent changes in relative prices and quantities. In this vein, a number of papers have attempted to uncover the relative importance of technological changes, institutions and international trade in shaping the changes of the US and UK distribution of wages[3].

[1] The comments of three referees helped us improve the paper substantially. We are also indebted to Manuel Arellano, Orazio Attanasio, Richard Blundell, Andrew Dilnot, David Green, Zvi Griliches, John Hills, Joel Horowitz, Paul Johnson, Tom MaCurdy, Charles Manski, Andrew Oswald, John Pencavel, Jim Powell, Mark Stewart, John Van Reenen and Edward Whitehouse for their comments and discussions. We thank participants in seminars at Warwick, the Poverty Institute at the University of Wisconsin and an ESRC half day conference on wage dispersion at the IFS. Finally we would also like to thank Alissa Goodman and Steve Webb for invaluable help with the FES data and Steve McIntosh for helping us collect the pupil teacher ratio data. Financial assistance from the ESRC centre for the Microeconometric Analysis of Fiscal Policy at the IFS and the Joseph Rowntree Foundation is gratefully acknowledged. We retain responsibility for all errors and interpretations.

[2] See for example Gosling, Machin and Meghir (1994), Katz Loveman and Blanchflower (1995) and Schmitt (1995) for earlier descriptions of the changes in the UK and the survey of Machin (1996) which places the UK rise into its historical and international context.

[3] See for example, Murphy and Welch (1992) Bound and Johnson (1992), Freeman and Katz (1994), Dinardo, Fortin and Lemieux (1996) and Gosling and Machin (1995)

In this paper, we provide a simple characterisation of the way that the distribution of male wages has evolved. We describe the distribution of wages using a set of quantiles[4]. In the same way as MaCurdy and Mroz (1991), who model the structure of wages in the US, we show that each of the chosen quantiles of the distribution of wages can be modelled as an additive function of cohort, life-cycle and time effects (constructed to average to zero within the sample period). The median closely relates to a standard human capital wage equation. The other quantiles allow us to infer how the dispersion of wages. Of course, if the distribution of unobservable determinants of wages is independent of the observed ones, then the entire set of changes in the distribution of wages can be explained by changes in the median. This hypothesis is strongly rejected in our data.

Cohort effects play an important part in our interpretation of the changes in the distribution of wages.[5] We show that conditional on the presence of cohort effects, life-cycle profiles of wages have not changed. Sources of these cohort effects can be changes in education policy, changes in the role of labour market institutions, as well as changes in the conditions at labour market entry. All such factors may affect the amount and type of human capital accumulated by young workers by the age of 23, which is the age at which we start modelling wages.[6] If such pre- and early labour market factors serve to shift the life-cycle of wages permanently, then they can be summarised by an additive cohort effect, conditional on age, education and cyclical time effects.

[4] This has some similarities to the work of Juhn, Murphy and Pierce (1993) and Bushinsky (1994) who also consider the evolution of quantiles of the wage distribution.

[5] See also Beaudry and Green (1996) for Canada, Fitzenberger, Hujer, MaCurdy and Schnabel (1995) for Germany and MaCurdy and Mroz (1991) for the US.

[6] Examples of the education reforms introduced in the UK are the formalisation of the "tripartite" education system after 1945 which streamed children at age 11 into different types of education; the expansion of higher education after the 1963 Robbins report; the partial and gradual dismantling of the tripartite system in the late 1960s; the introduction of youth training schemes in the early 1980s; and the "Assisted Places" scheme in the mid 1980s which partially financed private education for children from poorer backgrounds.

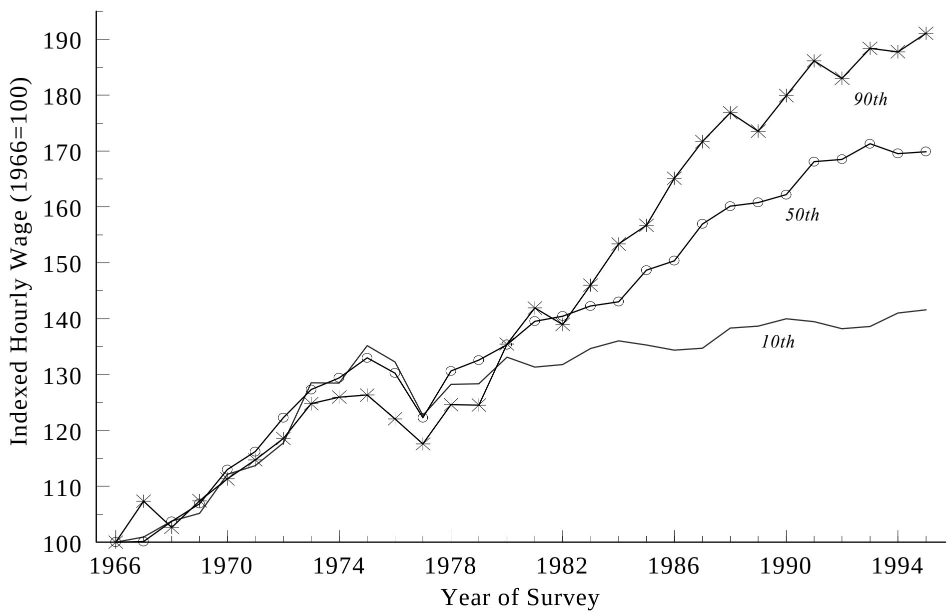

Figure 1.1: Indexed Real Hourly Wages by Percentile (Source: FES)

We show that changes in education differentials, together with cohort effects, can explain two thirds of the overall increase in wage dispersion over our time period. Nevertheless, there have also been important within-group shifts in wage structure. As successive generations of workers have entered the labour market they have done so with more and more dispersed wages which have persisted over time to create a more unequal wage distribution.

The structure of the rest of the paper is as follows. Section 2 describes the way we characterise the distribution of wages. In section 3 we give a basic description of the data and section 4 presents our econometric estimates. A concluding discussion is presented in section 5. All technical issues are discussed in the appendix.

2. Estimating The Wage Distribution Over Time

2.1. Modelling approach

We model the quantiles of the distribution of wages as functions of age, education, cyclical time effects and cohort effects. The latter reflects differences in productive characteristics across generations as well as general productivity growth. If all quantiles evolve in the same way (aside from an intercept shift) then the changing dispersion of pay can be explained by the changing returns to and the composition of observed skill characteristics. Using quantiles is an easy and intuitive way of characterising the distribution of wages. The median defines the location of the distribution while the quantiles around it can be used to describe changes in dispersion or other aspects of its shape. Differences across quantiles can thus be interpreted as an estimate of the changing importance of the unobserved component in wages[7].

We can define the qth quantile of the conditional wage distribution as:

$$q = \Pr[w_{it} < w^q | cohort_i, age_{it}, education_i] \tag{2.1}$$

where w_{it} refers to the log of the real hourly wage rate of individual i in time period t. In equation (2.1) *cohort* is the year of birth of a particular worker, *age* is his age in years at the time his wage is observed and *education* (abbreviated to *ed* below) is a measure of education (precise definitions are given below).

Following MaCurdy and Mroz (1991), we restrict equation (2.1) to have the following additive structure:

[7] It turns out that in this data, the conditional means and medians are indistinguishable.

$$w^q(cohort_i, age_{it}, ed_i) = A^q(age_{it}, ed_i) + C^q(cohort_i, ed_i) + T^q(time, ed_i) \quad (2.2)$$

where T^q $(time, ed_i)$ is constructed so that it is orthogonal to the cohort and age functions and hence includes no trends. All trends in the data will be included in the functions A^q and C^q. We examine this assumption in more detail below. By comparing the fit of equation (2.2) to that of the unrestricted quantile estimates, we can test whether the data can reject this restriction.

To justify the loglinear specification we suppose that for an individual i in education group ed human capital (H) is produced by

$$H_{it}^{ed} = \exp(A(age_{it}, ed_i) + C(cohort_i, ed_i) + u_{it}) \quad (2.3)$$

where u_{it} represents unobserved characteristics. We assume that the production function depends on the total amount of human capital employed for each education group (H_t^{ed}). Denote the equilibrium price of human capital for education group ed by $exp(T_t^{ed})$. Then a person with human capital H_{it}^{ed} will earn a wage rate of $w_{it} = H_{it}^{ed} exp(T_t^{ed})$.[8] The corresponding log wage equation will then be

$$\log w_{it} = T_t^{ed} + A(age_{it}, ed_i) + C(cohort_i, ed_i) + u_{it} \quad (2.4)$$

If u_{it} is independent of age, education and cohort then all quantiles are identical apart from an intercept shift and more importantly that observables are able to account for all changes in the distribution of wages. The specification in equation (2.2) relaxes this independence assumption but assumes that the additive structure is preserved for the quantiles. This assumption is testable. Thus, as well as

[8] The fact that any common trend in productivity growth is included in human capital rather than in its price is purely a normalisation. This is discussed further below.

being a simple way of characterising the wage distribution, using quantiles provides a direct way of assessing the contribution of unobservables in explaining the evolution of the entire wage distribution, not only at the conditional mean and variance.

This simplified specification has the attraction that we can attach clear interpretations to each component. The set of functions $C^q(cohort_i, ed_i)$ for each quantile will measure the cross generational differences in wages caused by the different characteristics of cohorts. Thus, the difference in these functions between the top and the bottom of the distribution captures changes in within group dispersion driven by cohort effects and by general trends in productivity. Similarly the differences across education groups will measure changes in the returns to education across generations.

The functions $A^q(age_{it}, ed_i)$ measure how the wage distribution changes as a cohort ages for each education group. We expect this to be important for the following reasons. First and foremost, age minus years of education is the level of potential experience for an individual, so if experience is important we should observe wages growing with age. Second, with differential rates of learning by doing we should also observe an increase in the variance of wages with age.[9] (Deaton and Paxson, 1993, make a similar argument in the case of consumption). Third, A^q will reflect changes in wages over the life-cycle for reasons other than accumulated experience, most importantly general productivity growth.

We define common shocks to the wage distribution $T^q(time, ed_i)$ as those changes in wages which are the same within all education groups regardless of age and which are orthogonal to the other functions in equation (2.2). This func-

[9] As Farber and Gibbons (1997) point out a learning model would also lead to increases in the variance with age: Age reflects learning about individual ability both by workers and their employers. As unobserved productivity is gradually revealed and translated into pay the variance of wages will increase with age.

tional form specification defines macroeconomic effects as relative changes within the sample which are common to all individuals in the same education group at any point in time regardless of age. This is the only way that lifecycle growth rates will differ across generations.

Since, by construction, the time effects $T^q(time, ed_i)$ do not contain any trends, macroeconomic wage growth will be reflected in the evolution of the other two (age and cohort) functions. Thus the rate of growth of entry level wages for the qth quantile is given by $\frac{\partial C^q}{\partial cohort} |_{age}$. Given entry, wages at the qth quantile grow at a rate $\frac{\partial A^q}{\partial age} |_{cohort}$. The latter is thus restricted to be the same across cohorts. Both these growth rates contain a constant term which is independent of both age and cohort and can be thought of as the effect of a constant change in productivity, which may differ across skill groups. If it differs across quantiles, conditional on education, this indicates the existence of a trend within our sample period towards increasing within group dispersion, i.e. productivity growth differs for individuals with different unobserved productivity characteristics.

Finally, an important note of caution is in order. Whilst it is possible to test that the restrictions we impose are not rejected by the data, it does not mean that this is the only parsimonious specification for which this is true. The interpretation we attach is fundamentally an identifying assumption. This is because cohort, age and time do not vary independently. We are driven to such an interpretation by our belief that the accumulation of human capital, aside from observed measures of education will depend on pre-labour market experiences such as the interaction of school quality, neighbourhood effects and family background. Dearden, Ferri and Meghir (1998) and Dearden, Machin and Reed (1997) both show that such pre-labour market factors are important determinants of future wages. Policy reforms to the schooling system, the changing distribution of work across households and trends in family composition all mean that these processes may differ systematically across different generations. Finally early labour market conditions at time of labour market entry will also have permanent effects if initial sectoral or occupational allocation is important (see Baker, Gibbs and Holmstrom, 1995, Beaudry and DiNardo, 1991 and Neal 1995), if entry factors affect the cost of skill acquisition or, following the learning by doing model of Aghion, Howitt and Violante (1999), if cyclical factors affect the distribution of new technologies across jobs. Once one assumes that cohort effects are important our specification is a natural way to proceed unless, of course, one can model cohort effects using observables which requires detailed information on individual characteristics, together with prior restrictions on what the determinants of cohort effects could be. In the end of the paper, we undertake a simple descriptive analysis and show that the detrended cohort effects relate to pre and early labour market conditions.

2.2. Estimation method

To characterise the distribution we estimate 13 quantiles (the 5th, the 25th, the 75th, the 95th percentiles and the 9 deciles). For the estimation of each one we use a two step estimation procedure which is implemented on each education group separately. The first step regressions include general functions of cohort and age and interactions thereof.

These quantiles are estimated using the Smoothed Least Absolute Deviations estimator (SLAD) recently suggested by Horowitz (1998). This has better small sample properties than the usual LAD estimator and since the criterion function is differentiable, optimisation can take place using standard gradient type methods.[10] The predictions from the first stage regression were then regressed against

[10] In fact there were two first step regressions, the first was a completely saturated model, whose parameters were simply the order statistics of each cohort, year and education cell. The second was a more restrictive one, relating wages to polynomials in age, cohort, a set

our chosen functional form above, using weighted least squares where the weights are based on the covariance matrix of the predicted quantiles from the first stage.[11]

We provide graphical descriptions of the fit of the model. We also construct goodness of fit statistics *vis a vis* a completely saturated model where all cohort age interactions are included. These χ^2 statistics give a measure of how well the restricted distribution tracks observed changes. In essence, the goodness of fit analysis evaluates the contribution of cohort-age interactions in explaining the changes in the dispersion of wages.

The estimated conditional quantiles were then used to construct the entire conditional distribution of wages, allowing us to construct counterfactual distributions of wages and consider the within and between group contributions to changes in dispersion. The details of our estimation procedure are given in the appendix.

3. The Data and the Basic Facts

We use two repeated cross-section datasets in our empirical work, the Family Expenditure Surveys (FES) and the General Household Surveys (GHS). They provide an excellent complementary analysis as FES contains good quality wage data with only a basic measure of education (age at the end of full time education) whilst the GHS only has a consistent series of weekly rather than hourly earnings, but contains information on highest educational qualifications. From both datasets we take all men aged between 23 and 59 (inclusive) who worked at least one hour in the past week.[12] For the FES we construct individual hourly

of interations between functions of cohort and age and a set of year dummy variables. The procedure used to obtain these estimates and their variances is described more fully in the appendix.

[11] This is similar to the minimum distance procedure in the context of quantile regressions suggested by Chamberlain (1993).

[12] We excluded all men who reported themselves to be self employed.

wages by dividing usual weekly pre-tax earnings by usual weekly hours of work (both measures including overtime). Due to changes in questionnaire design over time, we are unable to construct a consistent series of hourly wages in the GHS and so we model the usual weekly earnings of full time men.[13] The education measure available in the FES (since 1978) is age left full time education. The GHS reports the individual's highest formal educational qualification, as well as age left full time education, allowing a comparison with the FES.

Our initial analysis is based on FES data from 1978 to 1995 and for this we allocate workers into three education groups: those who left full time education at or before age 16, those who left at 17 or 18 and those who left after 18 (i.e. including college graduates). We then repeat our analysis on GHS data from 1978 to 1991 allocating workers into four education groups according to their highest educational qualification: no formal qualifications, "O" level, CSE or equivalent qualification (usually obtained at age 16), "A" Level or equivalent (usually obtained at age 18 and needed to get into most universities) and Degree or Teaching qualification.[14] There are 55501 observations in the FES and 53356 observations in the GHS samples.

The advantage of the FES is that it provides a measure of the hourly wage rate and it covers a longer time period. The GHS only provides a consistent series of weekly earning from 1978 to 1991. However the information on the actual qualification obtained can be very valuable in obtaining insights into why the dispersion of wages has increased within the educational categories based on the age that full time education ended (referred to from now as years of education). In Table 3.1 uses GHS data to show how the two education measures are related.[15]

[13] The earnings measure in the GHS includes payment for bonuses and overtime but the hours measure excludes overtime which varies across individuals and over the business cycle.

[14] The GHS series stops in 1991 because of a definition change in the way the earnings data was collected after then.

[15] Note that qualifications may be obtained on a part time basis as well. Hence it is possible,

The rows add up to one. Each cell shows the proportion of each group defined by age left full time education obtaining each level of formal qualification. It is evident from Table 3.1 that the composition of our defined three education groups has changed quite dramatically over time. The proportion of individuals with no qualification has declined in the age 16 group, from 62% to 43%. At the other end, the 19+ group contains proportionately fewer individuals with degrees. Finally in table 3.2 we document the change in the levels of education achieved by successive cohorts (in five year bands). There has been an impressive move in favour of higher levels of education across cohorts. This has to be borne in mind when interpreting the results, since the ability distribution within the education groups is likely to be changing.[16]

Figures 3.1 and 3.2 use FES and GHS data respectively to plot movements in education-related wage differentials over time. They plot median log(wages) for a given education group relative to a low education base category (workers who left school at or before 16 for the FES and those with no educational qualifications from GHS). In both cases, there is strong evidence of an upward trend in the unconditional educational wage differentials, with growth in median wages being positively correlated with the level of education. For example, in the GHS the raw log(wage) differential between workers with a degree and those with no qualifications rose from 0.320 in 1978 to 0.460 in 1991 (as depicted by a 0.140 rise in the cumulative differential in the Figure). However, it is noteworthy that in 1995 and 1996 the FES records a drop in the education differential for those who left full time education after 18. This is consistent with the increase in the supply of highly educated workers that has occurred in the UK since the late

for example, that a person declaring that full time education ended at 16 is observed with a degree.

[16]Note that the composition of the 16 education group has been changing because of changes in the statutory levels of education. This has been raised from 14 to 15 and finally to 16. The reforms did have a strong effect on the years of education achieved.

Year	Age left full-time education	**No Quals**	Qualifications achieved **O-Levels**	**A-Levels**	**Degree**
1978-1980	16	0.620	0.229	0.077	0.074
	17-18	0.159	0.196	0.254	0.390
	19+	0.051	0.048	0.093	0.809
1981-1983	16	0.600	0.220	0.089	0.091
	17-18	0.151	0.228	0.227	0.394
	19+	0.083	0.049	0.096	0.772
1984-1986	16	0.594	0.197	0.084	0.126
	17-18	0.101	0.308	0.270	0.321
	19+	0.069	0.054	0.124	0.752
1987-1989	16	0.509	0.250	0.110	0.131
	17-18	0.079	0.340	0.282	0.299
	19+	0.054	0.057	0.137	0.752
1990-1991	16	0.431	0.324	0.120	0.125
	17-18	0.060	0.357	0.348	0.235
	19+	0.064	0.089	0.152	0.695

Table 3.1: The relationship between age left full time education and observed qualifications in the GHS

	Education		
Year of Birth	**16**	**17-18**	**19+**
1926-30	86.48	6.74	6.78
1931-35	83.82	8.55	7.63
1936-40	80.83	9.27	9.90
1941-45	76.09	11.25	12.66
1946-50	70.74	13.18	16.08
1951-55	63.77	16.03	20.21
1956-60	64.30	17.63	18.07
1961-65	62.47	16.87	20.66
1966-70	57.07	20.11	22.81
1971-72	53.02	16.11	30.87

Table 3.2: Proportion of cohort in each education category (FES)

1980s. However, this graph alone shows no evidence that this fall is not cyclical.

Figure 3.3 reports age based median log(wage) differentials by age (35 or over versus less than 35) for both datasets. In the FES the differential rose from 0.03 in 1978 to 0.17 by 1995 and from the GHS the differential shows a very similar pattern, going from 0.03 in 1978 to 0.14 by 1991. If anything, the FES series displays a more cyclical pattern but this graph clearly demonstrates that the wages of older workers have risen rapidly relative to those of their younger counterparts. This increasing gap is open to a number of interpretations. In particular, it could mean that the returns to experience have increased, that there are significant cohort effects on wages and/or that “macro” factors are adversely affecting younger workers. In the next section we report estimates suggesting that these changing age differentials can be represented by cohort effects on wages which are predicted to persist over the life-cycle.

As noted earlier an important part of the rise in wage inequality has occurred within-groups. Figures 3.4 and 3.5 show the importance of this, documenting significant rises in wage inequality within education groups in both the FES and the GHS. The graphs report 90-10 log(wage) differentials within education groups and, in all cases, there are rises in the 90-10 differential. This is true for the comparisons based on the rather coarse education definitions in the FES data and for the qualifications based measures from the GHS data. These within-group trends in wage dispersion, and their different evolution across education definitions, stress the need to consider results based on both data sources and we examine this in some detail below.

4. Results

We now report results based on the estimation of equation 2.2 for a number of quantiles. The equations are estimated from FES and GHS data using the years

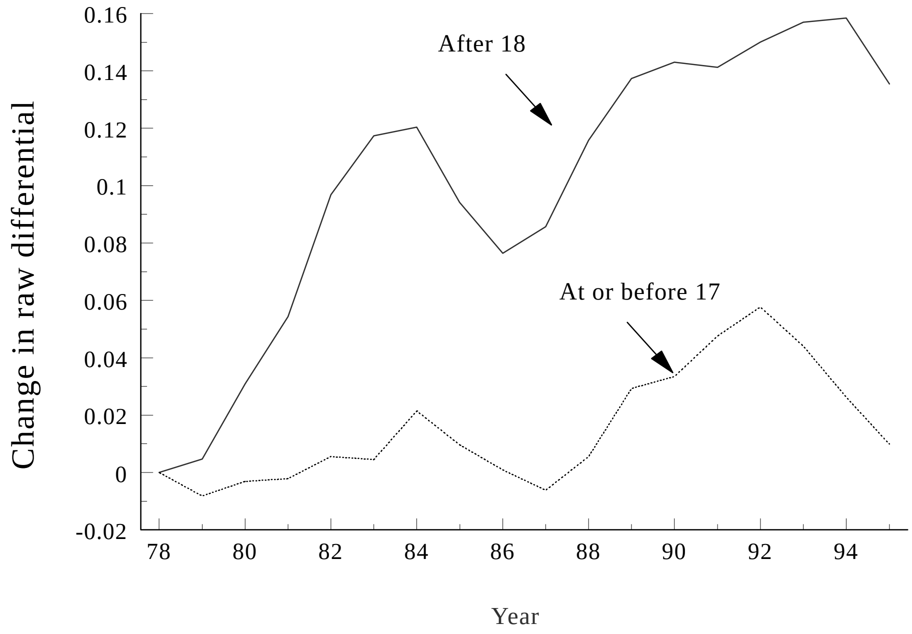

Figure 3.1: Cumulative changes in log(wage) differentials by education (FES)

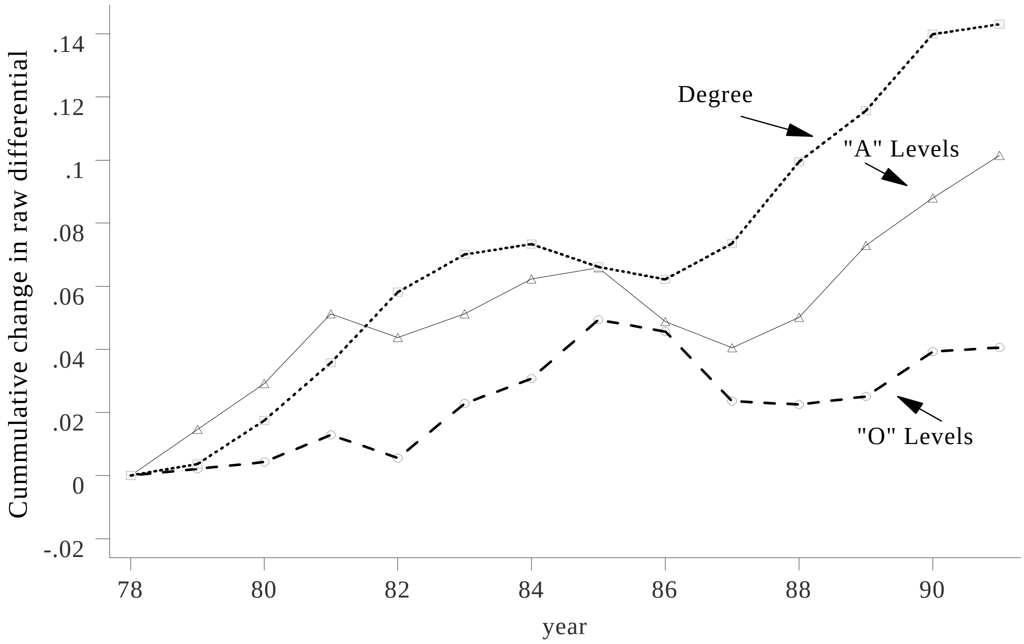

Figure 3.2: Cumulative changes in log(wage) differentials by education (GHS)

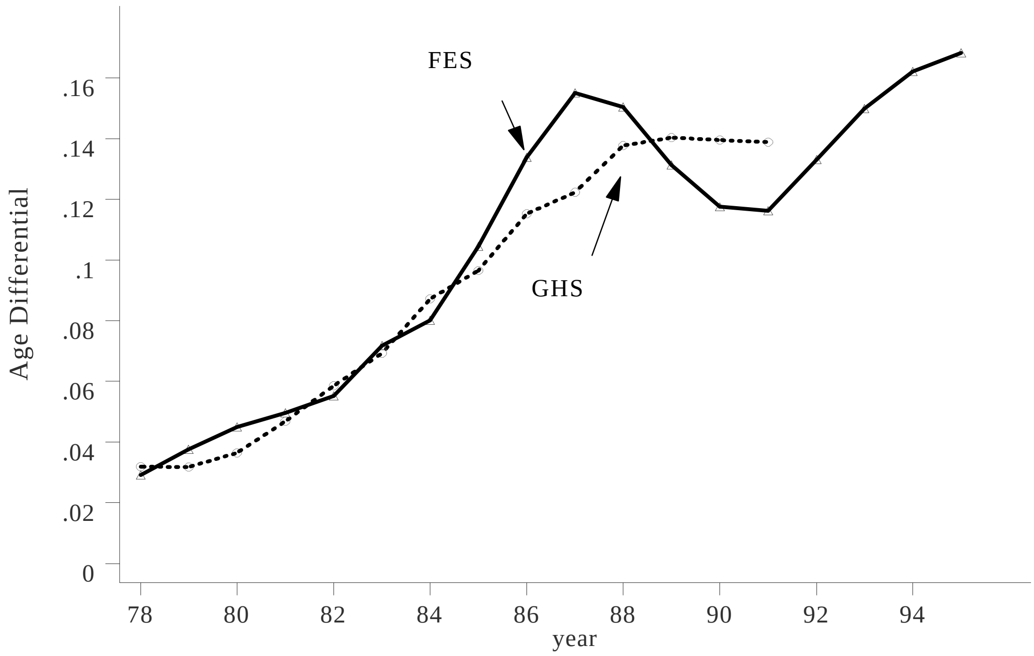

Figure 3.3: Difference in median log(wages) by age (aged 35 or over relative to under 35)

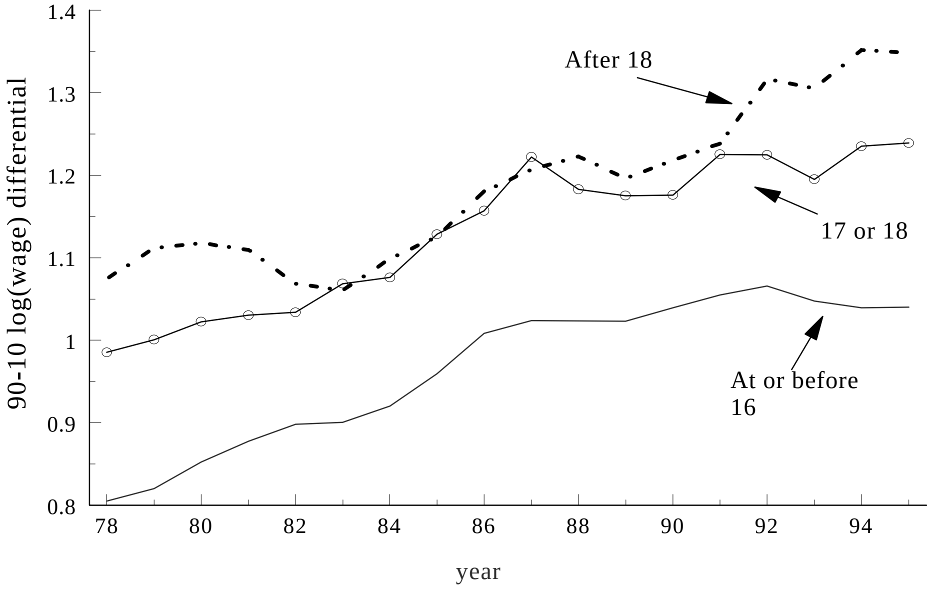

Figure 3.4: 90-10 log(wage) differentials within education groups (source FES)

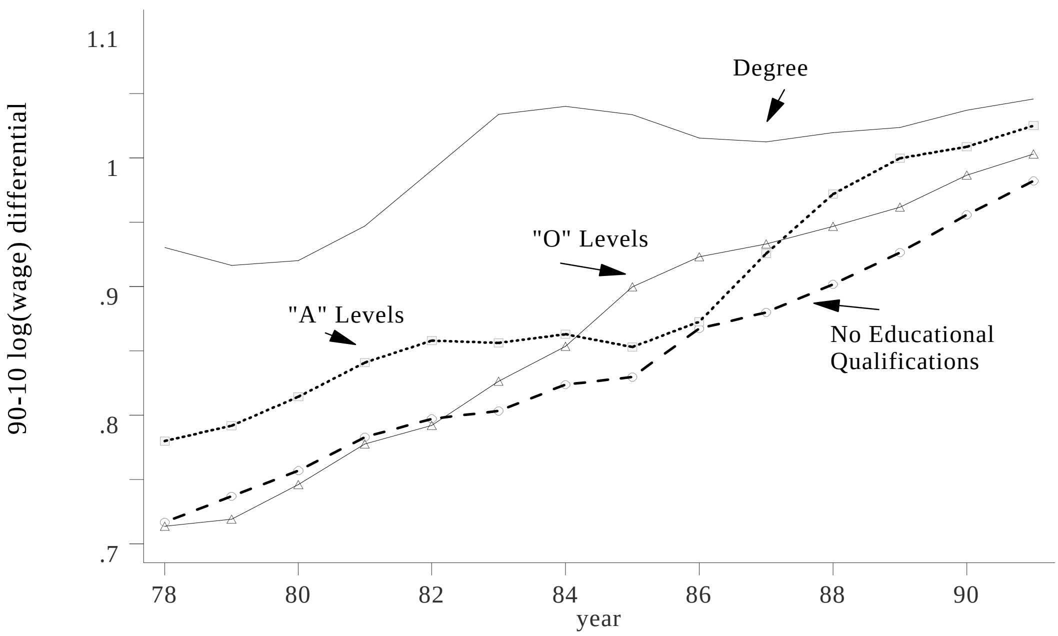

Figure 3.5: 90-10 log(wage) differentials within education groups (source GHS)

of education variable and the GHS data using the qualifications variable. In what follows we discuss the conclusions that can be drawn by our simple representation of the changes in the distribution of wages.

First, however, we provide evidence that our simplified model fits the aggregate moves in wage dispersion. We thus compare three specifications. One is a saturated model where all possible cohort age interactions are included as explanatory variables in the quantiles. The second model includes up to fifth order polynomials in age, cohort and (age $\times$ cohort) as well as unrestricted additive time effects. The third specification is our simple restricted model with just a cubic in age and cohort and orthogonal time effects (by construction these last include no trends). All specifications were estimated separately for each quantile of each education group[17].

4.1. The fit of the model

4.1.1. Estimating the distribution of wages using quantiles

Conditional and unconditional quantiles are related in the following way

$$q = \Pr(w < w^q) = \int_{R(z)} \Pr(w < w^q \mid z) \quad dF(z) \tag{4.1}$$

where $F(z)$ is the distribution of the observed vector of characteristics z (i.e.. cohort, education and age), $R(z)$ is the range of z and w^q is the point corresponding to the qth quantile of the unconditional distribution. Using the relationship expressed in equation 4.1 we can easily construct predicted unconditional distributions from the predicted conditional quantiles. Similarly, by changing the conditional quantiles $\Pr(w < w^q| \ z)$ or their weights $dF(z)$, we can construct counterfactual distributions of wages which will allow us to assess the impact of

[17] We have checked that all restricted conditional quantiles satisfy monotoncity conditions which is a necessary condition for them to be part of a well behaved distribution

specific variables on the observed changes. We use this idea later in the paper and the full procedure used is discussed in the appendix[18].

4.1.2. Comparing alternative specifications

Figures 4.1 and 4.2 show how well the two restricted specifications (ours and the one including some cohort age interactions - see appendix) do *vis a vis* the predictions from the saturated model. The solid line, is computed using the saturated model, whose predictions are simply the relevant order statistics of each year, age and education cell. The dashed line with circles uses the estimates from the third order polynomial model including the cohort age interactions. The solid line with squares, is computed using the estimates from our restricted model. The interdecile range we compute is a function of all estimated quantiles and is not fitted directly from the data. Looking at Figure 4.1 first, it is clear that there is no substantial difference in fit between any of the three models. At each year all three lines are within 0.02 points of each other. This comparison suggests that omitted interactions between age and cohort play practically no role in explaining the overall changes in wage dispersion.

Figure 4.2 considers the fit of the models within each education group. The first panel, which looks at those leaving school at or before 16 shows the increase in dispersion explained by the three models to be almost identical. In the next two panels which look at those workers with more education there is some gap between the saturated model and the two restricted specifications. Again, the differences are small except in 1989 and 1990. Moreover, the difference between 1978 and 1995 in wage dispersion within each of the groups is entirely picked up by the restricted model. The differences between the saturated model and the two restricted specifications has probably more to do with the small cell sizes

[18] This methodology has similarities to that used by DiNardo, Fortin and Lemieux, (1996).

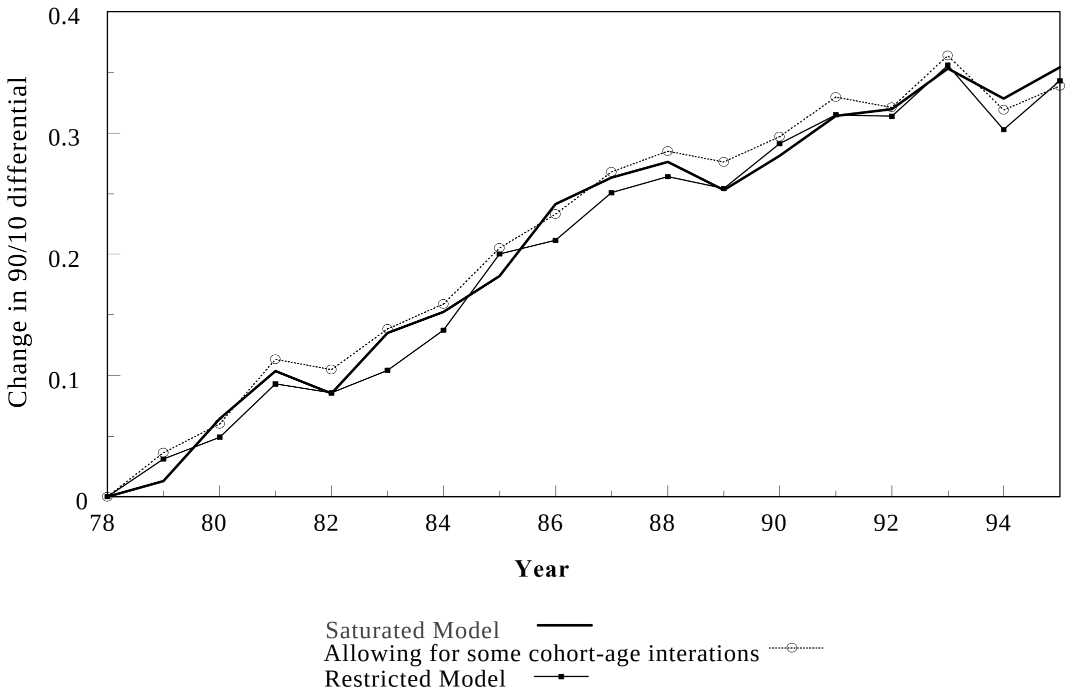

Figure 4.1: Predicted change in the 90-10 differential of log(hourly earnings) from 1978

in the two highest education groups than with the accuracy of the model. This view is corroborated by the fact that there is little difference in the fit of the intermediate model (which includes a number of cohort-age interactions) and that of our specification which excludes all cohort-age interactions (see the appendix for the precise specifications).

Thus the larger differences are between the saturated model and either of the two restricted models. We examine this issue further in Table 4.1, which presents χ^2 tests that the saturated model (represented by the solid line in Figures 4.1

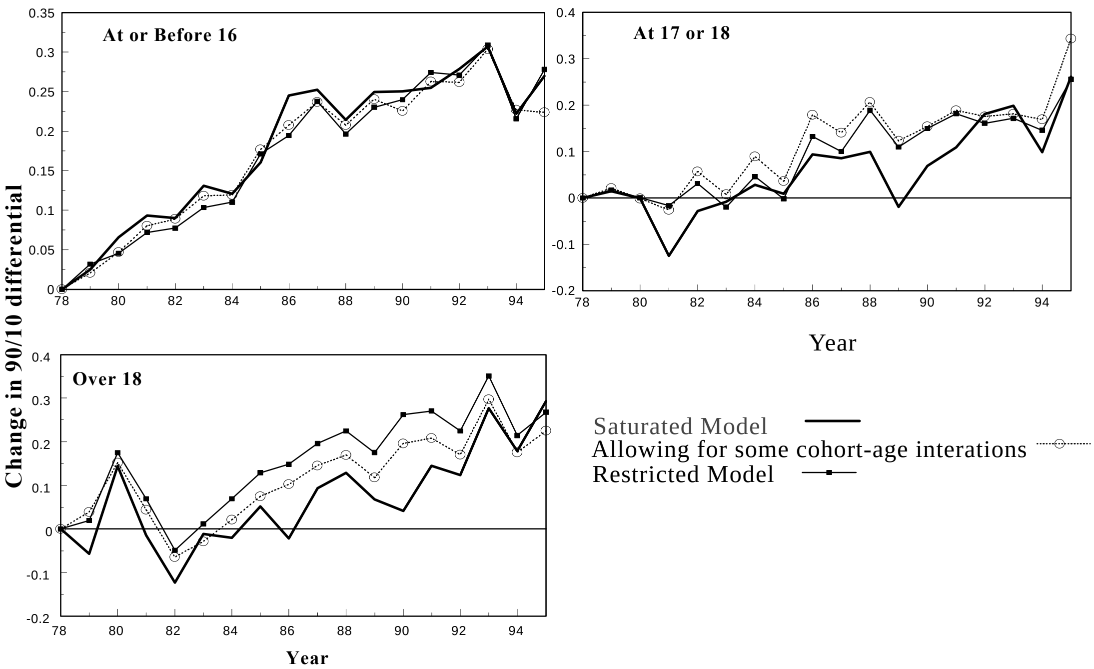

Figure 4.2: Predicted change in the 90-10 differential of log(hourly earnings) from 1978 by age left full time education

Percentile	**Education group**		
			Left school
	before 16	*at 17 or 18*	*After 18*
	(643 dof)	*(633 dof)*	*(623 dof)*
10th	768.81	1223.70	1414.11
	0.000	*0.000*	*0.000*
25th	596.43	573.15	685.05
	0.905	*0.957*	*0.043*
50th	530.85	353.43	543.97
	0.999	*0.999*	*0.999*
75th	678.23	432.81	817.289
	0.163	*0.999*	*0.000*
90th	980.37	1014.96	1842.58
	0.000	*0.000*	*0.000*

Probability values under the null given in italics

Table 4.1: Chi squared Goodness of Fit Tests (FES data)

and 4.2) can be restricted to our most restrictive specification. While the tests reject at the tails of the distribution, at the centre of the distribution there is no evidence of misspecification.[19]

At first glance, these results seem to be in conflict with some of the US literature which places a lot of emphasis on the growing difference between the pay of older versus younger workers over time as a major cause of rising wage inequality with a notable exception the paper of MaCurdy and Mroz (1991) (see, for example, Bushinsky, 1994, Katz and Murphy, 1992, or Juhn, Murphy and Pierce, 1993). However, the basic underlying facts are the same in our data; age differentials have grown over time (see Figure 3.3 above). Our modelling approach and the resulting good fit of the model shows that these changes can be represented

[19] In fact the test statistics may seem perhaps too small. This may be due to the small cell sizes for some cohorts at the higher education groups. However, the statistics are small even for the lower education group where there are plenty of observations per cell.

by cohort effects. Of course the functional forms we use can always be re-written as functions of time and age. Hence the interpretation that one attaches to the simple representation of these changes is to an extent arbitrary and differs in no way from the usual practice of imposing identifying assumptions to interpret the data. A good way of expressing our findings is that, given the presence of cohort effects, we find no evidence of changes in the life-cycle profiles.

Broadly speaking, our results imply that the changes can be simply described by changes in the rate of growth and dispersion of entry level wages across successive generations of workers, which then persist over the lifecycle. The rest of the story is explained by the changes in the education differentials. We now examine these changes in greater detail.

4.2. Cohort Effects and Changing Returns to Education

The coefficient estimates and associated standard errors on the cohort variables for both data sets are reported in Tables 4.2 and 4.3 and these effects are shown graphically in Figures 4.3 to 4.5.[20] Note that for the purposes of improving precision we have imposed restrictions across the quantiles of the two middle qualifications (O-levels and A-levels) in the GHS. The corresponding estimates appear in this table under the title "*Intermediate Qualifications*".

The growth in median wages across different generations of workers is shown in Figure 4.3. The lines in these graphs show the degree to which life cycle earnings profiles of successive generations are shifting upwards. The first panel uses FES data on hourly earnings using age left full time schooling as the education variable. The education group to have received the largest wage gains *in total* is shown to be the middle one (those leaving school between 17 and 18) with the top and the

[20] Parameter estimates using GHS data based on the years of education measure (as in the FES) are not reported but are available from the authors on request.

bottom groups receiving the same increase over the entire period. However, the shape of the growth rates of these two groups is very different with most of the growth in the lower education group occurring across cohorts born up to about 1956 and an acceleration of growth rates of the highly educated cohorts born after 1940. The second panel uses GHS data on weekly earnings, defining education by qualifications[21]. This shows that the degree group have experienced the highest growth in wages and those with no formal qualification the lowest. Amongst those with an intermediate qualification it appears to be the case that those with "O" levels have done relatively better, compared to earlier generations than those with "A" levels.

Finally this figure shows how our model interprets the observed increases in age differentials. The cohort profiles for all education groups are concave at least for cohorts born after 1952. Hence the relative gap between the wages of older cohorts and younger ones is increasing. This effect is most dramatic for the lower education groups

Figure 4.4 translates these relative growth rates into education differentials for 40 years olds across successive generations. [22] The FES data shows differentials between the post 18 and the 16 and below group to be falling across cohorts born before 1940, rising across cohorts born between 1940 and 1968, and then flattening off. Differentials between the middle and the lower education group are rising continuously across all cohorts in our data. The GHS data shows a steady increase in differentials between those with some qualification and those without. In comparing the GHS results to the FES ones recall that the lowest education group in the FES includes most of those who have an O-Level qualification. Once

[21] A comparison of the FES data with GHS data using years of schooling rather than educational attainment yields similar results to those just described. Results available from the authors on request.

[22] The changes in differentials will reflect both changing returns over time as well as the effect of changes in the quality of education.

this is taken into account the implications from the two data sets are not that different. Below we present evidence showing that the increased heterogeneity in terms of qualifications of the FES educational groups goes some way to explaining the observed changes in within group dispersion. The GHS also shows slight evidence of a "flattening off" of the differential for the highest qualification group. However, because the data does not include cohorts born after 1968, a complete comparison with the FES cannot be made.

Figure 4.5 shows how differentials *within* education groups have changed across cohorts, again using both FES and GHS data. The hypothesis that the slopes of cohort profiles are the same across quantiles is strongly rejected and from the figure the differences are quantitatively very important; i.e. there are significant increases in within group wage inequality across cohorts: Wage dispersion within each education group has increased over successive generations, with a possible exception being the degree group; we return to this issue later.

4.3. The Effects of Education on Wage Dispersion

The results above suggest that education has had interesting and complex effects on changes in the distribution of wages. To fully understand these we construct a number of counterfactual experiments which illustrate the magnitude of the effects of education on between-group and within-group dispersion. Such decompositions are best achieved using the variance of wages. This is because the within group and between group interquantile ranges (say) do not necessarily add up to the interquartile range of the unconditional distribution.

Our estimated quantiles allow the computation of any moment of the distribution in the following way: The set of predicted quantiles was used to estimate the entire conditional distribution of wages for each year, age and education cell as described earlier. These conditional distributions were then used to predict the

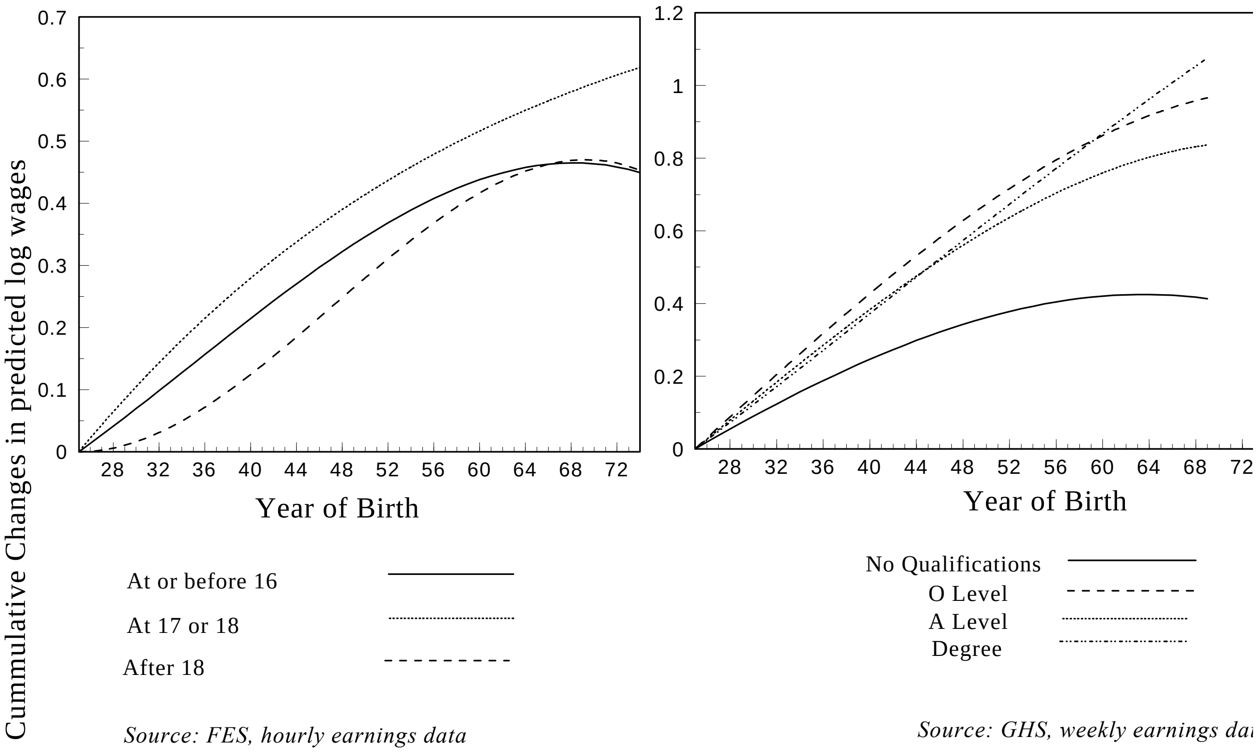

Figure 4.3: Cumulative growth in median wages by education group and year of birth

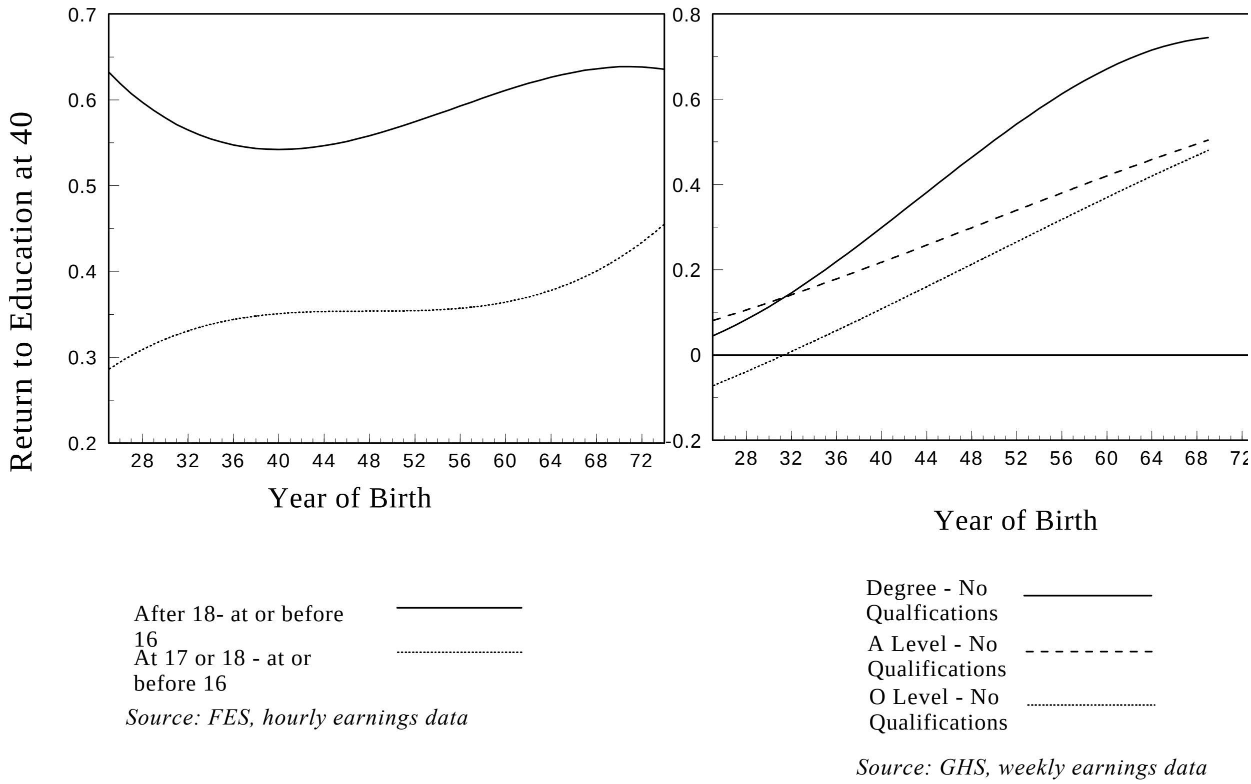

Figure 4.4: Educational differentials by year of birth

	Percentile				
	10th	25th	50th	75th	90th
Left School at or before 16					
Constant	1.063	1.109	1.165	1.212	1.338
	0.027	*0.021*	*0.021*	*0.023*	*0.032*
Cohort	0.045	0.077	0.143	0.184	0.203
	0.010	*0.008*	*0.010*	*0.009*	*0.017*
$Cohort^2$	0.013	0.015	0.006	0.013	0.013
	0.007	*0.006*	*0.007*	*0.006*	*0.012*
$Cohort^3$	-0.004	-0.005	-0.004	-0.006	-0.006
	0.002	*0.001*	*0.001*	*0.001*	*0.002*
Left School at 17 or 18					
Constant	1.216	1.166	1.098	1.233	1.369
	0.077	*0.054*	*0.046*	*0.051*	*0.065*
Cohort	0.136	0.158	0.197	0.233	0.272
	0.050	*0.042*	*0.033*	*0.035*	*0.037*
$Cohort^2$	-0.052	-0.036	-0.024	-0.041	-0.067
	0.030	*0.025*	*0.020*	*0.020*	*0.023*
$Cohort^3$	0.004	0.003	0.001	0.005	0.010
	0.005	*0.005*	*0.004*	*0.003*	*0.005*
Left School after 18					
Constant	0.926	0.952	1.087	1.214	1.093
	0.077	*0.052*	*0.049*	*0.048*	*0.057*
Cohort	-0.028	-0.070	0.063	0.111	0.153
	0.048	*0.036*	*0.031*	*0.031*	*0.033*
$Cohort^2$	0.060	0.110	0.056	-0.025	0.046
	0.027	*0.018*	*0.016*	*0.017*	*0.020*
$Cohort^3$	-0.011	-0.018	-0.011	-0.004	-0.007
	0.005	*0.003*	*0.003*	*0.003*	*0.003*

Notes
1. Dependent variable: Log hourly wage. Sample Size: 16: 39218, 17/18: 7847, 19+ 8436
2. Cohort Coefficients scaled so that cohort = (year of birth-1930)/10
3. Standard errors in italics

Table 4.2: Quantile Estimates of the Cohort Effects (FES data)

	Percentile				
	10th	25th	50th	75th	90th
No Qualifications					
Constant	4.715	4.841	4.887	4.725	4.844
	0.037	*0.030*	*0.024*	*0.029*	*0.035*
Cohort	0.113	0.123	0.173	0.237	0.280
	0.010	*0.008*	*0.008*	*0.009*	*0.013*
Cohort^2	-0.029	-0.042	-0.014	-0.030	-0.038
	0.010	*0.007*	*0.006*	*0.007*	*0.012*
Cohort^3	0.002	0.006	-0.002	0.006	0.005
	0.002	*0.002*	*0.002*	*0.001*	*0.003*
Intermediate Qualifications					
Constant	4.509	4.509	4.528	4.556	4.447
	0.033	*0.031*	*0.029*	*0.033*	*0.052*
Cohort	0.183	0.245	0.291	0.330	0.480
	0.017	*0.016*	*0.015*	*0.020*	*0.045*
Cohort^2	-0.008	-0.016	-0.006	0.024	-0.038
	0.012	*0.010*	*0.010*	*0.013*	*0.028*
Cohort^3	-0.003	-0.002	-0.004	-0.012	0.000
	0.002	*0.002*	*0.002*	*0.002*	*0.005*
High Qual	0.136	0.114	0.139	0.166	0.177
	0.013	*0.012*	*0.013*	*0.023*	*0.03*
High Qual X	-0.032	-0.017	-0.029	-0.037	-0.043
Cohort	*0.007*	*0.006*	*0.006*	*0.010*	*0.015*
Degree					
Constant	3.944	4.208	4.418	4.504	4.549
	0.043	*0.032*	*0.028*	*0.035*	*0.057*
Cohort	0.197	0.264	0.247	0.250	0.317
	0.027	*0.022*	*0.019*	*0.022*	*0.030*
Cohort^2	0.053	-0.005	0.005	0.010	-0.034
	0.017	*0.014*	*0.011*	*0.014*	*0.018*
Cohort^3	-0.008	0.000	-0.001	-0.001	0.008
	0.003	*0.002*	*0.002*	*0.002*	*0.003*

Notes
1. Dependent variable: Log weekly earnings. Sample Size: No qualifications: 24393, Intermediate: 16746, Degree: 13017
2. Cohort Coefficients scaled so that cohort = (year of birth-1930)/10. High Qual: O-levels and A-levels.
3. Standard errors in italics

Table 4.3: Quantile Estimates of the Cohort Effects (GHS data)

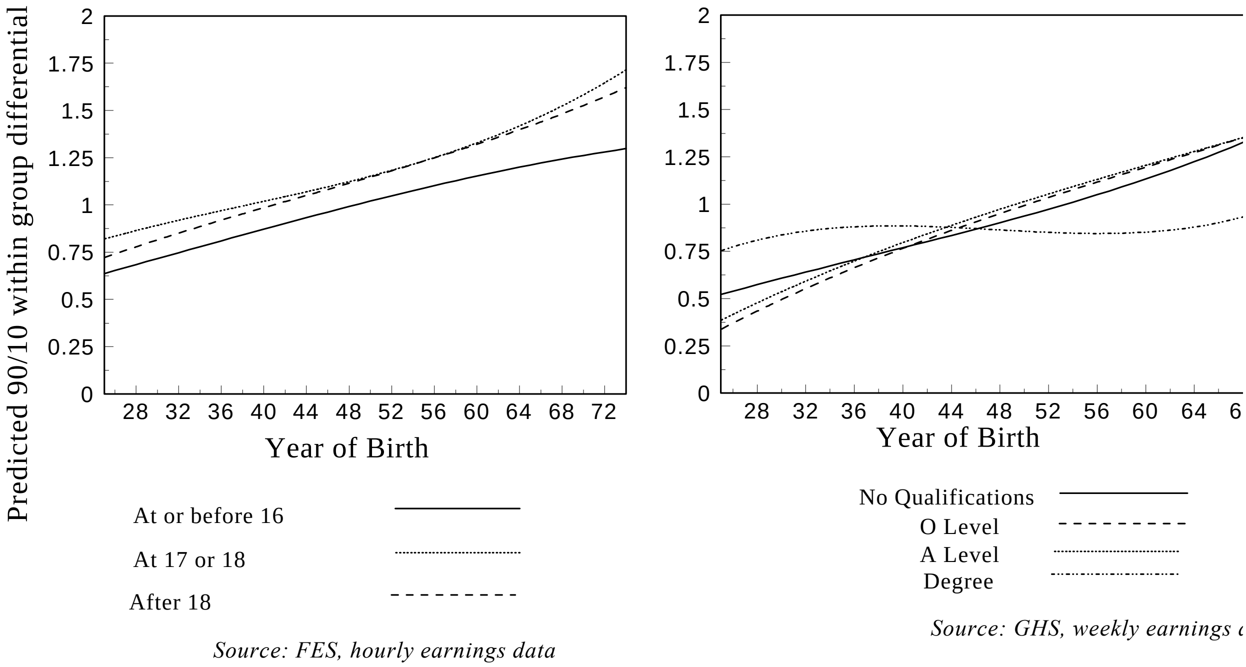

Figure 4.5: 90-10 differentials within education groups by year of birth

full set of group means and within group variances of wages. The overall within and between group components to changes in wage inequality, are then easily estimated by switching off within group inequality (see the appendix). Our model predicts very accurately the changes in the sample standard deviation of wages.

Figure 4.6 shows the relative changes in between and within group standard deviation of wages from 1978 using the FES sample. The last panel which looks at the whole FES sample suggests that over 2/3 of the overall increase in the standard deviation of wages can be explained by the changing returns to education, changes in the educational composition of the workers and cohort effects. In addition it shows that most of the increase in within group dispersion occurred between 1984 and 1988.

Now note that age effects can also be important in explaining changes in wage dispersion if the age composition of a group is changing over time, even when the growth rate of wages with age is the same across cohorts. As we show below dispersion increases a lot over the lifecycle of a given cohort. In fact, amongst those leaving school at or before age 16, cohort and age effects explain about half of the overall increase in dispersion. This is driven both by the fact that this group is getting older on average and that the growth rate of wages of new entrants into the labour market is falling over time. For those workers with more education, cohort and age effects are, if anything, mitigating the increase in wage dispersion; this is because the average age of this group is declining rapidly. Thus for the high education group dispersion is rising sharply but this is all within group. This may also reflect the increasing heterogeneity in the ability distribution for that group as the number of individuals in that group has increased over time and is shown in table 3.1

Figure 4.7 decomposes the overall increase in wage inequality into that driven by education and that driven by cohort and age effects (due to changes in age

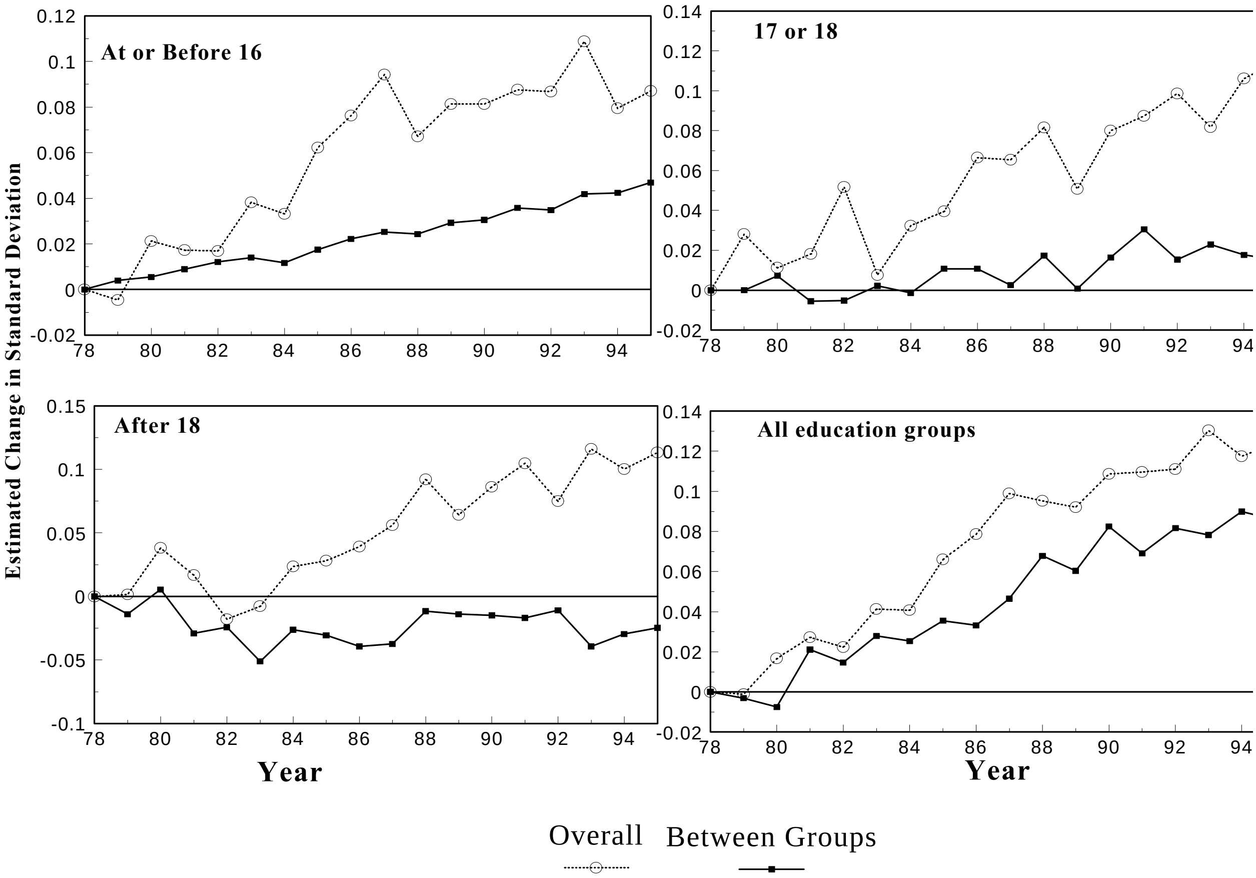

Figure 4.6: Within and Between group components to the overall change in the standard deviation of wages by year and age left full time education: Source: FES data

composition). The top line represents the increase in the standard deviation predicted if there was no change in within group wage inequality. The solid line represents the predicted increase, holding the age composition of the workforce constant and setting cohort effects to zero. This quantifies the part of the between group change due to the changing returns to education. Similarly the dashed line predicts the increase in wage inequality setting the composition and returns to education at the 1978 level. The figure shows that up to 1983 most of the increase in between group inequality is attributable to education, between 1983 and 1988 cohort effects play an increasingly important role in explaining wage inequality. Between 1988 and 1995, both cohort and education effects have played an equal role in the increase in wage inequality, although the pace of this increase is much slower than in earlier years. Finally, note the slight fall in the between education group dispersion from 1994 to 1995 shown by the solid line in figure 4.7. This is consistent with the observed fall in the unconditional returns to a degree shown in figure 3.1.

4.3.1. The Role of Educational Qualifications

There are a few sources of increasing heterogeneity within each education group. One is the set of qualifications received within each of the groups based on the years of education measure; another is the possible changes in the distribution of other skills across cohorts and the last is changes in the composition of ability in each group solely driven by the fact that the process determining educational choice may have changed. We now look at our results using GHS data, which has information on educational qualifications, to assess whether part of the increase in within group inequality is a result of the first source, i.e. increasing heterogeneity in terms of qualification of education groups defined by years of schooling.

There are two hypotheses that we wish to examine. First we should expect

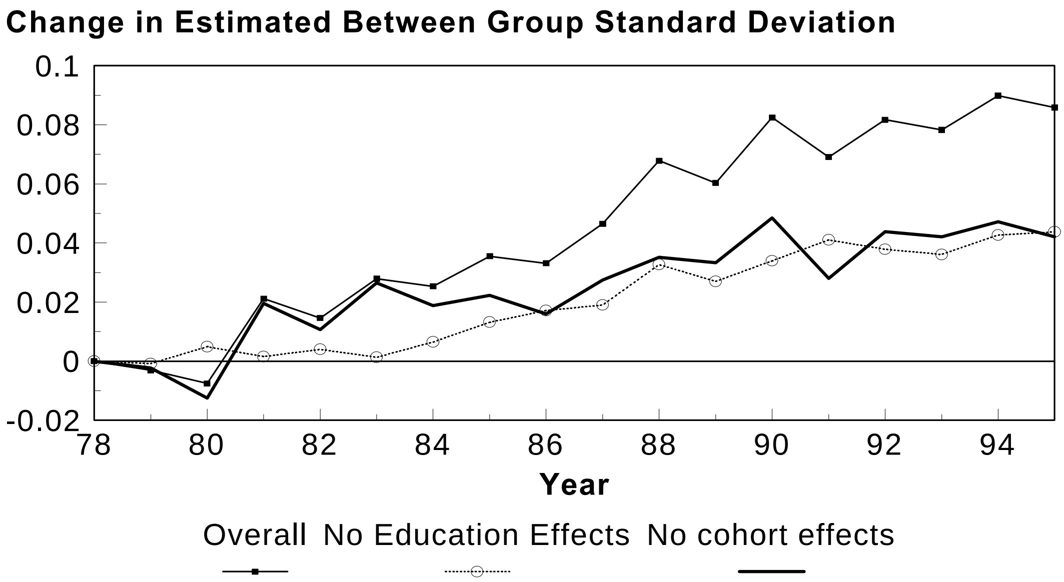

Figure 4.7: The impact of changes in the effect of age at the end of education on the changes in between group inequality, FES data

cohort effects to be less important when we use GHS data as the changing government policy towards education will be reflected in part by a shift in the distribution of qualifications received by children leaving school at a particular age. Second we should expect the overall increase in within group dispersion to be less important once one controls for educational qualifications rather than just years of schooling.

Figure 4.8 thus reports the same decompositions as figure 4.7 but using GHS data. This shows the role of education (measured as the age that full time education ended) in shaping changes in “between group” inequality to be much the same in the GHS as in the FES sample.

Given this similarity, in Figure 4.9 we compare the growth in dispersion within age and education groups across generations for workers using the two different education measures on GHS data. This allows us to assess directly whether part of the increase in within group wage dispersion shown in figure 4.5 can be explained by the increasing heterogeneity of qualifications received by each group. The first panel of the figure shows within group wage inequality to be growing at more or less the same rate across education groups when these are defined by years of schooling. The second panel shows that the overall increase in within group inequality is smaller when defined by educational qualification. More interestingly, using years of education rather than qualifications obscures the fact that there has been no increase in within group inequality amongst the group of workers with a degree. The confusion occurs because the post 18 group includes individuals with high school qualifications such as A-Levels as well as degrees. Conditional on median cohort effects, within group inequality has increased among high school graduates, but not much it seems among University graduates. On the low end of the educational ladder, it seems that part of the increase in within group inequality is due to the fact that the proportion of individuals with just statutory years of education who leave with some qualifications (O-Level) has increased (from a very

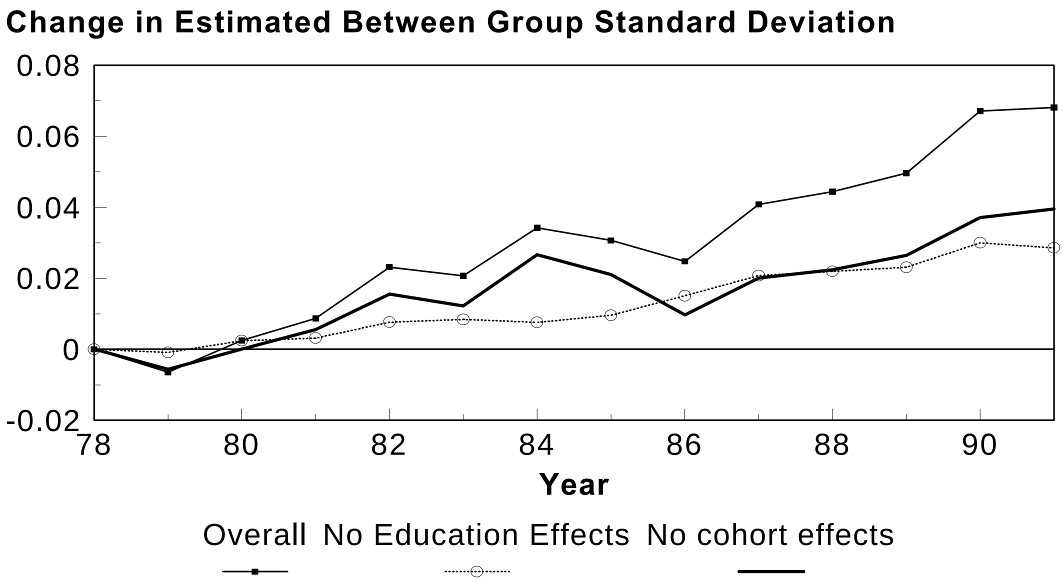

Figure 4.8: The impact of changes in the effect of age at the end of education on the changes in between group inequality, GHS data

low base). This translates into a higher increase in within group inequality when using the years of education measure.[23]

4.4. The Distribution of Real Wages Over the Life-cycle

In Tables 4.4 and 4.5 we present the coefficients on age for each quantile by education group for the FES and the GHS respectively. The age functions are quantitatively very different across quantiles and the differences are highly significant.

The estimates imply that dispersion of wages increases over the life-cycle, particularly for the lower education group. A number of models of wage determination imply this evolution of wage dispersion over the life-cycle. This is consistent with a simple human capital model where individuals learn on the job at a stochastic rate with a possibly heterogeneous mean. It is also consistent with a model where information about individuals is revealed on the job.[24] Suppose for instance that initial pay, at labour market entry is quite homogeneous for individuals with similar observable characteristics, as employers have limited information on individual productivity. As individuals' ability gets revealed, the variance of productivity is transmitted into the dispersion of wages.

The implied life-cycle growth of wages are presented in figure 4.10 (FES) and 4.11 (GHS). The results are quite consistent across the two data sets: Median wages grow for all education groups, but they grow substantially faster for the higher education individuals: The returns to experience seem to increase with education. This pattern is even stronger when we look at the top deciles. At the bottom of that distribution wages grow at a very low rate over the lifecycle. For the two lowest education groups in the FES the growth is about 25% over

[23] See Ginther (1994) for some US work that finds larger increases in wage inequality within lower skill groups.

[24] See Jovanovic (1979), Farber and Gibbons (1997) and Baker, Gibbs and Holmstrom (1994).

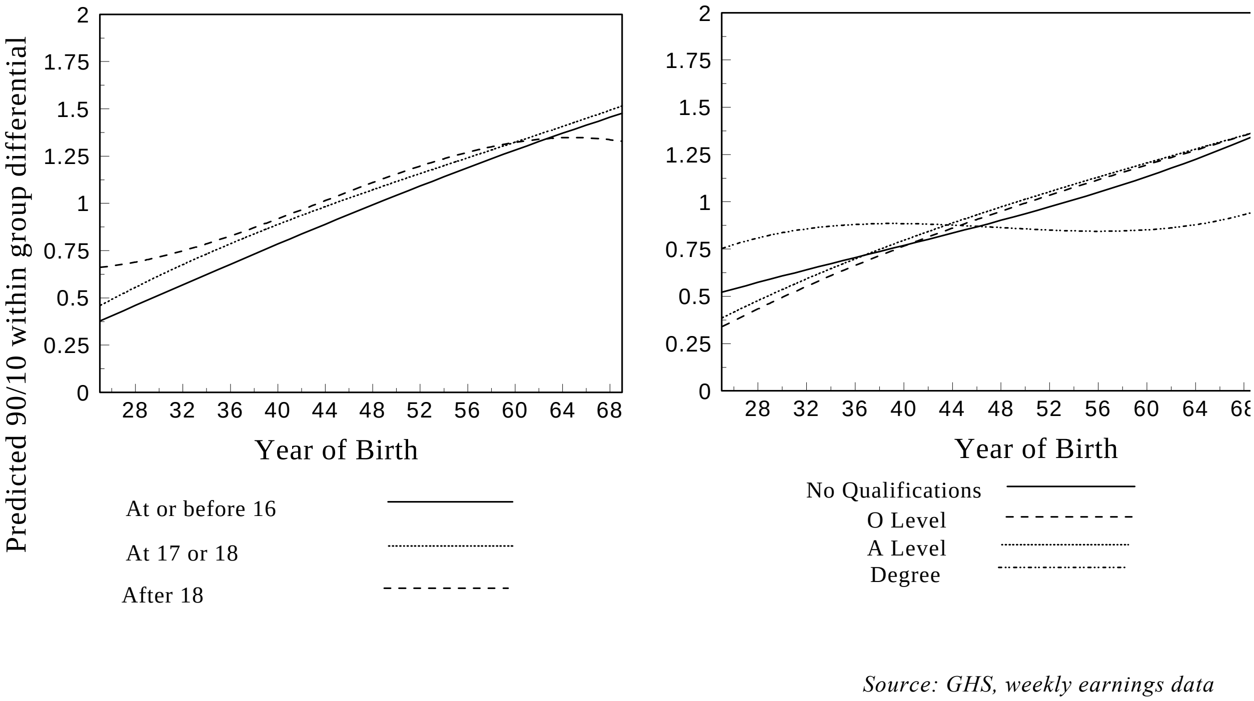

Figure 4.9: The role of qualifications in affecting within group inequality

38 years. For the top education group, even the lower decile of the distribution grows substantially.

Figure 4.12 shows the implication of the estimates for dispersion over the lifecycle within each education group. Quite clearly dispersion increases in all education groups but most in the higher ones, with a notable exception the degree group in the GHS where the increase in dispersion is moderate, particularly early on in the life-cycle.

The results are consistent with the basic human capital model: Workers with some post 18 education have much less experience at age 22 (when we start observing). At that point, depending on the cohort they have no advantage or even a negative differential over those who left school earlier (not shown). This is particularly true for earlier cohorts. The returns to age (a proxy for experience) are then much higher for those in the highest education group and keep rising for the entire working life. This is also consistent with the idea that workers who invested more in pre-labour market education continue investing when working. On the other hand it seems that the wage returns to experience for low education workers are much lower although they seem to be positive throughout.

4.5. The Cyclical Time Effects

The cyclical time effects play only a minor role in explaining within sample differences in wage dispersion. These are defined to be common detrended effects across cohorts on each quantile of the wage distribution. They can differ across education groups. The maximum implied impact we observed from such effects on the unconditional wage distribution is 4 percentage points on the interdecile range. Their relative impact is larger towards the end of the 80s and the early 90s.

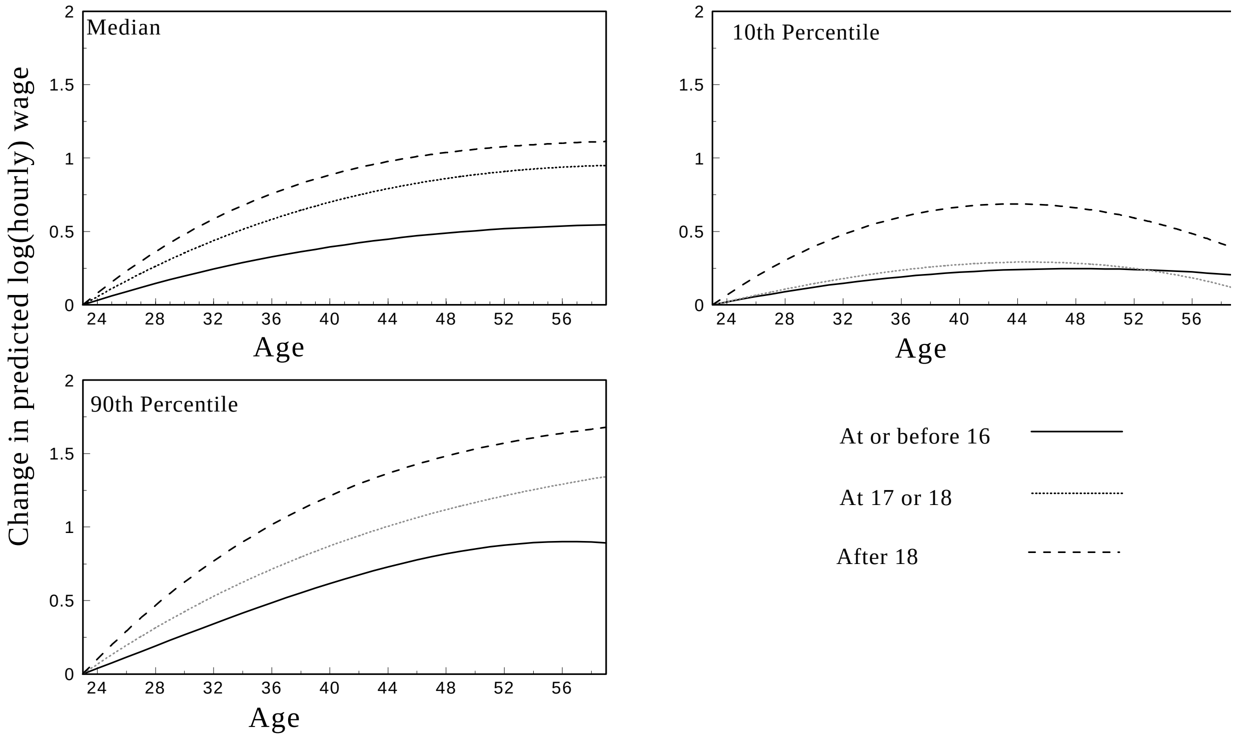

Figure 4.10: Wage growth over the lifecycle by percentile and age left full time education: Source FES data

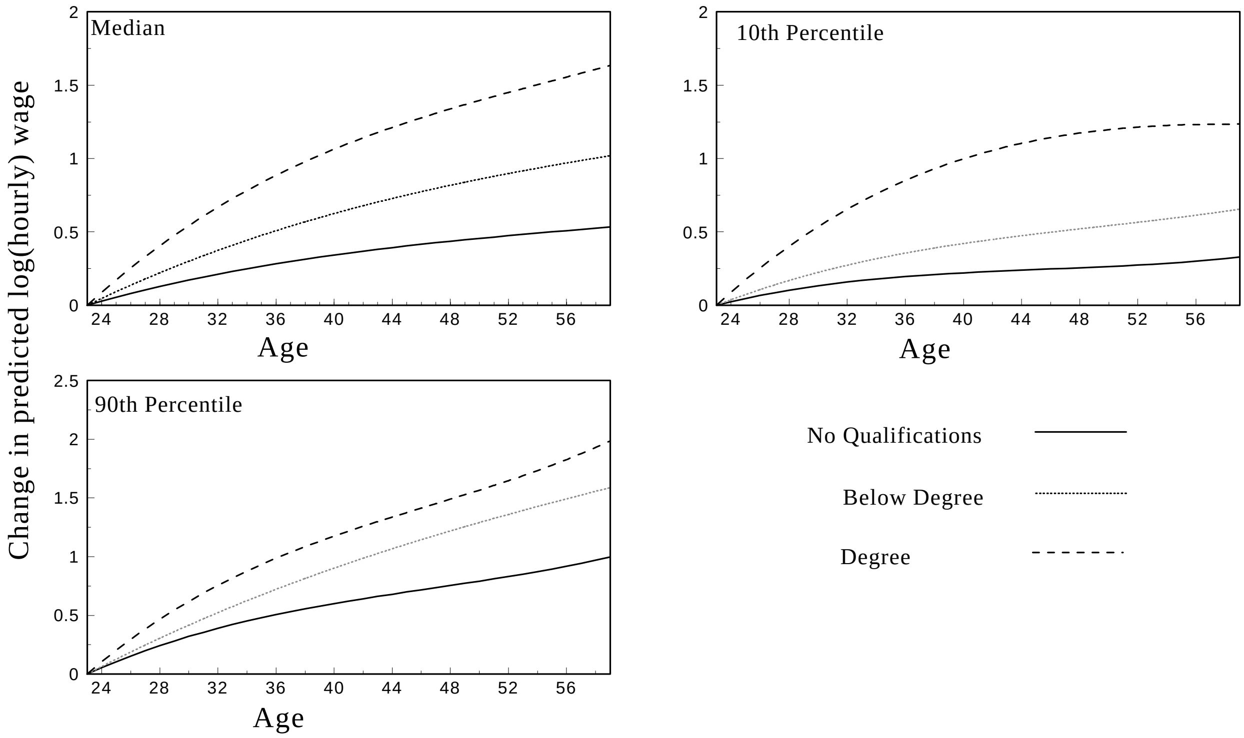

Figure 4.11: Wage growth over the lifecycle by percentile and education qualificaiton: Source GHS data

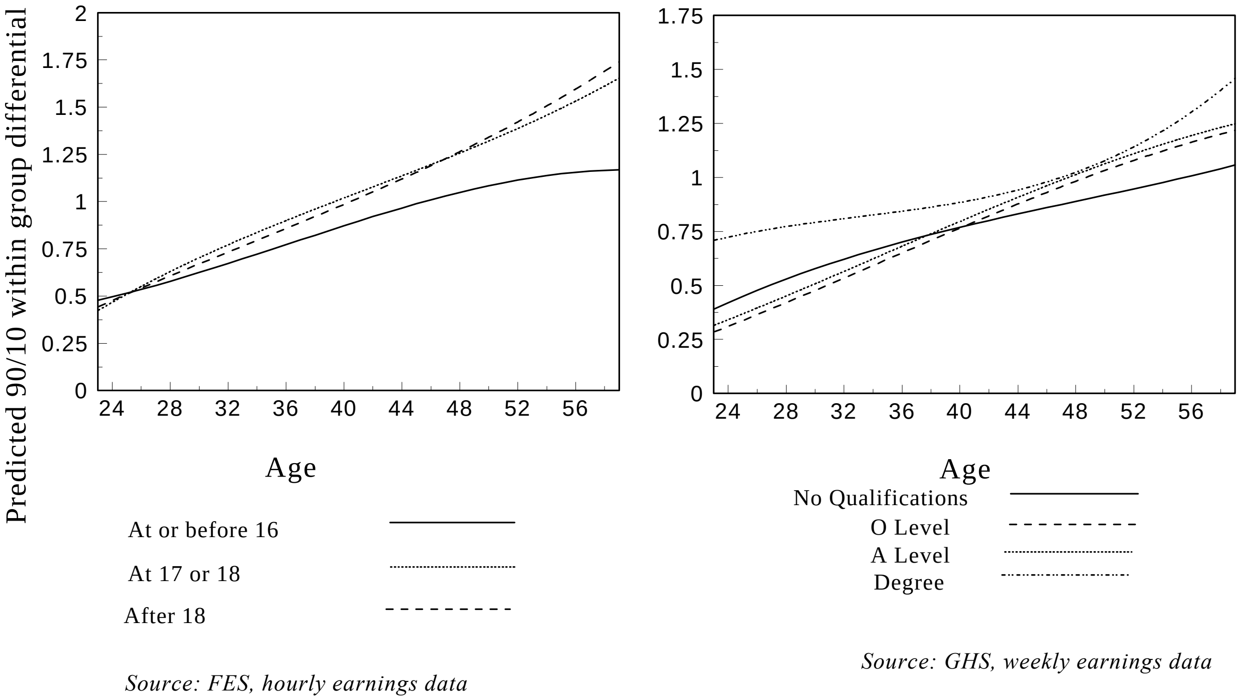

Figure 4.12: Within group wage inequality over the lifecycle by education group

	Percentile				
	10th	25th	50th	75th	90th
Left School at or before 16					
Constant	1.063	1.109	1.165	1.212	1.338
	0.027	*0.021*	*0.021*	*0.023*	*0.032*
Age	0.230	0.302	0.357	0.418	0.364
	0.040	*0.031*	*0.031*	*0.032*	*0.047*
Age^2	-0.045	-0.053	-0.060	-0.045	0.029
	0.022	*0.017*	*0.017*	*0.017*	*0.016*
Age^3	0.001	0.001	0.003	-0.001	-0.014
	0.003	*0.002*	*0.003*	*0.002*	*0.005*
Left School at 17 or 18					
Constant	1.216	1.166	1.098	1.233	1.369
	0.077	*0.054*	*0.046*	*0.051*	*0.065*
Age	0.245	0.405	0.648	0.723	0.753
	0.118	*0.080*	*0.064*	*0.068*	*0.090*
Age^2	-0.020	-0.026	-0.114	-0.134	-0.120
	0.067	*0.046*	*0.040*	*0.042*	*0.052*
Age^3	-0.008	-0.010	0.006	0.009	0.007
	0.010	*0.008*	*0.007*	*0.007*	*0.008*
Left School after 18					
Constant	0.926	0.952	1.087	1.214	1.093
	0.077	*0.052*	*0.049*	*0.048*	*0.057*
Age	0.824	0.931	0.957	1.009	1.183
	0.102	*0.073*	*0.064*	*0.068*	*0.075*
Age^2	-0.205	-0.204	-0.228	-0.229	-0.241
	0.062	*0.044*	*0.039*	*0.039*	*0.048*
Age^3	0.009	0.010	0.019	0.018	0.018
	0.010	*0.008*	*0.007*	*0.007*	*0.008*

Notes
1. Age Coefficients scaled so that age = (age in years-20)/10
2. See notes in Table 4.2

Table 4.4: Quantile Estimates of the Age Effects (FES)

	Percentile				
	10th	25th	50th	75th	90th
No Qualifications					
Constant	4.715	4.841	4.887	4.725	4.844
	0.037	*0.030*	*0.024*	*0.029*	*0.035*
Age	0.306	0.395	0.317	0.581	0.657
	0.048	*0.044*	*0.035*	*0.039*	*0.048*
Age^2	-0.104	-0.142	-0.061	-0.130	-0.177
	0.025	*0.022*	*0.018*	*0.021*	*0.027*
Age^3	0.013	0.018	0.005	0.013	0.022
	0.003	*0.003*	*0.003*	*0.003*	*0.005*
Intermediate Qualifications					
Constant	4.509	4.509	4.528	4.556	4.447
	0.033	*0.031*	*0.029*	*0.033*	*0.052*
Age	0.454	0.511	0.525	0.500	0.699
	0.048	*0.043*	*0.042*	*0.047*	*0.067*
Age^2	-0.117	-0.113	-0.080	-0.019	-0.086
	0.028	*0.024*	*0.025*	*0.029*	*0.043*
Age^3	0.013	0.012	0.006	-0.006	0.006
	0.005	*0.003*	*0.004*	*0.005*	*0.008*
Degrees					
Constant	3.944	4.208	4.418	4.504	4.549
	0.043	*0.032*	*0.028*	*0.035*	*0.057*
Age	1.048	1.026	0.946	1.033	1.258
	0.065	*0.045*	*0.040*	*0.047*	*0.080*
Age^2	-0.235	-0.267	-0.207	-0.219	-0.331
	0.040	*0.027*	*0.023*	*0.028*	*0.047*
Age^3	0.017	0.029	0.020	0.021	0.042
	0.007	*0.005*	*0.004*	*0.005*	*0.008*

Notes
1. Age Coefficients scaled so that age = (age in years-20)/10
2. See notes in Table 4.3

Table 4.5: Quantile Estimates of the Age Effects (GHS)

5. Some results with an alternative identification restriction

Our analysis has been based on imposing that life-cycle profiles are the same across cohorts and on normalising the time effects to be orthogonal to cohort and age effects. This attributes aggregate common trends to cohort and age effects. It is interesting to see if our basic conclusions are robust to an alternative identifying assumption which allows all time effects to be separated out from cohort and age effects. Thus based on the idea that life-cycle growth of wages may be due to active investments by individuals, and that such investments are likely to stop closer to retirement, we pursue the following strategy: We assume that all human capital acquisition stops at the age of 50; all observed wage growth is then attributed to general productivity growth and cohort effects beyond that point. We assume this is true for all education groups and all quantiles, which implies that life-cycle growth in dispersion also stops. Subject to this assumption the time and cohort effects can be identified from the wage growth of those older than 50. The advantage of these strong assumptions is that we can separate out differences in the life-cycle growth of wages across cohorts from time effects.

Table 5.1 presents test statistics for the exclusion of cohort age interactions. These are only significant for the lowest education group. Figure 5.1 presents the median life-cycle profiles for the three education groups and for four cohorts spaced 10 years apart each. While there is evidence of steepening life cycle wage profiles for the lowest education group, consistent with the significance of the test statistics quoted in table 5.1, the profiles are basically parallel for the two higher education groups.[25] We have already noted that the educational composition of the lowest educational group changes quite a lot over time towards obtaining more qualifications; this may explain this apparent steepening of the profiles for

[25] Note that the median and mean (not reported) are very similar in all cases.

Table 5.1: Tests of shifting age profiles across cohorts under an alternative identification restriction

	Left School at: At or Before 16	At 17 or 18	Past 18
10th	16.078	4.257	1.789
	0.013	*0.642*	*0.938*
25th	28.673	15.931	10.951
	0.000	*0.014*	*0.090*
50th	28.154	13.427	14.130
	0.000	*0.037*	*0.028*
75th	28.125	9.160	17.399
	0.000	*0.165*	*0.008*
90th	24.703	3.538	16.094
	0.000	*0.739*	*0.014*

Tests distributed $\chi^2(6)$ asymptotically.
P-values in italics. FES data.

younger cohorts in the lowest educational group, while no such evidence appears for the other groups.

Interestingly, Figure 5.2 shows that the predicted rise in the standard deviation of wages from the new version of the model is basically the same as that obtained from a restricted model which excludes cohort age interactions while holding constant the estimated time effects[26]. So what one model interpretes as changes in between group inequality the other attributes to within group changes, leaving the overall predicted increase unchanged.

It is more informative to consider Figure 5.3 where we decompose the overall growth in inequality into within and between group effects and then decompose these between group effects. As with our earlier results about half the overall growth in inequality is attributable to an increase in between group inequality.

[26]see the appendix for details

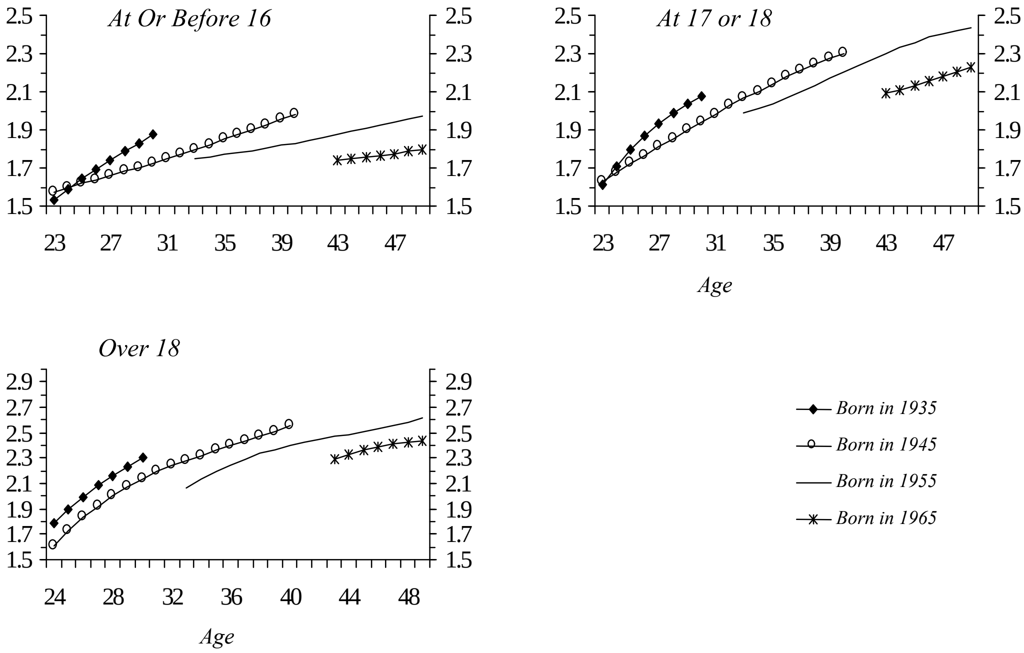

Figure 5.1: Predicted wage growth over the life cycle for 4 cohorts, abstracting from time effects

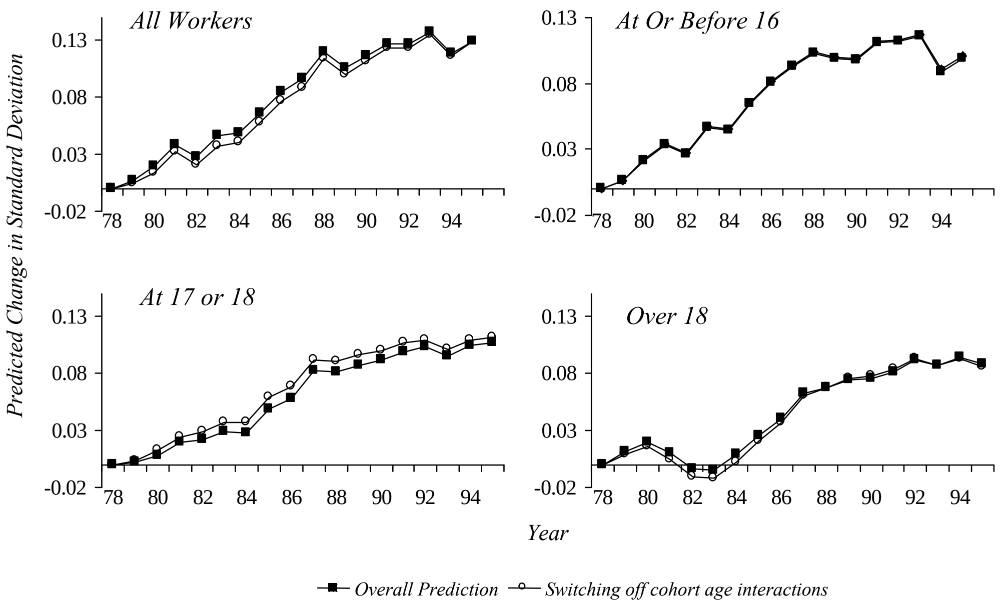

Figure 5.2: The impact of changing life-cycle profiles on changes in wage inequality 1978-1995.

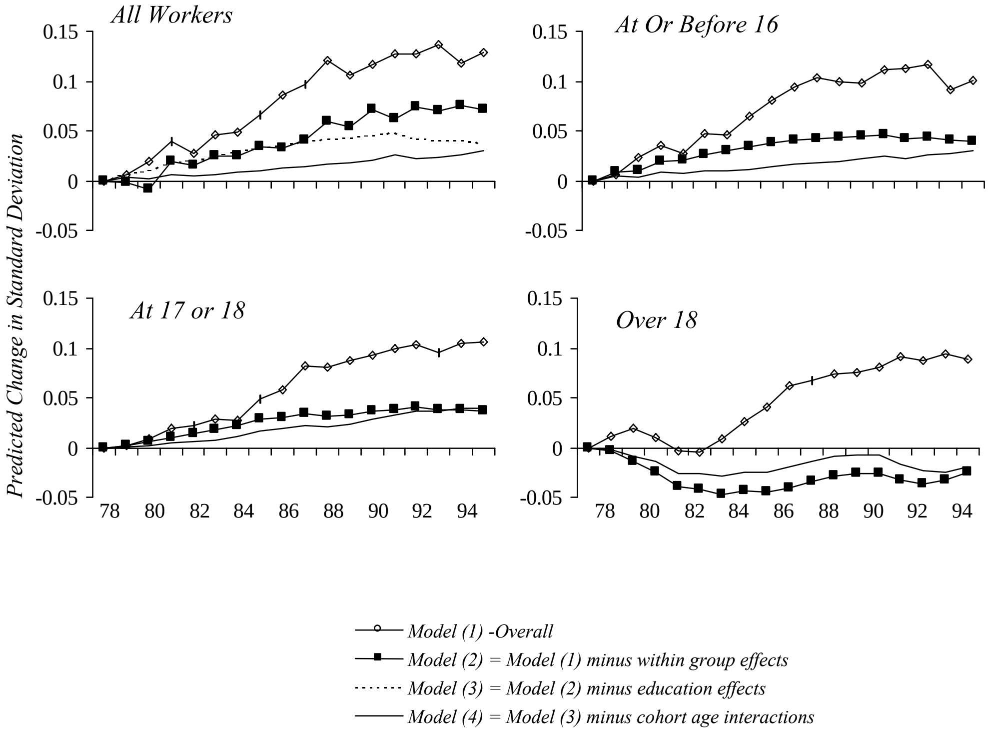

Figure 5.3: Decompositions of the rise in inequality 1978 to 1995.

About half of the latter is attributable to an increase in the returns to education. Switching off the changes in the life-cycle profiles reduces the predicted growth in between group inequality by about 13% in 1995 but the effect is bigger in the mid 1980s.

We next consider within and between group growth in inequality for each education group. The between group is due to differences across cohorts as the cohort composition changes and to differences induced by the overall change in the age composition of the sample. For the lowest two education groups between 30%

and 40% of the total increase in inequality is due to the growth in between group inequality. For the highest education group on the other hand all the increase in inequality is attributed to within groups. The conclusions remain basically unaffected when we ignore cohort-age interactions.

These findings, under both sets of identifying assumptions, are consistent with the idea that the increase in inequality between groups is interpretable as increases in the price of human capital i.e. T_t^{ed} in equation 2.4, in changes in the quantity of human capital reflected in changes in the education and age composition of the sample and by the cohort effects. Shifts in the life cycle profile of wages across cohorts, controlling for time effects would occur if the *type* of human capital changes over the life-cycle (or across cohorts), meaning that the log wages of more experienced workers at any point in time would be differently affected by changes in the relative prices of skills. As this does not appear to have occurred at least for the two higher education groups, age (potential experience) can be interpreted as simply increasing the *quantity* of human capital within each education group. Clearly this conclusion relies on the way we identify cohort effects. However, we have used two different identifying assumptions, that both allow us to test the absence of cohort/age interactions, and obtained similar results.

6. Concluding Remarks

Wage dispersion in the UK has increased rapidly from the late 1970s, with the magnitude of the increase being large in international terms (along with the United States). In this paper we model changes in the UK distribution of wages, treating specific percentiles of the wage distribution as a function of education, cohort, age and cyclical time effects. Most of our analysis is based upon the UK Family Expenditure Survey (FES), which permits us to analyse hourly wage rates on a consistent basis from 1978 onwards. We also report results based upon a data source containing more precise data on educational attainment, the General Household Survey. For both data sources, and for the various modelling approaches we adopt, the restricted and simple model we use to characterise the wage distribution is able to track very closely the changes in the distribution of wages despite its highly parsimonious structure.

About a third of the increase in wage dispersion from the late seventies to mid-nineties is due to increases in educational differentials. Another third is due to a continuous decline in the growth rate of median wages of successive cohorts entering the labour market, an effect which persists over their time in the labour market. The remaining third is within group: successive cohorts enter the labour market with increased dispersion of wages. This happens across all education groups. Using GHS data on earnings we show that some of this increase in within group heterogeneity may have to do with the increased heterogeneity in terms of qualifications of education groups when classified by years of education (as in the FES).

Attributing a portion of the rise in wage inequality to cohort effects results from an important assumption that allows us to separate out such effects from time and age effects. This assumption implies that life-cycle wage profiles do not change over time. However, following an alternative identification strategy, assuming that observed wage growth after 50 is entirely due to aggregate productivity growth, reaches similar conclusions at least for all but the lowest education group. We interpret the different behaviour of the lowest education group to changes in its composition towards obtaining more qualifications in later cohorts.

According to our interpretation the very rapid widening of the UK wage distribution that has occurred in recent history is therefore due to differences across generations in the process of human capital acquisition before (or just after) labour market entry, coupled with increases in the wage returns to education and other productivity enhancing skills (e.g. worker ability or new technologies that are biased in favour of more skilled workers). It will be interesting to see the extent to which such widening continues into the future, despite the rapid expansion of the education system and enrolment in higher education that has been experienced in the UK.

Appendices

A. The estimation procedure

To estimate the conditional quantiles of the wage distribution we use a least absolute deviations estimator (LAD), the asymptotic properties of which are described in Koenker and Basset (1988). For each model we estimate each quantile of the distribution separately to avoid the effect of possible measurement error at the extremes of the distribution on our results. The procedure we used to estimate each of the various models is now described:

A.1. Model 1. Unrestricted model with complete cohort time interactions

The first model is a saturated one, including all possible interactions between year of birth, age in years and education group. The coefficients of this are simply the order statistics of each age, year and education cell. Following Koenker and Basset (1988), the variances of these saturated estimates are defined for each cell as follows:

$$Var(\hat{\beta}^q) = \frac{q(1-q)}{Nf_q^2}$$

where q is the quantile to be estimated, N is the number of observations in the cell and f_q is the density of wages *in each cell* at the q^{th} quantile. These densities were estimated non parametrically using a Gaussian kernel with a fixed bandwidth of half the standard deviation of wages in each cell i.e.

$$f_q = \frac{1}{Nh}\sum_{i=i}^{N}\phi(\frac{w_i - \hat{\beta}_q}{h})$$

where $\phi(.)$ is the density function of the standard normal.

A.2. Model 2. Restricted Model with some cohort age interactions

In the second model the quantiles are specified as

$$y_i^q = \begin{array}{l} c^q + \alpha_1^q age_i + \alpha_2^q age_i^2 + \alpha_3^q age_i^3 + \alpha_4^q age_i^4 + \alpha_5^q age_i^5 + \\ \rho_1^q cohort_i + \rho_2^q cohort_i^2 + \rho_3^q cohort_i^3 + \rho_4^q cohort_i^4 + \rho_2^q cohort_i^5 + \\ \beta_1^q(cohort_i^2 \times age_i) + \beta_2^q(cohort_i^2 \times age_i^2) + \beta_3^q(cohort_i^2 \times age_i^3) + \\ \beta_4^q(cohort_i^3 \times age_i) + \beta_5^q(cohort_i^3 \times age_i^2) + \beta_6^q(cohort_i^3 \times age_i^3) + \\ (time \quad dummies)^q \end{array} \tag{A.1}$$

To estimate model A.1 we use the smoothed LAD estimator suggested by Horowitz (1998): Define the indicator function $I(a)$ which is one when a is true and zero otherwise. The LAD objective function

$$L^q = \sum_{i=1}^{N}\left(q + I(y_i - x'\beta^q > 0) - 1\right)(y_i - x'\beta^q) \tag{A.2}$$

is replaced with one that is differentiable, i.e.

$$SL^q = \sum_{i=1}^{N}\left(q + K(z) - 1\right)(y_i - x'\beta^q) \tag{A.3}$$

where:

$$K(z) = [I(z \geq 1)] + \overset{\Box}{} I(-1 \,\Box\, z \,\Box\, 1) \times (0.5 + \frac{105}{64}\left(z - \frac{5}{3}z^3 + \frac{7}{5}z^5 - \frac{3}{7}z^7\right)]$$

and

$$z = \frac{y - x'\beta^q}{h}$$

The integrated kernel function in equation (A.3) has the properties that

$$\lim_{h \to 0} k\left(\frac{y - x'\beta^q}{h}\right) = I(y - x\beta^q > 0)$$

Thus if the bandwidth (h) is allowed to decrease with sample size N then $\hat{\beta}$ in equation A.3 is asymtopically normally distributed as $N(\beta^q, V^q)$ where β^q is the true coefficient vector and V^q the variance-covariance matrix. As Horowitz shows, a consistent estimate of V is given by

$$V^{SLAD}(\beta^q) = (X'PX)^{-1}(X'WX)(X'PX)^{-1}$$

where P is a $N \times N$ block diagonal matrix with elements:

$$P[i,i] = \frac{1}{h}\frac{\partial K(z_i)}{\partial z_i} = \frac{1}{h}\left(I(-1 \Box z \Box 1)\frac{105}{64}\left(1 - 5z^2 + 7z^4 - 3z^6\right)\right)$$

and W is a block diagonal matrix with elements:

$$W[i,i] = \left(\frac{\partial L^q}{\partial(y_i - x_i'\beta)}\right)^2 = q + K(z_i) - 1 + \frac{\partial K(z_i)}{\partial z_i}(y_i - x_i\beta^q)$$

when $h \to 0$ as $N \to \infty$ the covariance matrix is asymptotically equivalent to that of the usual LAD estimator i.e.

$$V^{LAD}(\beta^q) = q(1-q)(X'\Lambda X)^{-1}E(X'X)(X'\Lambda X)^{-1}$$

where Λ is a diagonal matrix with the density of the residual at zero conditional on x_i (see Bushinksy (1996)). Similar ideas are used to compute the covariance matrix across quantiles which is employed when testing cross quantile restrictions.

Our choice of bandwidth was one third of the sample standard deviation for the higher education groups and a tenth of the standard deviation for those with basic education, where the sample size is much larger.

A.3. Model 3. Restricted model with no cohort age interactions and orthogonal time effects.

The quantiles from the restricted model are modelled as:

$$y_i^q = \begin{array}{l} c^q + \alpha_1^q age_i + \alpha_2^q age_i^2 + \alpha_3^q age_i^3 + \\ \rho_1^q cohort_i + \rho_2^q cohort_i^2 + \rho_3^q cohort_i^3 + \\ (orthogonal \;\; time \;\; dummies)^q \end{array} \tag{A.4}$$

To fit model A.4 we apply minimum distance to the estimated parameters from models 1 and 2[27]. To construct the orthogonal time effects we use a stepwise procedure.

We start by obtaining preliminary consistent estimates of α^q and ρ^q in A.4 by regressing the cell quantiles on age, age squared, age cubed, cohort, cohort squared and cohort cubed. Next we take the predictions from these model away from the original cell quantiles to obtain a first stage estimate of

$$\varepsilon_1^q = y^q - E(y^q|age \text{ and } cohort)$$

where E denotes an expectation. We then regress these residual on the set of 17 year dummy variables. The predictions $\hat{E}(\varepsilon_1^q|\text{time effects})$ are the first round estimate of the orthogonal time effects. This will allow one to obtain an estimate of

$$\varepsilon_2^q = y^q - E(\varepsilon_1^q|\text{time})$$

We then take these new residual and regress them using GLS against the age and cohort functions as above now weighted by the estimated variances of the cell quantiles adjusted for the estimated variance of the time effects. This will give a our consistent final estimate of the α^qs and ρ^qs and

$$\varepsilon_3^q = y^q - E(y^q|age \text{ and } cohort)$$

[27] see Ferguson, 1958 or Rothenberg, 1971 in general and Chamberlain(1993) in the context of quantiles

Regressing these against the set of year dummy variables, weighted by the estimated variances of the cell quantiles and the variances of the estimates in step 3 will give our final estimates of the time effects. By construction, these will sum to zero and be orthogonal to the included age and cohort effects.

In the second method where we employ minimum distance on Model 2, the procedure is more complicated.

The procedure adopted is as follows:

Let

Q^1 be a $N \times K^1$matrix of the age and cohort variables where N is the number of observations in each education group

Q^2 be a $N \times K^2$matrix of the time dummies for the raw data

Z be a $N \times K$ matrix of all the variables in equation A.1, where $K > K^1 + K^2$ since K includes the cohort age interactions.

$\hat{y}^q$ be a $N \times 1$ matrix of the estimated predictions from Model 2

$(V^q)^{-1}$ be the inverse of the variance covariance matrix of Model 2

$\hat{\alpha}_1^q$ and $\hat{\rho}_1^q$ be consistent 1st step estimates of the age and cohort effects for quantile q in A.4

$\hat{t}_1^q$ be the 1st step estimates of the orthogonal time effects for quantile q in A.4

$\hat{\alpha}_2^q$ and $\hat{\rho}_2^q$ be the final estimates of the age and cohort effects for quantile q in A.4

$\hat{t}_2^q$ be the fina1 estimates of the orthogonal time effects for quantile q in A.4

A first step estimate of the age and cohort effects is:

$$\begin{bmatrix} \hat{\alpha}_1^q \\ \hat{\rho}_1^q \end{bmatrix} = (Q_1'Q_1)^{-1}Q_1'\hat{y}^q$$

After computing:

$$\hat{\varepsilon}_1^q = \hat{y}^q - Q_1 \begin{bmatrix} \hat{\alpha}_1^q \\ \hat{\rho}_1^q \end{bmatrix}$$

a first step estimate of the time effects can then by obtained as:

$$\hat{t}_2^q = (Q_2' Q_2)^{-1} Q_2' \hat{\varepsilon}_1^q$$

After computing

$$\hat{\varepsilon}_2^q = \hat{y} - Q_2 \hat{t}_2^q$$

The final estimate of the age and cohort effects is given by:

$$\begin{bmatrix} \hat{\alpha}_2^q \\ \hat{\rho}_2^q \end{bmatrix} = \left[(Q_1' Z)((Z'Z)^{-1}(V^q)^{-1}(Z'Z)^{-1})(Z'Q_1) \right]^{-1} (Q_1' Z)((Z'Z)^{-1}(V^q)^{-1}(Z'Z)^{-1})(Z'\hat{\varepsilon}_2^q)$$

And finally with:

$$\hat{\varepsilon}_3^q = \hat{y} - Q_1 \begin{bmatrix} \hat{\alpha}_2^q \\ \hat{\rho}_2^q \end{bmatrix}$$

the final estimate of the time effects can be obtained

$$\hat{t}_2^q = \left[(Q_2' Z)((Z'Z)^{-1}(V^q)^{-1}(Z'Z)^{-1})(Z'Q_2) \right]^{-1} (Q_2' Z)((Z'Z)^{-1}(V^q)^{-1}(Z'Z)^{-1})(Z'\hat{\varepsilon}_3^q)$$

A.4. Model 4: identification of the time effects from the wage growth of older workers

In section 5 of the paper, time effects are identified by the assumption that human capital accumulation is assumed to stop over the age of 49. This means that wage growth within cohorts past the age of 49 is solely a time effect. The following general specification is estimated

$$
y_i^q = \begin{array}{l} c^q + \alpha_1^q e_i + \alpha_2^q e_i^2 + \alpha_3^q e_i^3 + \\ \rho_1^q cohort_i + \rho_2^q cohort_i^2 + \rho_3^q cohort_i^3 + \\ \beta_1^q (cohort_i \times e_i) + \beta_2^q (cohort_i^2 \times e_i^2) + \beta_3^q (cohort_i^2 \times e_i^3) + \\ \beta_4^q (cohort_i^2 \times e_i) + \beta_5^q (cohort_i^3 \times e_i^2) + \beta_6^q (cohort \times e_i^3) \\ +(time \;\; dummies)^q \end{array}
$$

where e is defined as age for those under 50 and 50 otherwise. The procedure used to estimate this was otherwise exactly the same as in Model 2. The cohort-age interactions were switched off by obtaining predictions from the following model:

$$
y_i^q - estimated\ time\ effects = \begin{array}{l} c^q + \alpha_1^q e_i + \alpha_2^q e_i^2 + \alpha_3^q e_i^3 + \\ \rho_1^q cohort_i + \rho_2^q cohort_i^2 + \rho_3^q cohort_i^3 \end{array}
$$

B. Constructing the unconditional distributions and computing counterfactuals

Unconditional quantiles can be estimated from conditional ones in the following way

$$
q = \Pr(w < w^q) = \sum_{z=1}^{Z} \frac{N_z}{N} \Pr(w < w^q \mid z) \tag{B.1}
$$

where z is a year, age education cell, N_z (N) is the number of observations in cell z (the full sample) and w^q is the point corresponding to the qth quantile of the unconditional distribution. Given a set of predicted conditional quantiles $z\beta^q$, we can then estimate by interpolation what conditional quantile a given wage level (ω) would correspond to. This is estimated as:

$$
q_z^{\omega} = \frac{1}{2}\left(\max_q\{q|z\beta^q \;\square\; \omega\} + \min\{q|z\beta^q \geq \omega\right)
$$

where $q_z^{\omega} = \Pr(w < \omega \mid z)$ is the quantile of cell z corresponding to the wage ω. The unconditional quantile corresponding to the wage ω can then be estimated as

$$q^{\,\omega} = \sum_{z=1}^{Z} q_z^{\,\omega} \frac{N_z}{N}$$

where N is the number of observations in the raw data and N_z is the number of observations in cell z. Estimating enough of these $(\omega,\, q)$ pairs will then enable us to estimate the unconditional quantiles of the distribution as:[28]

$$w^q = \frac{1}{2}\left(\max\{\omega | q^{\,\omega} \leq q\} + \min\{\omega | q^{\,\omega} \geq q\}\right)$$

where w^q is the wage corresponding to the q^{th} quantile. We can estimate the mean (M) as

$$M = \sum_{q^{\,\omega}=\min+1}^{\max} (w^q)(q^{\,\omega} - q^{\,\omega}_{-1})$$

where q^w_{-1} denotes the quantile corresponding to the next lowest wage. Similarly the standard deviation (Sd) is estimated as

$$Sd = \sqrt{\sum_{q^w=\min+1}^{\max} (w^q - M)^2 (q^w - q^w_{-1})}$$

We estimate 13 quantiles[29] and tie the distribution down with the observed minimum and the maximum for each year and age group. This gave a good fit of the data.

B.1. Obtaining counterfactual comparisons

The relative size of between versus within group effects on inequality was established by constructing the mean wage in each cell from the predicted quantiles

[28] We estimate the quantiles correponding to 200 wage points evenly spaced across the range of wages in the sample.

[29] i.e. the 5th, 25th, 75th and 95th percentiles and the 9 deciles

and using these predicted means to estimate what the standard deviation of wages would be if there was no within group inequality.

For each cell, the mean (M_z) was estimated as:

$$M_z = \sum_{q_z^w = \min + 1}^{\max} (w_z^q)(q_z^w - q_{z-1}^w)$$

where, as before, q_{z-1}^w denotes the quantile corresponding to the next lowest wage in cell z. The predicted standard deviation of wages with no within group inequality in the sample is then simply

$$\sqrt{\sum_{z=1}^{Z} \frac{N_z}{N} (M_z - M)^2}$$

The relative role of education and cohort effects on changes in between group inequality was estimated by setting the returns to education to zero, that is giving each age, year cell the predicted wages for the lower education groups.

References

[1] Aghion Phillipe Howitt Peter and Violante Gianluca (1999) Technology, Knowledge and Inequality, University College London, mimeo presented at the ELSE/IFS/MacArthur conference on "Inequality and Labor Markets", May 1999.

[2] Baker, George, Michael Gibbs and Bengt Holmstrom (1994) "The wage policy of a firm", *Quarterly Journal of Economics*, 109, 921-955.

[3] Becker, Gary (1975) *Human Capital: A Theoretical and Empirical Analysis,* Chicago: University of Chicago Press.

[4] Beaudry, Paul and John DiNardo (1991) "The effect of implicit contracts on the movement of wages over the business cycle: Evidence from micro-data", *Journal of Political Economy*, 99, 665-88.

[5] Beaudry, Paul and David Green (1996) "Cohort patterns in Canadian earnings and the skill-biased technical change hypothesis", University of British Columbia mimeo.

[6] Bound, John and Johnson, George (1992) "Changes in the structure of wages in the 1980s: An evaluation of alternative explanations", *American Economic Review,* 82, 371-92.

[7] Bushinsky, Moshe (1994) "Changes in the US wage structure 1963 to 1987; an application of quantile regressions", *Econometrica,* 62, 405-458.

[8] Bushinsky, Moshe (1996) "Recent advances in quantile regression models: A practical guideline for empirical research", Yale University mimeo.

[9] Chamberlain, Gary (1993) "Quantile regressions, censoring and the structure of wages" in Christopher Sims (ed.) *Advances in Econometrics: Proceedings of the 6th World Congress in Barcelona Volume 1,* Cambridge: Cambridge University Press..

[10] Dearden, Lorraine, Javier Ferri and Costas Meghir (1998) "The effect of school quality on educational attainment and wages", Institute for Fiscal Studies Discussion paper 98/3.

[11] Dearden, Lorraine, Stephen Machin and Howard Reed (1997) "Intergenerational mobility in Britain", *Economic Journal*, 107, 47-66.

[12] Deaton, Angus and Christina Paxson (1993) "Intertemporal choice and inequality", National Bureau of Economic Research Working Paper 4328.

[13] DiNardo, John, Nicole Fortin and Thomas Lemieux (1996) "Labour market institutions and the distribution of wages, 1977-1992", *Econometrica, 64, 1001-44.*

[14] Farber, Henry and Robert Gibbons (1997) "Learning and wage dynamics", *Quarterly Journal of Economics,* 112, 1007-47.

[15] Ferguson, (1958) "A Method for Generating Best Asymptotically Normal Estimates with Application to the Estimation of Bacterial Densities," *Annals of Mathematical Statistics,* 29*,* 1046-1061.

[16] Fitzenberger, Berndt, Reinhard Hujer, Thomas MaCurdy and Reinhold Schnabel (1995) "The dynamic structure of wages in Germany, 1976-1984; A Cohort Analysis", University of Konstanz Discussion Paper 22.

[17] Freeman, Richard and Lawrence Katz (1994) "Rising wage inequality: The United States versus other advanced countries", in Richard Freeman (ed.) *Working Under Different Rules*, New York: Russell Sage Foundation.

[18] Ginther, Donna (1994) "A non-parametric analysis of the US earnings distribution", University of Wisconsin mimeo.

[19] Gosling, Amanda and Stephen Machin (1995) "Trade unions and the dispersion of earnings in British establishments, 1980-90", *Oxford Bulletin of Economics and Statistics, 57, 167-84.*

[20] Gosling, Amanda, Stephen Machin and Costas Meghir (1994) "What has happened to men's wages since the mid-1960s?", *Fiscal Studies*, 15, 63-87.

[21] Horowitz J L (1998) "Bootstrap models for median regression models", *Econometrica*, Vol.66, No.6, pp.1327-1351

[22] Jovanovic, Boyan (1979) "Job matching and the theory of turnover", *Journal of Political Economy 87,* 972-990.

[23] Juhn, Chinhui, Kevin Murphy and Brooks Pierce (1993) "Wage inequality and the rise in returns to skill", *Journal of Political Economy, 101, 410-42.*

[24] Katz, Lawrence, Gary Loveman and David Blanchflower (1995) "An international comparison of changes in the structure of wages: France, the United Kingdom and the United States", in Richard Freeman and Lawrence Katz (eds.) *Changes and Differences in Wage Structures,* Chicago: University of Chicago Press.

[25] Katz, Lawrence and Kevin Murphy (1992) "Changes in the structure of wages 1963-87: Supply and demand factors", *Quarterly Journal of Economics,* 107, 35-78.

[26] Koenker, Roger and Gilbert Bassett Jr. (1978) "Regression Quantiles", *Econometrica,* 46, 33-50.

[27] Machin, Stephen (1996) "Wage inequality in the UK", *Oxford Review of Economic Policy*, 7, 47-62.

[28] MaCurdy, Thomas and Thomas Mroz (1991) "Estimating macroeconomic effects on wages from cohort specifications", University of Stanford mimeo.

[29] Murphy, Kevin and Finis Welch (1992) "The structure of wages", *Quarterly Journal of Economics,* 107, 285-326.

[30] Neal, Derek (1995) "Industry-specific human capital: Evidence from displaced workers", *Journal of Labor Economics*, 13, 653-77.

[31] Rothenberg, Thomas (1971) *Efficient Estimation With A Priori Information,* Cowles Foundation Monograph 23, New Haven: Yale University Press.

[32] Schmitt, John (1995) "The changing structure of male earnings in Britain, 1974-1988", in Richard Freeman and Lawrence Katz (eds.) *Changes and Differences in Wage Structures,* Chicago: University of Chicago Press.