

IZA DP No. 1350

## Changes in the Distribution of Male and Female Wages Accounting for Employment Composition Using Bounds

Richard Blundell  
Amanda Gosling  
Hidehiko Ichimura  
Costas Meghir

October 2004

# Changes in the Distribution of Male and Female Wages Accounting for Employment Composition Using Bounds

**Richard Blundell**

*University College London,  
IFS and IZA Bonn*

**Amanda Gosling**

*University of Essex, IFS and CEPR*

**Hidehiko Ichimura**

*University College London and IFS*

**Costas Meghir**

*University College London,  
IFS, CEPR and IZA Bonn*

Discussion Paper No. 1350

October 2004

IZA

P.O. Box 7240  
53072 Bonn  
Germany

Phone: +49-228-3894-0

Fax: +49-228-3894-180

Email: [iza@iza.org](mailto:iza@iza.org)

Any opinions expressed here are those of the author(s) and not those of the institute. Research disseminated by IZA may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit company supported by Deutsche Post World Net. The center is associated with the University of Bonn and offers a stimulating research environment through its research networks, research support, and visitors and doctoral programs. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

## ABSTRACT

### **Changes in the Distribution of Male and Female Wages Accounting for Employment Composition Using Bounds\***

This paper examines changes in the distribution of wages using bounds to allow for the impact of non-random selection into work. We show that bounds constructed without any economic or statistical assumptions can be informative. However, since employment rates in the UK are often low they are not informative about changes in educational or gender wage differentials. Thus we explore ways to tighten these bounds using restrictions motivated from economic theory. With these assumptions we find convincing evidence of an increase in inequality within education groups, changes in the “return” to education and increases in the relative wages of women.

JEL Classification: J31, C24

Keywords: wage differentials, selection models, bounds

Corresponding author:

Costas Meghir  
Department of Economics  
University College London  
Gower Street  
London WC1E 6BT  
United Kingdom  
Email: [c.meghir@ucl.ac.uk](mailto:c.meghir@ucl.ac.uk)

---

\* We would like to thank Jaap Abbring, Daron Acemoglu, Manuel Arellano, James Heckman, Joel Horowitz, Chuck Manski, Whitney Newey for useful comments and discussion as well as participants in numerous seminars including the Econometric Society European meeting in Madrid 2004, the meeting of the European Association of Labour Economists in Lisbon 2004, the Stanford Institute for Theoretical Economics 2002 and seminars at UCL, Johns Hopkins and Northwestern Universities and the Econometric Study Group in London 2004. Funding for this research was provided by the ESRC centre for the analysis of Fiscal Policy at the IFS. Data from the FES made available from the CSO through the ESRC data archive has been used by permission of the Controller of HMSO. We are responsible for all errors and interpretations.

## Executive Summary

Since the late 1970's, wage inequality among those working rose in the UK has risen dramatically. However, this has been accompanied by large changes in employment rates. Male employment has declined by at least 10 percentage points, this decline being even more dramatic for those with lower education levels. On average female employment has risen but this increase is concentrated amongst more educated women, particularly those with children. This means the composition of the workforce has changed significantly in terms of the observable skill characteristics of workers. It is very likely that such changes are also present in the dimensions of skill that researchers do not observe. Factors such as motivation and aspects of ability not picked up by observed education measures may also be important determinants of productivity and hence wages. It is thus possible that the observed rise in inequality and changes in differentials across education and experience levels and between men and women could simply be driven by changes in the kinds of people that are observed working. The focus of our paper is to address this very important issue.

Put simply the problem we face is that we are interested in the overall distribution of employment rewards, but we only observe the pay of those actually taking up work. These are two very different things. We use and develop further a method that allows us to define ranges in which the changes of interest fall. Our approach allows us to draw conclusions about the changing structure of the labour market without making any assumptions about the determination of employment or the distribution of wages faced by those not working. We find, however, that some minimal economic restrictions are necessary to draw conclusions about changes to educational differentials and the relative growth rates of male and female wages. The advantage of the approach we use is not that it is assumption free but rather that we need only impose plausible restrictions with strong economic content rather than those based on statistical or functional form assumptions.

Using 20 years of the Family Expenditure Survey, we provide upper and lower bounds to: The level and changes to inequality, the wage growth over age and across cohorts and levels and changes in differentials associated with gender and education. Some key conclusions are that wages of the upper quartile of the wage distribution grew by at least 25 percentage points more than the lower quartile, the returns to college relative to high school grew by at least 8 percentage points for the young and the wages for young unskilled females grew at least 24 percentage points faster than for a similar group of men.

# 1 Introduction and motivation

There is a vast literature on the distribution of wages and the way this has evolved over time. This includes work on the returns to education, on gender wage differentials and on the evolution of wages over the life-cycle all of which relate to factors that can lead to wage dispersion. In addition, economies such as the US and the UK have seen large and unprecedented increases in wage inequality amongst workers over the last 30 to 40 years. This is illustrated in Figure 1 where we show the way that the interquartile ranges of male and female log hourly wages have evolved for those who work in the UK.

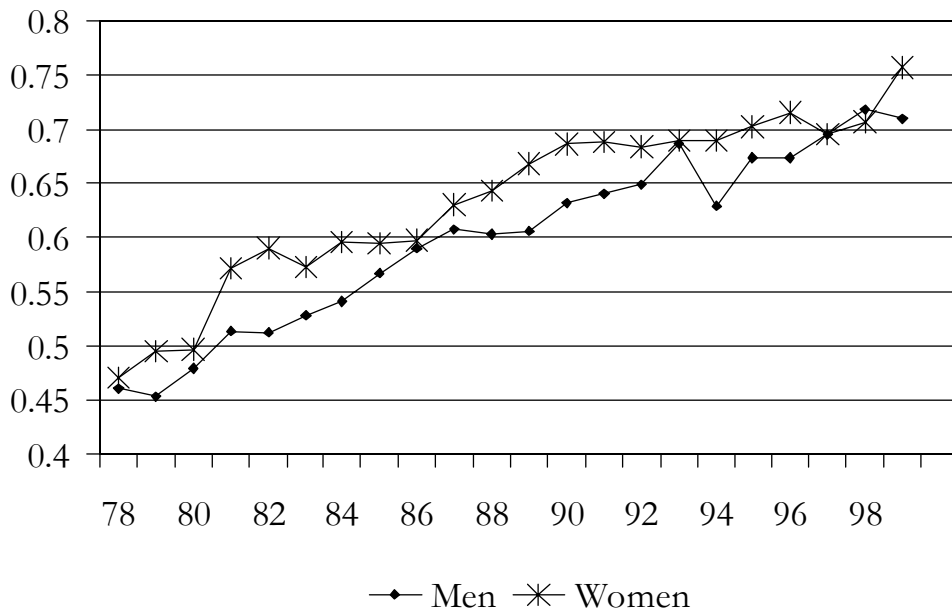


Figure 1: Interquartile range of log wages

These increases in inequality have been associated with increased returns to education, cohort effects and increases in the returns to unobserved skill.<sup>1</sup> A variety of interpretations have been

<sup>1</sup>See Juhn, Murphy and Pierce (1993) for the US and Gosling, Machin and Meghir (2000) for the UK.

given as to why these events have occurred; these include skill biased technical change, globalization induced increased competition for low skill workers and changes in the supply of graduates. Gosling, Machin and Meghir (2000) show that the increases in the UK can be attributed to permanent differences across cohorts and in changes in the returns to education. At the same time the gap between the wages of working men and working women have fallen.<sup>2</sup>

In parallel with these momentous changes in the distribution of observed wages, employment rates for males and females as well as the composition of the workforce have changed. Employment for men aged 22 to 59 has decreased from 93% in 1978 to 83% in 2000. This decline is not confined to older men and so reflects more than the increase in early retirement. In Figure 2 we show the age profile for employment for 1978 and for 2000 by gender: The decline for males occurs at all ages, although it is even more pronounced for men over 50. Women saw an increase in their employment, especially for those below 30. Figure 3 illustrates that the change in employment has been heavily skill biased. The least skilled are those who stopped full time education at or before 16 (statutory schooling). In this paper those who completed education sometime between 17 and 18 are called high school graduates and those who completed past 18 are called the college graduates. Although male employment has declined for all education groups the highest decline has been for the least skilled group. For women most of the increase in employment can be accounted for by the increase in employment of women with more than the minimal level of education. The employment rate of the statutory schooling group has shown a slight decline over the entire time period.

---

<sup>2</sup>See Harkness (1996) for the UK and Blau (1998) and Blau and Kahn (1997) for the US.

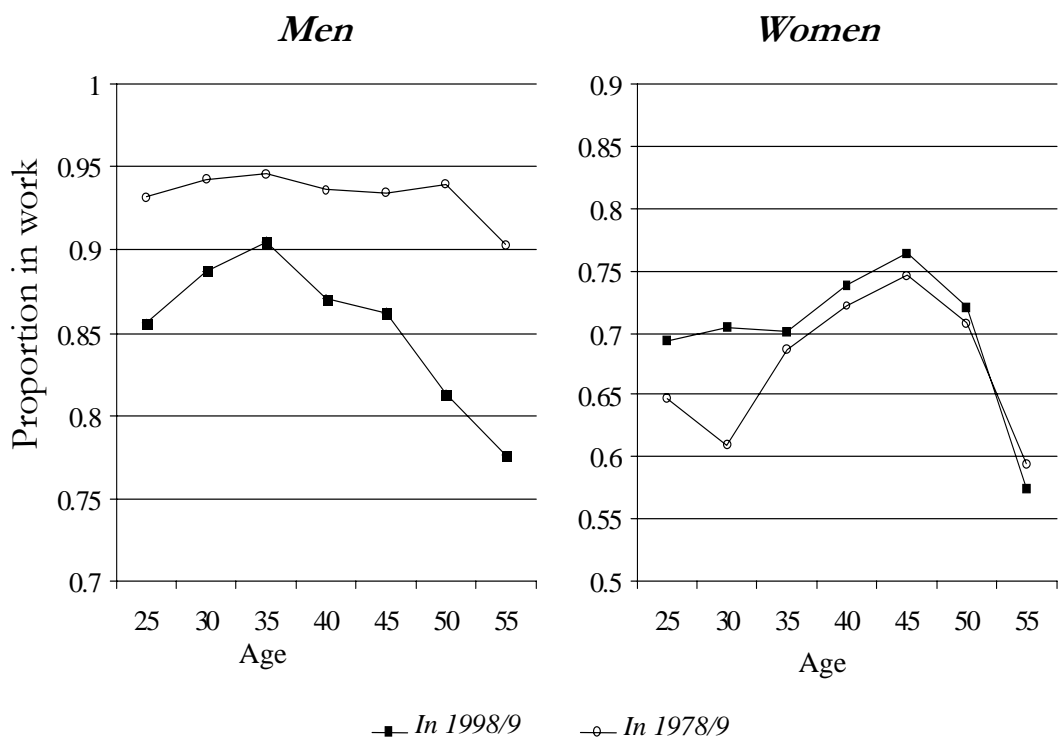


Figure 2: Employment by age in 1978 and 1998

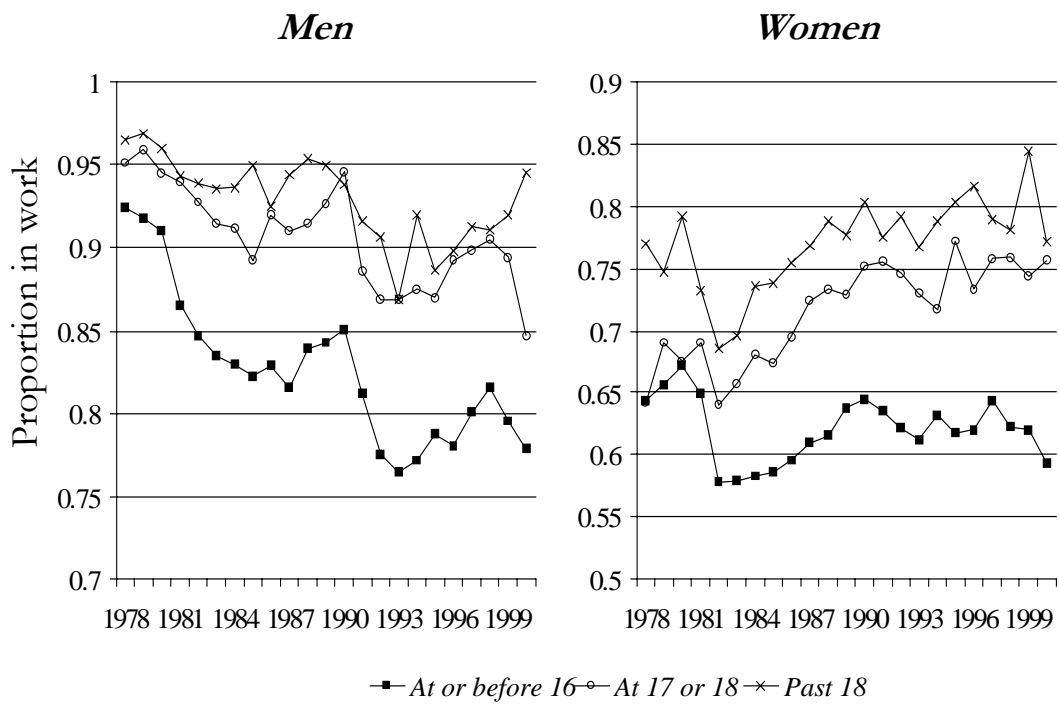


Figure 3: Employment by education level for men and women over time



Just as observable characteristics differ between the workers and non-workers, those not working may have different unobserved skill characteristics to those working. Any change in employment may thus affect the distribution of wages directly through composition effects. Wage inequality amongst workers could increase if employment rates have fallen less for those at the top and the bottom of the skill distribution without any change in the underlying structure of demand or distribution of skill. This point has a bearing on the interpretation we give these observed changes in the distribution of wages. For example, in order to interpret the change in return to education as being driven by an increase in the relative demand for higher educated workers (due say to skill biased technical change) we need to establish that the change in the observed return is not an artifact of changes in the composition of employment.

It is often believed that individuals with lower offered wages are the ones less likely to work. However, in many circumstances a proportion of non-workers may be drawn from the middle or top of the distribution of unobserved skill. This may occur due to wealth effects or possibly unobservable preference characteristics that can affect reservation wages. In addition, when examining changes in the distribution of wages as large as the ones observed in the US and the UK one has to consider the possibility that the selection mechanisms into work may have changed substantially.

The aim of this paper is to analyze wage dispersion and its evolution, allowing for the sample selection induced by individuals' non-employment. To address the issue of selection without relying on strong assumptions we use bounds to the wage distribution and its quantiles. We start from the worst case bounds. We then employ a number of restrictions motivated by economics to tighten the bounds. Our theoretical work is related to a number of earlier papers. These include Manski (1994) on worst case bounds, Manski and Pepper (2000) who show how exclusion and monotonicity restrictions can be used to tighten the bounds; Koenker and Bassett (1978), Buchinsky (1994, 1998) and Gosling, Machin and Meghir (2000) on quantiles and their use in characterizing changes in the wage distribution; Heckman Smith and Clements (1997) and Heckman LaLonde and Smith (1998) who investigate bounds for the joint distribution when only the marginals are identified,

and Heckman and Vytlacil (2001, 2004) who consider the use of instruments combined with index restrictions. Our approach allows us to obtain very tight bounds on parameters of interest and in some cases close to point estimates, while preserving the robustness of a non-parametric approach and imposing only relatively weak and transparent restrictions.

The empirical analysis focuses on inequality, on wage differentials related to education, gender and age groups and on the way these have changed over time. We seek to bound relevant aspects of the wage distribution accounting for the presence of non-employment in the labour market. As a data source we use the UK Family Expenditure Survey from 1978 to 2000.

It is important to stress at the outset that the analysis in this paper is not an attempt to provide a structural model of wage determination which could provide answers to counterfactual questions relating say to the impact of some policy intervention on the distribution of wages. We are simply trying to estimate the distribution of offered wages accounting for sample selection induced by non-work.

At least for analyzing means, the selection problem is well understood in the literature and it is usually dealt with either in a parametric<sup>3</sup> or in a nonparametric framework.<sup>4</sup> A recent example which shows how important such selection issues can be is the paper by Blundell, Reed and Stoker (2003). They use changes in the policy framework for out of work welfare benefits, exploiting the differences generated by the variability in housing costs across individuals to sort out composition effects from genuine wage growth. They find that the picture of wage growth is very different after accounting for selection effects.

Some of our substantive findings are as follows: Wages of the upper quartile of the wage distribution grew by at least 25 percentage points more than the lower quartile, the returns to college relative to high school grew by at least 8 percentage points for the young and the wages for young unskilled females grew at least 24 percentage points faster than for a similar group of men.

---

<sup>3</sup>See Heckman (1979) or Heckman and Sedlacek (1985).

<sup>4</sup>See for example, Powell (1987), Newey (1988), Heckman (1990), Choi (1990), Ichimura and Lee (1991), Coslett (1991), Ahn and Powell (1993).

We also show that for young unskilled women the improvement of wages relative to similar men is underestimated due to selection bias, in the sense that the (restricted) lower bound to the change in the male-female differential suggests a larger improvement for women than the one based on wages of working men and women.

In what follows we first show how bounds to the distribution of wages and its quantiles can be derived with restrictions. We then discuss estimation of these bounds, as well as relevant specification tests. Finally, we discuss the data and our empirical results.

## 2 Worst case bounds and bounds with restrictions

Let  $W$  and  $X$  denote the dependent random variable and the conditioning vector respectively. In our case the dependent variable  $W$  should be taken to be the log wage and  $X$  should be understood to include gender, age, education and year. When  $W$  is observed, the indicator variable  $E$  equals 1 and when  $W$  is not observed,  $E$  equals 0. In our case  $E$  indicates whether the person is employed or not. The probability of  $E = 1$  given  $X = x$  is written as  $P(x)$ . In our analysis this is the employment probability for individuals with characteristics  $x$ . We write the cumulative distribution function (CDF) of  $W$  given  $X = x$  by  $F(w|x)$ , that given  $X = x$  and  $E = 1$  by  $F(w|x, E = 1)$ , and that given  $X = x$  and  $E = 0$  by  $F(w|x, E = 0)$ . While  $F(w|x)$ , the object of interest, is not identified (because of non-random sample selection), we can write

$$F(w|x) = F(w|x, E = 1)P(x) + F(w|x, E = 0)[1 - P(x)]. \quad (1)$$

Given that the data identify both  $F(w|x, E = 1)$  and  $P(x)$  the problem can be respecified as one in which only  $F(w|x, E = 0)$  is unknown. In our application this should be understood as the distribution of wages rejected by those not taking up employment and  $F(w|x)$  is an equilibrium distribution at a point in time which is not invariant to changes in  $P(x)$ .

Our starting point for the analysis is the work by Manski (1994) who notes that once the

inequality:

$$0 \leq F(w|x, E = 0) \leq 1$$

is substituted into equation (1), the bounds to the cumulative distribution function become:

$$F(w|x, E = 1)P(x) \leq F(w|x) \leq F(w|x, E = 1)P(x) + [1 - P(x)]. \quad (2)$$

The bounds can then be translated to give the worst case bounds on the conditional quantiles.<sup>5</sup>

Denoting by  $w^q(x)$  the  $q$ th quantile of  $F(w|x)$ , we have

$$w^{q(l)}(x) \leq w^q(x) \leq w^{q(u)}(x)$$

where  $w^{q(l)}(x)$  is the lower bound and  $w^{q(u)}(x)$  is the upper bound which solve the equations

$$q = F(w|x, E = 1)P(x) + [1 - P(x)] \quad \text{and} \quad q = F(w|x, E = 1)P(x), \quad (3)$$

with respect to  $w$ , respectively.<sup>6</sup> Unless there are restrictions on the support of  $W$ , we can only identify the lower bound to quantiles  $q \geq 1 - P(x)$  and upper bounds for quantiles  $q \leq P(x)$ .

In general we can not identify bounds to means and variances or higher order moments unless we impose further restrictions. We thus focus on quantiles to characterize the bounds to the distribution.

## 2.1 Imposing restrictions to tighten the bounds

A higher  $P(x)$  implies tighter bounds on quantiles. With employment rates in our data, the worst case bounds can be informative about certain aspects of wages, such as life-cycle wage growth for men. However, in many cases they are uninformative. We seek to tighten the bounds on quantiles by imposing restrictions motivated by theoretical considerations about the selection process; in our case the employment process. Some of these restrictions will have testable implications.

---

<sup>5</sup>Proposition 3, p.152.

<sup>6</sup>The approach of Brown (1984) is quite similar to this but it is less general. Essentially Brown compares the evolution of the observed median with the worst case lower bound.

### 2.1.1 The median restriction

In a labour supply model with homogeneous reservation wages (given  $X$ ), individuals with higher offered wages are more likely to work. Thus the first set of economic restrictions we could make would be those that restricted the lower bound to the distribution of wages in some way. For the worst case bounds  $F(w|x, E = 0)$  is set as lying between 0 and 1. Let  $W$  denote the random variable representing potential log-wage. When  $0 < P(x) < 1$ , assuming the absence of negative selection into employment, i.e.

$$\Pr(E = 1|W \leq w, x) \leq \Pr(E = 1|W > w, x)$$

for each  $w$  is equivalent to assuming

$$F(w|x, E = 1) \leq F(w|x, E = 0) \quad \forall w, x \quad (4)$$

for each  $w$  with  $0 < F(w|x) < 1$ . This follows from the equality

$$\begin{aligned} & F(w|x, E = 1) - F(w|x, E = 0) \\ &= \frac{F(w|x)(1 - F(w|x))(\Pr(E = 1|W \leq w, x) - \Pr(E = 1|W > w, x))}{P(x)(1 - P(x))} \end{aligned}$$

which is a consequence of the Bayes rule.

The likely positive relationship between offered and reservation wage does not constitute a rejection of this restriction; this would only occur if the relationship between reservation wages and offered wages is strong enough to undermine the direct relationship of wages and employment. We are concerned that for some groups, however, this may be the case. Thus we impose a weaker restriction that for individuals with observed characteristic  $x$ , the median wage offer for those not working is not higher than the median observed wage,  $w^{50(E=1)}(x)$ . This implies that

$$\begin{aligned} 0 \leq F(w|x, E = 0) \leq 1 & \quad \text{if } w < w^{50(E=1)}(x) \\ 0.5 \leq F(w|x, E = 0) \leq 1 & \quad \text{if } w \geq w^{50(E=1)}(x) \end{aligned} \quad (5)$$

The restriction implied by equation (5) sets the bounds as

$$\begin{aligned}
& F(w|x, E = 1)P(x) \\
& \leq F(w|x) \leq && \text{if } w < w^{50(E=1)}(x) \\
& F(w|x, E = 1)P(x) + (1 - P(x)) \\
& \\
& F(w|x, E = 1)P(x) + 0.5(1 - P(x)) \\
& \leq F(w|x) \leq && \text{if } w \geq w^{50(E=1)}(x) \\
& F(w|x, E = 1)P(x) + (1 - P(x)).
\end{aligned} \tag{6}$$

As equation (6) shows, this restriction provides tighter bounds to every quantile at the median and above.

The restriction we impose is not incontrovertible. There are reasons to believe that reservation wages may well be positively related with offered wages for some groups of people. First, if wages are correlated over time, higher wage people are more likely to have accumulated assets which may increase their reservation wages. This may be an issue for older workers, for example. Second if high wage women are matched with high wage men, then out of work income of women could be increasing in their potential earnings and thus associated with higher reservation wages. Lastly as higher wage women tend to delay rather than avoid childbirth, those not working could be the higher wage women in older age groups who have pre-school children. We present some corroborative evidence for the assumption. Moreover, we still present worst case bounds.

### 2.1.2 Using determinants of employment to tighten the bounds

**An exclusion restriction** Manski (1994) shows that if  $W$  is independent of  $Z$  conditional on  $X$  i.e.

$$F(w|x, z) = F(w|x) \quad \forall w, x, z$$

then the bounds to the conditional distribution of  $W$  given  $X$  is given by

$$\max_z F(w|E = 1, x, z)P(x, z) \leq F(w|x) \leq \min_z F(w|E = 1, x, z)P(x, z) + 1 - P(x, z). \tag{7}$$

This formula can easily be modified to combine the exclusion restriction with the median restriction, by replacing  $F(w|E = 1, x, z)P(x, z)$  by the lower bound given in equation (6).

In general since  $F(w|E = 1, x, z)$  depends on  $z$ , finding the minima and the maxima of  $P(x, z)$  over  $z$  will not identify the tightest bounds.

Note that we can rewrite the lower bound  $F(w|E = 1, x, z) \Pr(x, z)$  as

$$F(w|x) \Pr(E = 1|W \leq w, x, z)$$

when  $W$  and  $Z$  are independent given  $X$ . This implies that the lower bound is maximized at  $z$  that maximizes  $\Pr(E = 1|W \leq w, x, z)$ . Analogously, the upper bound is minimized at  $z$  that *maximizes*  $\Pr(E = 1|W > w, x, z)$ . Thus in order for the lower bound to be tightened at  $w$ ,  $\Pr(E = 1|W \leq w, x, z)$  needs to vary over  $z$  and for the upper bound to be tightened at  $w$ ,  $\Pr(E = 1|W > w, x, z)$  needs to vary over  $z$ . If neither of these conditions are met then the exclusion restriction does not help in tightening the bounds. A necessary condition for the bounds to be tightened is that  $P(x, z)$  does depend on  $z$ .

There is nothing in the definition of the bounds obtained above that forces the minimum of the upper bounds to lie above the maximum of the lower bounds if the exclusion restriction is false. Thus in cases where  $W$  and  $Z$  are not conditionally independent it is possible for the bounds to cross and the upper bound for some values of  $Z$  to lie below the lower bounds. Later we construct a test of the null hypothesis that the bounds are equal against the alternative that the upper bound is lower than the lower bound. If we reject the hypothesis this is evidence against the exclusion restriction. However it is possible that the restriction is false but that the bounds do not cross.

**Weakening the exclusion restriction: Monotonicity** Strong exclusion restrictions of the type discussed above may not always be credible. As pointed out by Manski and Pepper (2000), however, we might be prepared to assume the direction of the relationship between  $w$  and  $z$ . Thus we now derive bounds under the assumption that the distribution of wages decreases monotonically

with the wage, i.e.

$$F(w|x, z') \leq F(w|x, z) \quad \forall w, x, z, z' \text{ with } z < z'. \quad (8)$$

This means that a higher value of the instrument  $Z$  will lead to a distribution of wages that stochastically dominates the distribution of wages with lower values of  $Z$ .<sup>7</sup> To exploit this restriction we can find tightest bounds over the support of  $Z$  and then integrate out  $Z$ . For a value of  $Z = z_1$  the best lower bound is the largest lower bound over  $z \geq z_1$  in the support of  $Z$ . This is given by

$$F(w|x, z_1) \geq F^l(w|x, z_1) \equiv \max_{z \geq z_1} F(w|E = 1, x, z)P(x, z). \quad (9)$$

Similarly we can obtain a best upper bound at  $Z = z_1$  by choosing the smallest possible upper bound over the support of  $Z$  such that  $z \leq z_1$ :

$$F(w|x, z_1) \leq F^u(w|x, z_1) \equiv \min_{z \leq z_1} F(w|E = 1, x, z)P(x, z) + 1 - P(x, z). \quad (10)$$

The bounds to the distribution of  $F(w|x)$  may then be obtained by integrating over the distribution of  $Z$  given  $X = x$ , i.e.

$$E [F^l(w|x, Z)|x] \leq F(w|x) \leq E [F^u(w|x, Z)|x]. \quad (11)$$

For the bounds to the distribution of wages  $F(w|x)$  to be tightened using the monotonicity restriction at some value of  $W = w$  it has to be that either the lower or the upper bound is increasing over some range of the support of the instrument  $Z$ , subject to them not crossing at any value of  $Z$ . To interpret what this means, observe that

$$F(w|E = 1, x, z)P(x, z) = F(w|x, z)\Pr(E = 1|W \leq w, x, z)$$

and that  $F(w|x, z)$  is *decreasing* in  $z$  by the monotonicity assumption. Thus in order for the lower bound function to be increasing in  $z$ ,  $\Pr(E = 1|W \leq w, x, z)$  needs to be increasing for some  $z$ . Analogously, in order for the upper bound function to be increasing in  $z$ ,  $\Pr(E = 1|W > w, x, z)$  needs to be *decreasing* for some  $z$ .

---

<sup>7</sup>The reverse assumption is covered in the analysis because we can choose the sign of  $Z$ .



### 2.1.3 Bounds to within group inequality

Let  $q_1 < q_2$  with  $P(x) < q_1$  and  $q_2 < 1 - P(x)$  and denote corresponding quantiles given  $x$  by  $w^{q_2}(x)$  and  $w^{q_1}(x)$ . To measure inequality within our framework we will be estimating bounds to the differences between quantiles:

$$D(x) = w^{q_2}(x) - w^{q_1}(x). \quad (12)$$

An example is the interquartile range. To obtain the bounds, note that

$$F(w|E = 0, x) = \frac{F(w|x) - P(x)F(w|E = 1, x)}{1 - P(x)}.$$

Since  $F(w|E = 0, x)$  is non-decreasing in  $w$ , the equality places a restriction on  $F(w|x)$ :  $F(w|x)$  cannot increase slower than  $P(x)F(w|E = 1, x)$ . This provides the upper bound on  $D(x)$ .

To be more precise, let  $w^{q_1(u)}(x)$  and  $w^{q_1(l)}(x)$  be the upper and lower bounds to the  $q_1$ th quantile of  $F(w|x)$ . For any  $w_0$  between  $w^{q_1(l)}(x)$  and  $w^{q_1(u)}(x)$ ,  $F(w|x)$  with  $F(w_0|x) = q_1$  is a candidate CDF. The slowest it can increase is by  $P(x)F(w|E = 1, x)$  and when it does, the implied CDF lies entirely between the bounds. This class of CDF can be denoted by

$$P(x) [F(w|E = 1, x) - F(w_0|E = 1, x)] + q_1.$$

Any  $F(w|x)$  which is “parallel” to  $P(x)F(w|E = 1, x)$  must be one of these. The  $q_2$ th quantile of this CDF is  $F^{-1}(F(w_0|E = 1, x) + (q_2 - q_1)/P(x)|E = 1, x)$  and thus the upper bound of  $D(x)$  is

$$\sup_{w_0 \in [w^{q_1(l)}(x), w^{q_1(u)}(x)]} F^{-1}(F(w_0|E = 1, x) + (q_2 - q_1)/P(x)|E = 1, x) - w_0.$$

Clearly the lower bound is  $\max(0, w^{q_2(l)}(x) - w^{q_1(u)}(x))$ .

It turns out that imposing the implication of  $F(w|E = 0, x)$  being a CDF on obtaining the bounds makes them considerably tighter in practice.

### 2.1.4 Bounding wage differentials between groups and their change over time

We will present bounds to the difference in median wages across education groups, gender, cohort and age. In contrast to the case where we bound differences in the quantiles of the same distribution, there is no restriction on the bounds to differences in the quantiles across different values of  $x$ . Thus the bounds to the difference between any groups has to nest the possibility that the true quantile is at the upper bound for one group and the lower bound for the other. Similarly when we look at changes in differentials over time then we have to allow for the differential to be at the lower bound in one year and the upper bound in the other. Consider the case with two conditioning variables, education and time. We are interested in  $D_t^q = w^q(ed_1, t) - w^q(ed_0, t)$ . This is given by:

$$w^{q(l)}(ed_1, t) - w^{q(u)}(ed_0, t) \leq D_t^q \leq w^{q(u)}(ed_1, t) - w^{q(l)}(ed_0, t). \quad (13)$$

Similarly  $\Delta D_t^{q(l)}$ , the lower bound to  $D_t^q - D_s^q$ , is given by

$$\left\{ w^{q(l)}(ed_1, t) - w^{q(u)}(ed_0, t) \right\} - \left\{ w^{q(u)}(ed_1, s) - w^{q(l)}(ed_0, s) \right\}$$

and  $\Delta D_t^{q(u)}$ , the upper bound, is given by

$$\left\{ w^{q(u)}(ed_1, t) - w^{q(l)}(ed_0, t) \right\} - \left\{ w^{q(l)}(ed_1, s) - w^{q(u)}(ed_0, s) \right\}.$$

Thus, even if the bounds to the quantiles are tight, as the bounds to differentials will be much larger and those of the *change* in differentials larger still, it could be that the bounds to the economic parameter of interest (say the changes in educational wage differentials over time) become too far apart to draw any interesting conclusions. We therefore consider two types of restrictions that will make these bounds narrower.

The first is to assume observables are independent of unobservables and they are additively separable. This is a fairly common assumption made in both the parametric and semiparametric selection literature. However we do not report the results exploiting this restriction as it leads the bounds to cross in most cases, implying the restriction is invalid.

The second restriction that helps tighten the bounds to the *change* in educational differentials is also on the functional form of wages. Let  $\Delta D_{ts}^q(a)$  denote the change in the educational differential between period  $t$  and  $s$  at age  $a$ . We assume that the change in education differentials over time are the same across an age group denoted by  $A$ . Thus for a given quantile  $q$ , we have:

$$\Delta D_{ts}^q(a) = \Delta D_{ts}^q \quad \forall a \in A. \quad (14)$$

Since this would hold if the age effect on the  $q$ th log-wage quantile for the relevant group is additively separable we call equation (14) additivity assumption. The bounds to the change in the education differentials are then

$$\max_{a \in A} \Delta D_{ts}^{(l)q}(a) \leq \Delta D_{ts}^q \leq \min_{a \in A} \Delta D_{ts}^{(u)q}(a).$$

### 2.1.5 The relationship of our restrictions to those of the parametric approach

The median restriction imposes some degree of positive selection among workers which is not an implication of the standard bivariate normal selection model (Heckman, 1974, 1979). The latter in principle also allows for workers to be negatively selected or randomly selected although the direction of the selection has to be always the same for every quantile. In this sense the results are not directly comparable. However, the model with the exclusion restrictions nests directly any nonparametric selection model that uses such restrictions and allows any form of selection, positive, negative or none. This is also true for the model with the monotonicity restriction which nests all models that use exclusions. Once these restrictions are combined with the median restriction this nesting property is no longer valid since positive selection (to a degree) is imposed.

## 3 Estimation

Our main focus will be the bounds to the quantiles. To estimate these we first estimate the bounds to the distribution of wages. We now describe the non-parametric estimation procedure we have used.

The conditioning vector  $X$  includes gender, education, age and time. Estimation of the worst case bounds and bounds with monotonicity require the estimation of the employment probability and the distribution of wages observed amongst the workers for each possible set of characteristics  $X$ . For tractability we limit the number of cells as follows. We define three education groups: Those who finished full time education at the age of 16 (statutory schooling), those who continued until 18 (high school graduates) and those who completed after 18 (at least some college). We limit the number of ages at which we estimate the bounds for to the ages of 25, 30, 35, 40, 45, 50 and 55 smoothing over neighboring age groups using a quartic kernel as described below. We also pool years in pairs from 1978/1979 until 1998/2000.<sup>8</sup>

Thus the probability of employment for an individual with characteristics  $x_k$  (age, education, gender and time period) is estimated by

$$P(x_k) = \frac{\sum_{i=1}^N I(E_i = 1) \kappa_k(x_i)}{\sum_{i=1}^N \kappa_k(x_i)}$$

where  $I(A)$  is the indicator function which equals one whenever  $A$  holds and zero otherwise and the weights  $\kappa_k(x_i)$  are defined by

$$\begin{aligned} \kappa_k(x_i) &= I(\text{year}_i = \text{year}_k) I(\text{ed}_i = \text{ed}_k) I(\text{gender}_i = \text{gender}_k) \mu_k(\text{age}_i) \text{ and} \\ \mu_k(\text{age}_i) &= \left( \frac{\text{age}_i - \text{age}_k}{3} + 1 \right)^2 \left( \frac{\text{age}_i - \text{age}_k}{3} - 1 \right)^2 I(|\text{age}_i - \text{age}_k| < 3). \end{aligned}$$

To estimate the empirical distribution of wages for workers we found it advantageous to allow for some smoothing. Thus the estimator we use is

$$F(w|E_i = 1, x_k) = \frac{\sum_{i=1}^N \Phi\left(\frac{w-w_i}{h}\right) I(E_i = 1) \kappa_k(x_i)}{\sum_{i=1}^N I(E_i = 1) \kappa_k(x_i)}$$

where  $h$  is set at a fifth of the standard deviation of wages.

The next estimation problem, relevant for computing the bounds with exclusion or monotonicity restrictions is the estimation of the probability of employment and the distribution of wages

---

<sup>8</sup>As we only have one quarter of data for the year 2000, we pool these individuals with those sample in 1998 or 1999

conditional on the instrument  $Z$  which in our case is the out of work income and can be regarded as continuous.

To reduce the computational burden we use the percentile ranks of out of work income  $Z$ . We then estimate the bounds to the distribution of wages and the probability of employment only at a subset of the percentile ranks, every five percentile. The weights for estimation are now given by

$$\kappa_k(x_i, z_i) = I(\text{year}_i = \text{year}_k)I(\text{ed}_i = \text{ed}_k)I(\text{gender}_i = \text{gender}_k)\mu_k(\text{age}_i)\phi_k(z_i)$$

where

$$\phi_k(z_i) = \left(\frac{z_i - z_k}{0.2} + 1\right)^2 \left(\frac{z_i - z_k}{0.2} - 1\right)^2 I(|z_i - z_k| \leq 0.2).$$

The estimated bounds are then substituted in equations (7), (9) and (10) to obtain estimates of the bounds under the exclusion and the monotonicity restriction, respectively. In the latter case we need to integrate over the distribution of the instrument (see (11)) which we do using the empirical distribution of the instrument  $Z$  given  $X = x$ . The bounds to the quantiles are then estimated by solving the analogous equations to those in (3).

We construct confidence intervals for the parameters of interest, namely the quantiles, the differentials across groups and the changes in the differentials over time using the bootstrap and applying the results of Imbens and Manski (2004). These are narrower than the confidence intervals for the estimated identification region itself.

### 3.1 Specification tests

#### 3.1.1 Testing that the bounds do not cross

When the upper and lower bounds are estimated invoking the exclusion or the monotonicity restrictions, there is nothing to prevent the upper bound to be less than or equal to the lower bound, which means that the restrictions do have a rejectable implication. Nevertheless it should be stressed that the restrictions may be invalid and the bounds may never cross. Hence the specification tests are not asymptotically uniformly powerful tests of the restrictions themselves

but rather that the bounds do not cross. If they do cross in the population, however, this does suggest that the restrictions may be invalid.

Even if in the population the bounds are equal, in any finite sample they will often cross just because of sampling error. Thus we need a formal test for the hypothesis of a point estimate against the alternative that bounds cross. To achieve this we use the following test statistics: Denoting the sum over all discrete values of  $X$  by  $\Sigma_X$ ,

**For Exclusion**

$$T_E = \sum_X \left[ \max_w \left\{ I(\hat{F}^{ub}(w|x) - \hat{F}^{lb}(w|x) < 0) (\hat{F}^{ub}(w|x) - \hat{F}^{lb}(w|x))^2 \right\} \right], \quad (15)$$

**For Monotonicity**

$$T_M = \sum_X \left[ \max_w \left\{ \max_z \left[ I(\hat{F}^{ub}(w|x, z) - \hat{F}^{lb}(w|x, z) < 0) (\hat{F}^{ub}(w|x, z) - \hat{F}^{lb}(w|x, z))^2 \right] \right\} \right] \quad (16)$$

where  $\hat{F}^{ub}(w|x)$  and  $\hat{F}^{ub}(w|x, z)$  are the estimate of the upper bound under exclusion and under monotonicity, respectively and similarly for the lower bounds. We then estimate the asymptotic distribution of the test statistic under the null using the bootstrap. To achieve this we draw bootstrap samples by

$$T_E^* = \sum_X \max_w \left[ I(\Delta \tilde{F}(w|x) - \Delta \hat{F}(w|x) < 0) (\Delta \tilde{F}(w|x) - \Delta \hat{F}(w|x))^2 \right] \text{ and}$$

$$T_M^* = \sum_X \left[ \max_w \left\{ \max_z \left[ I(\Delta \tilde{F}(w|x, z) - \Delta \hat{F}(w|x, z) < 0) (\Delta \tilde{F}(w|x, z) - \Delta \hat{F}(w|x, z))^2 \right] \right\} \right]$$

where  $\tilde{\cdot}$  denotes an estimate based on a bootstrap sample and  $\Delta$  denotes a difference between the upper and lower bounds; for example  $\Delta \hat{F}(w|x) = \hat{F}^{ub}(w|x) - \hat{F}^{lb}(w|x)$  and similarly for all other expressions preceded by  $\Delta$ . From the simulated values of this centered bootstrap test statistic we can then derive the p-values of our test.

In this way, we approach this testing problem analogously to testing for the location of multiple means. We have not proved that the above centering procedure provides the appropriate critical values for our test. However in the appendix we do provide some Monte Carlo evidence that these test statistics have good size and power properties in our context. See Appendix A.

**Deriving point estimates when the bounds cross** It is also necessary to derive consistent estimates of the bounds in the case where they cross. Under the null that the difference between the upper and lower bounds is zero, then the both the upper and lower bound estimates are consistent estimates of the actual quantile of wages we are interested in. Choosing either the estimate of upper or of the lower bound would give us consistent estimates of this quantile then, but a more efficient approach would be to use a weighted combination of the upper and lower bounds i.e.

$$\hat{w}^q = \alpha \hat{w}^{q(u)} + (1 - \alpha) \hat{w}^{q(l)}$$

where  $\alpha$  is chosen optimally to minimize the asymptotic variance of  $\hat{w}^q$  using the estimated asymptotic distribution of  $\hat{w}^{q(u)}$  and  $\hat{w}^{q(l)}$ . We use the bootstrap to estimate the distribution of the upper and lower bounds.

### 3.1.2 Testing for the absence of selection effects

The worst case bounds cannot be informative about whether there is a non-random selection problem or not; they can just tell us the extent to which selection can bias the results. However, when we impose the monotonicity restriction or the exclusion restriction the bounds to the distribution may not include the observed distribution which is evidence that selection does bias the results. However the difference may not be statistically significant and thus we develop a test statistic for the null hypothesis that the observed distribution is within the bounds.

Under the null hypothesis we have that

$$F^{lb}(w|x) \leq F(w|x, E = 1) \leq F^{ub}(w|x) \quad .$$

The test statistic we use then

$$T_S = \sum_X \max_w \left\{ \max_x \left[ \hat{F}^{lb}(w|x) - \hat{F}(w|x, E = 1), \hat{F}(w|x, E = 1) - \hat{F}^{ub}(w|x) \right]^2 \right. \\ \left. \times I \left( \max_x \left[ \hat{F}^{lb}(w|x) - \hat{F}(w|x, E = 1), \hat{F}(w|x, E = 1) - \hat{F}^{ub}(w|x) \right] \right) > 0 \right\} .$$

This provides a joint test for all the values of  $x$ . To obtain the critical values we again use the bootstrap after recentering so that the null is satisfied.

## 4 Results

### 4.1 Data and variable definitions

The data we used for the analysis is the pooled repeated cross sections of the UK Family Expenditure Surveys (FES) from 1978 to the first quarter 2000. We included in our sample all men and women between the ages of 23 and 59 who were not in full time education.<sup>9</sup> This gave us a sample of 187,467 individuals in total. Hourly wages, which are the object of the analysis are defined as usual weekly earnings divided by usual weekly hours (inclusive of overtime) and are deflated by the consumer all items quarterly retail price index. Deflated wages lower than 50p an hour were also treated as missing. We defined individuals to be in “work” (i.e.  $E = 1$ ) if they were reported themselves as being employed, whether full or part time or self employed over the last week. We treated the self-employed as employed in estimating employment probabilities. However, since wages and hours of work are not reliably measured for this group we assume their wages are missing at random and so exclude them from the calculation of  $F(w|E = 1, x)$ .

**Constructing out of work income** For the models where we use an exclusion or monotonicity restriction the instrument for employment will be the welfare benefits that the person would be eligible for when out of work. This variable has been used before by Blundell, Reed and Stoker (2003) and is constructed from the IFS tax and welfare-benefit model. More specifically for singles we use the benefit level for which they would be entitled if they did not work. For married or cohabiting individuals we take the household benefit level that they would be eligible for if neither worked. We show below that this instrument is a very significant source of variation in employment rates across individuals.

---

<sup>9</sup>In the UK university education is completed for most students by the age of 22.



The source of variation for out-of-work income is the demographic composition of the household and the housing costs that the household faces. They vary by region and time and the numerous policy changes that have occurred. The key policy changes have been as follows: First allowance for children has increased so that the childless have experienced relative falls in their out of work income. Second there was a switch in housing policy away from rent ceilings towards subsidies. Thus as rent increased to market levels at differential rates across regions the subsidies paid to the unemployed also increased at differential rates, substantially increasing the replacement rates and reducing the incentives to work for those with high levels of housing.

The dependence of the instrument on demographic composition clearly undermines it as an excludable instrument, unless one controls for demographics in the wage equation as in Blundell, Reed and Stoker (2003). We define the instrument as the residual in a regression of out of work income on household composition. However, it is still the case that the out of work income depends on household characteristics including region of residence and in particular housing. The latter is the key point since better housing implies more out of work income because of the way housing benefit operates, covering housing costs for individuals facing financial hardship in the short run. Hence the instrument may well shift the wage distribution downwards which causes a problem with its use as an excluded variable but motivates the monotonicity restriction.

Of course it is imperative that the instrument is related to employment. We test the hypothesis that our measure of out of work income does not affect employment conditional on the other observable characteristics (age, time, gender and education) and we reject both overall and within education and gender cells with a p-value of zero in all cases. The distributions of our test statistics under the null are approximated by the bootstrap.

## 4.2 The validity of our restrictions

### 4.2.1 The median restriction

The median restriction will play an important role in some of the results that follow. It is an identifying assumption at least for the data we have at hand. In this section we present some circumstantial evidence in support of this restriction.

The median restriction is most likely to be invalid for two different groups of workers. First high productivity women who do not work as they are married to high productivity men (as argued by Neal (2004)) and second older individuals with high productivity for whom the wealth effect due to lifetime asset accumulation may dominate the substitution effect in the decision to continue working.

In Figure 4 we use longitudinal data from the British Household Panel Survey 1991-2001 (BHPS) to show that such effects may not be that important. We have regressed log wages for each year of the panel separately on age and education and allocated workers a residual (i.e. actual wage minus predicted wage). We then split the sample into those who were ever unemployed over the sample and those who were continuously employed. Figure 4 shows that the quantiles of the distribution of wages of those continuously in work lie above the quantiles of the distribution for those who have had a work interruption, even controlling for factors such as age and education which are important determinants of unemployment. The median residual wage for workers is always higher than the median for non-workers. In fact this graph, taken at face value would provide support for the stochastic dominance assumption. Nevertheless we are missing individuals who never work, which is approximately 5.63% of men and 11.32% of women which highlights the fact that identifying assumptions are not testable.

Further support for the median restriction is given below when it is tested in combination with the monotonicity restriction.

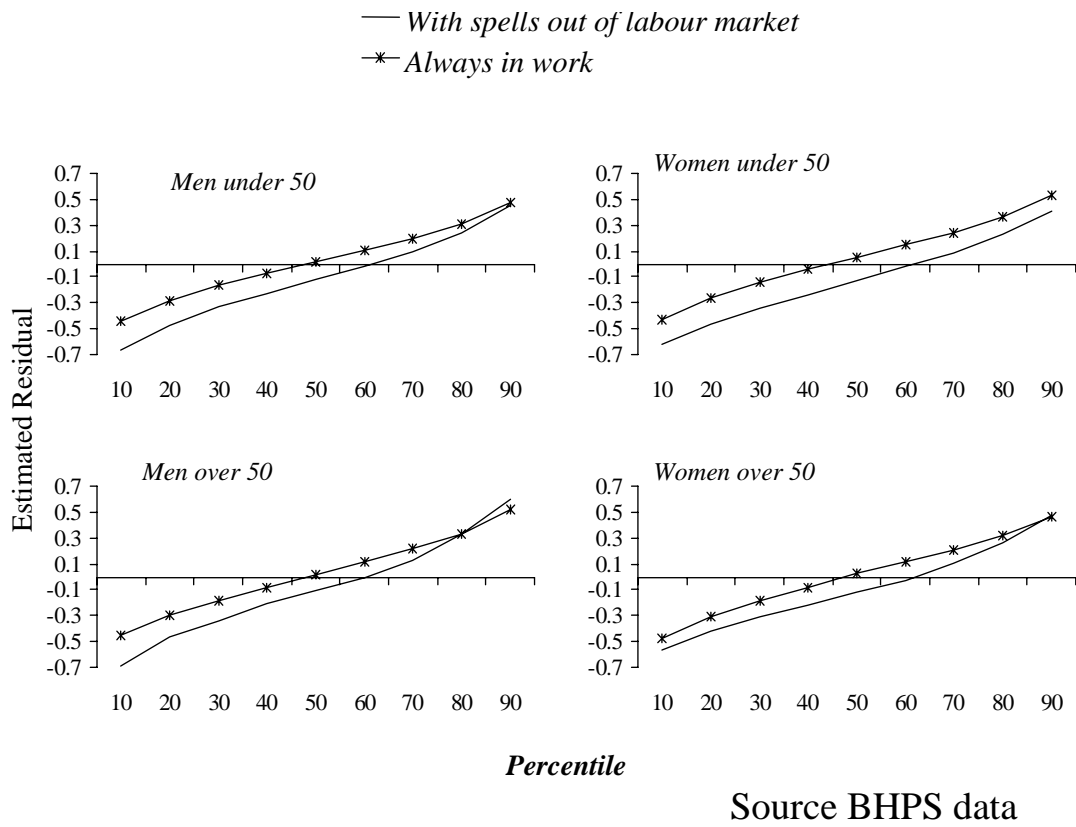


Figure 4: Distributions of residual wages by gender, education and work histories/futures

### 4.2.2 Restrictions about the relationship between wages and out of work income

There are two restrictions we test. First is that the bounds do not cross when we assume that wages are independent of out of work income, denoted by  $Z$ , conditional on  $X$ ; this is the exclusion restriction. The second is that the bounds do not cross when we impose the monotonicity restrictions, namely that the distribution of wages for those whose (potential) out of work income is higher stochastically dominates the distribution of wages for those with lower (potential) out of work income.

For women, no restriction can be rejected. This is not surprising given the low employment rate of women and hence the very wide worst case bounds. For men the results are more conclusive. The exclusion restriction is rejected both overall and within each education group (in each case the significance levels are below 0.1%). On the other hand the monotonicity restriction is never rejected and in each case the p-values are larger than 99%. This leads us to reject the exclusion restriction as an identification strategy and to base our results on the monotonicity restriction.

We can also test whether the combined monotonicity and median restriction is rejected. This is never the case and our preferred model is therefore one which imposes both this and the monotonicity restriction.

### 4.2.3 Selection effects

Most of our tests here failed to reject the absence of non-random selection. There was one notable exception, the significance level for men over 40 in the statutory schooling group was 2.2%. Remember that this is a test with weak power against some alternatives, since we are not testing the null of no selection effects but that the observed distribution lies inside the upper and lower bounds.

### 4.3 Changes in the distribution of wages

We now use our empirical approach to reexamine the changing distribution of wages in the UK 1978-2000 in the light of these selection issues. There are three broad questions of interest:

1. **The overall increase in inequality:** The increase in inequality amongst the wages of male workers in both the UK and the US has been well documented, but as this increase has come in tandem with a 10 percentage point fall in employment for men, it is important to establish whether such changes could have been driven by the changing composition of the workforce. Although changes to the female wage distribution are of interest, we focus simply on men, as the employment rates for women are too low to allow informative bounds on the lower percentiles of their wage distribution.
2. **The life cycle profile of wages.** How wages change with age, how this process may change over time and differ across skill groups are some of the key questions in empirical labour economics. (see Becker 1964). As we show above, there are large changes in employment over the lifecycle. This means that any observed increase/fall in wages with age could simply be driven by more/less skilled workers entering or leaving the labour market.
3. **Wage differentials between education groups and men and women.** The increase in pay differentials across education groups has been a key candidate for explaining changes in inequality. In the UK and the US, changes in the structure of male wages have occurred at the same time as a fast increase in female relative wages. We simply wish to establish whether these changes are robust to selection issues.

All results are accompanied by 95% confidence intervals for the unidentified parameter as in Imbens and Manski (2004). These intervals are constructed using the bootstrap.

### 4.3.1 Trends in inequality

We start by considering whether the oft cited conclusion that wage inequality has risen since the late 1970s is robust to compositional or selection effects. Figure 5 thus plots the upper and lower bound to the interquartile range, our inequality measure, from 1978 to 2000 for the male wage distribution. The central line shows, for comparison, what has happened to wage inequality

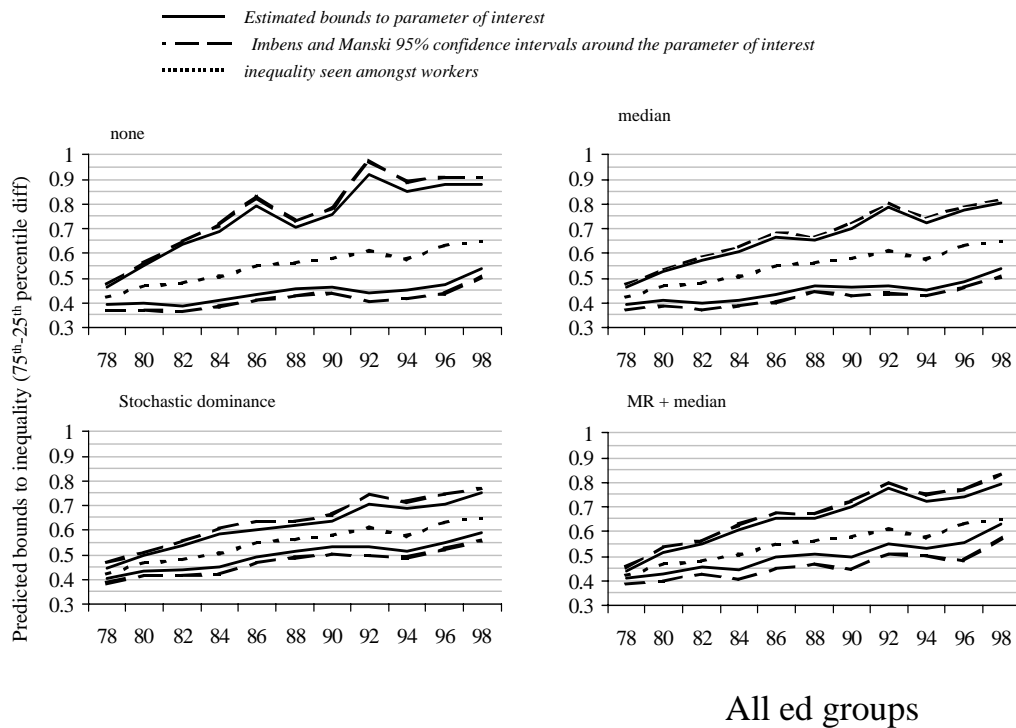


Figure 5: Bounds to interquartile range of  $\ln(\text{wages})$  1978-2000

amongst workers and the dotted lines give 95% confidence intervals. We can only say for certain that inequality has gone up if the lower bound at the end of the period is higher than the upper bound at the beginning of the period. The worst case lower bound in 1998-2000 is higher than the worst case upper bound in 1978, suggesting that inequality as measured by the interquartile

range must have risen by at least 0.075 log points (*0.017*).<sup>10</sup> This means that selection effects alone cannot explain the rise in inequality. We then show results using the median restriction, the assumption that workers wages stochastically dominate those for non-workers and the combination of the median and monotonicity restriction. Under the latter the interquartile range must have risen by at least 0.193 (*0.026*) log points which is less than the rise of the interquartile range of the distribution of wages for workers (0.226, *0.018*).

### 4.3.2 Within group inequality

A feature of the increase in inequality in Britain (as well as in the US) has been the large increase in within education group inequality. For the statutory schooling group the worst case bounds are just about capable of confirming that inequality did not decline over the entire sample period. Otherwise they are quite uninformative - which is not surprising given the large declines in employment. However, once we impose some structure on the employment process in the form of the median and the monotonicity restriction or in terms of stochastic dominance, inequality rose by at least 0.154 (*0.024*) log points in the former case and 0.064 (*0.021*) log points in the latter. This compares to an observed increase of 0.160 (*0.025*) log points for the workers. For the high school graduates inequality increased by at least 0.172 (*0.035*) log points even when we look at the worst case bounds. When we impose the median restriction combined with monotonicity the lower bound to the increase in inequality is slightly higher (0.190 log points, *0.036*). This compares to 0.233 log points (*0.038*) for the observed distribution of wages.

A similar picture emerges for the college graduates where the worst case bounds also imply a minimum increase of 0.098 (*0.039*) log points. The point estimates obtained using the median and monotonicity restrictions imply an overall increase in inequality of 0.174 (*0.039*) log point. This compares to an increase of 0.201 (*0.038*) log points for the distribution of wages for workers. Thus there is clear evidence of a substantial increase in within group wage inequality. The results imply

---

<sup>10</sup>The italicised number presented in parenthesis here and below are the standard error of the estimate computed using the non-parametric bootstrap.

that inequality increased by at least as much as implied by the observed distribution of wages. In Figure 6 we present bounds to the interquartile range over time together with confidence intervals as well as the interquartile range for those observed working. The bounds have been estimated under the assumptions of median and monotonicity restrictions.

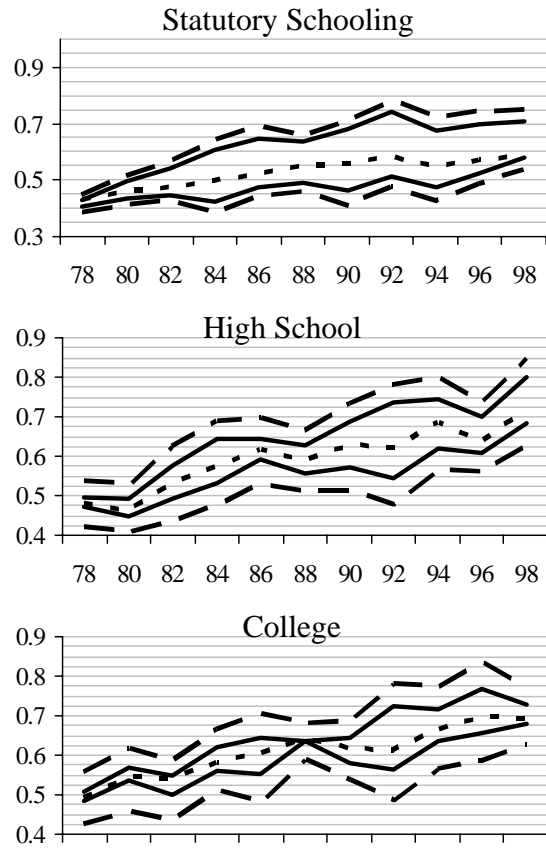


Figure 6: Inequality by schooling level



#### 4.4 Life-Cycle Wage profiles

In Figures 7 to 9 we present bounds to the median wages by education group over age for five cohorts each 10 years apart (1925, 35, 45, 55 and 65). Within each graph we present results based on the worst case bounds, the median restriction, median combined with monotonicity and for the case where selection into work is assumed random. Following these in Figure 10 we show results for females based only on the monotonicity restriction combined with the median. In these graphs moving from left to right we obtain the growth of wages by age. Moving from one cohort to the next at the same age we obtain the cohort/time effect. To establish growth over age we need to compare the upper bound of the median at the lower age to the lower bound at the higher age; Similarly for inter-cohort growth.

Those with just statutory schooling<sup>11</sup> (see Figure 7) have the lowest employment rates and therefore the widest bounds. Nevertheless even with the worst case bounds we can detect a growth of wages of at least 15% (2.4%) over ages 25-35 for the 1955 cohort and 7% (2.8%) for ages 35-45 for the 1945 one. The bounds become considerably tighter when we impose restrictions and almost all ambiguity is eliminated with the combination of median and monotonicity. Combining information across cohorts these show unambiguous growth of wages up until 45.

We can also detect substantial inter-cohort growth between the 1935 and 1945 cohorts (at least 10%, 2.6%), while between the more recent cohorts (1945 and 1955) growth was at least 5% (2.5%) at the age of 35 but the lower bound is below zero at the age of 30. Thus there is evidence that inter-cohort growth is declining for those with just statutory schooling, although we are forced to compare these at different ages.

---

<sup>11</sup>Those who left school by 16.

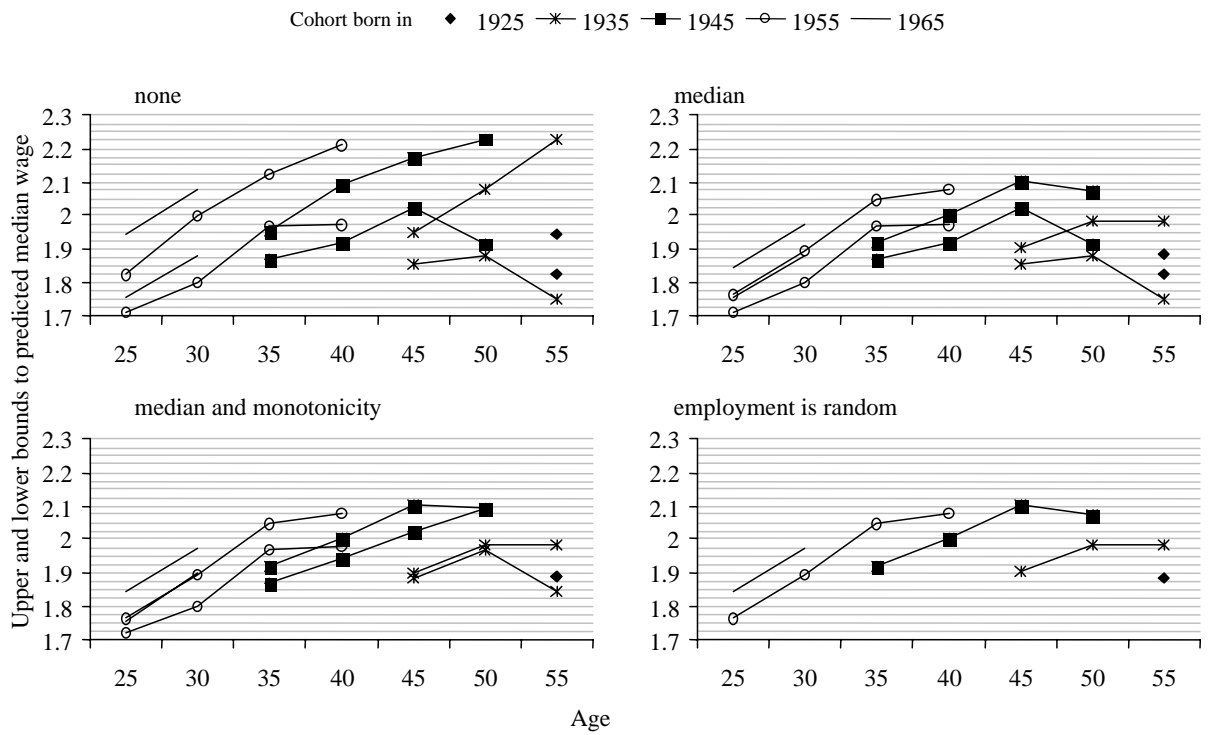


Figure 7: Life-cycle and cohort wage profiles for males who left school at 16 (the statutory schooling group)

For the high school graduates<sup>12</sup> (Figure 8) the median and the monotonicity restriction provide almost point estimates. The results there indicate that there has not been much inter-cohort growth since the 1945 cohort." However there is substantial "life-cycle" wage growth of the median of the wage distribution between the age of 25 and 40 for those born in 1955 and between the age of 35 and 45 for those born 10 years earlier.

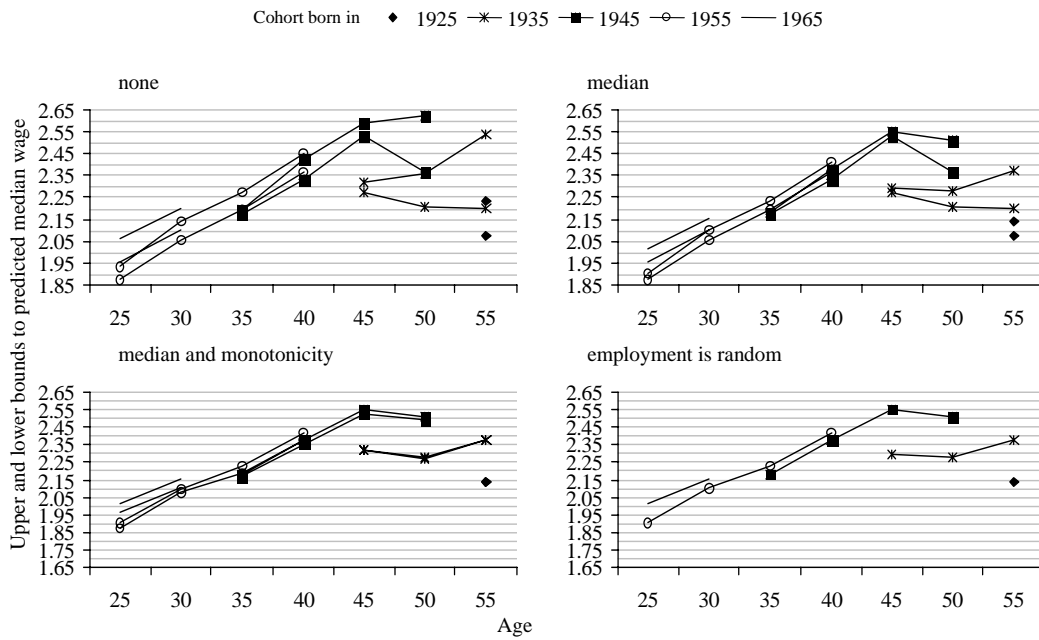


Figure 8: Life-cycle and cohort wage profiles for males who left school by 18 years of age

<sup>12</sup>Those who left school at 17 or 18.

For males with the highest education level (Figure 9) there also appears to be a declining growth of wages between cohorts and a declining life-cycle growth. There is no evidence of wages declining at older ages up to 55. Indeed for the oldest cohort wages keep growing between the ages of 50 and 55, and this is not an artifact of selection bias.

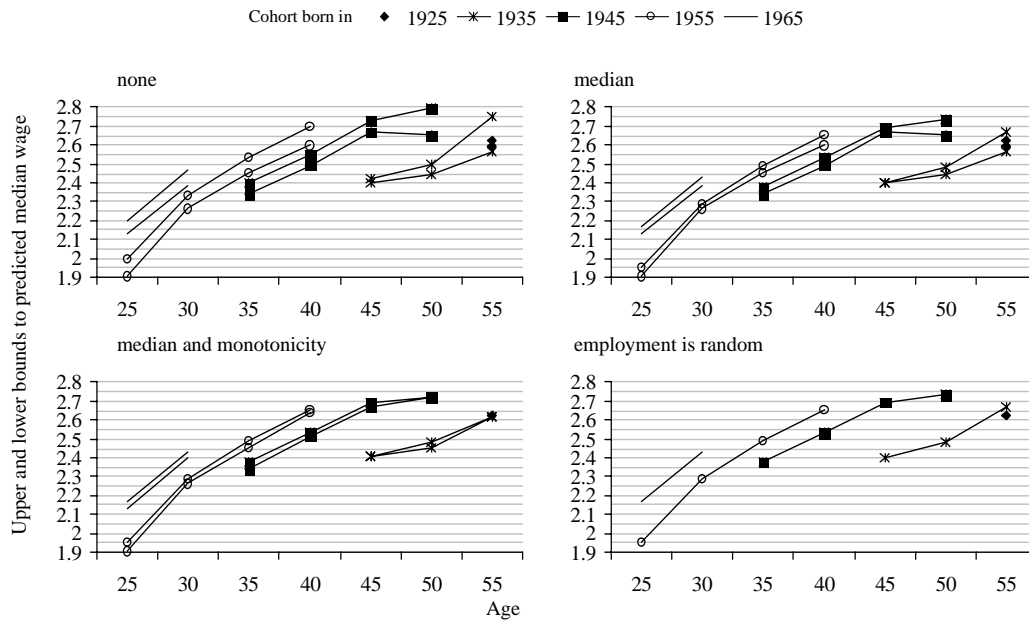


Figure 9: Life-cycle and cohort wage profiles for males who left school after 18 years of age

For females with just statutory schooling (top panel of Figure 10) it is hard to say anything for the young cohorts even with our restricted bounds. However for the 1945 cohort there has been a clear growth of wages with age. Moreover the growth between the 1945 and 1955 cohort at age 40 has been at least 13%. Some lifecycle and inter-cohort growth has been evident also for the two higher education groups (middle and lower panels of Figure 10).

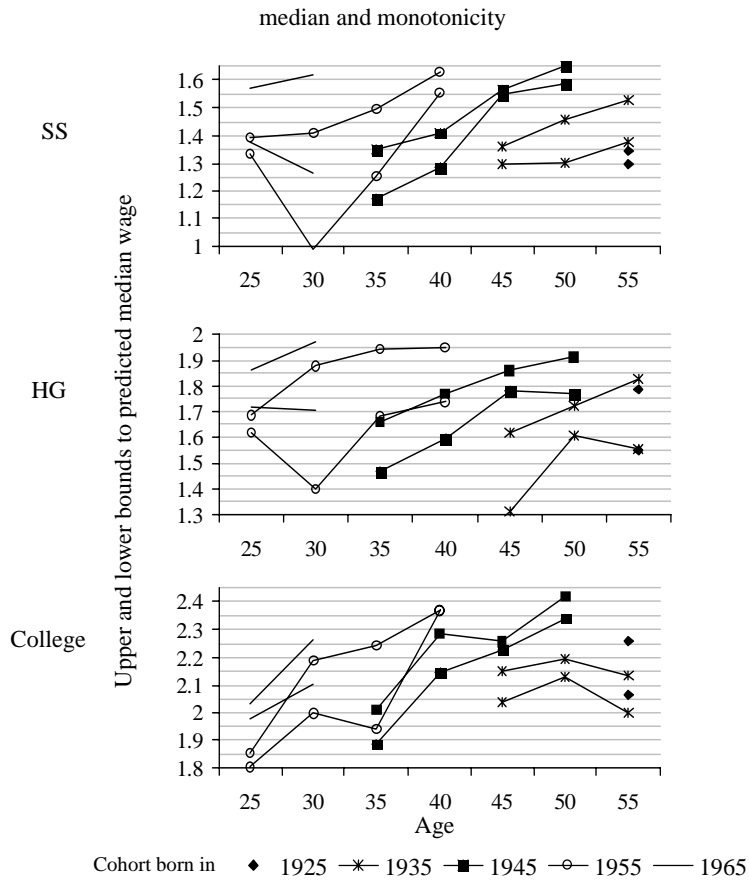


Figure 10: Lifecycle and cohort wage profiles by education group obtained from imposing both the median and monotonicity restrictions

## 4.5 Educational and gender differentials

### 4.5.1 Educational differentials

In Figure 11 we provide a picture of the educational differentials and how they change over time and across cohorts based on the monotonicity and median restrictions.

The difference in the medians of the wage distribution between those with some college education and the statutory schooling group clearly increase with age. However the picture is less clear after 45. We cannot confirm the presence of cohort or time effects in the returns although the bounds do not exclude that these could be quite high. This is a key point because it relates to whether returns to education have increased - we cannot be sure on the basis of these results. When we turn to the returns of college relative to high school a clear picture emerges of cohort effects implying that returns to college are increasing for younger cohorts, which is a source of increasing inequality.

To see this better in Figure 12 we present bounds to the changes in the differential between college and high school at 25 and 40 using the monotonicity and median restriction, between the years 1978 and 1998. We also improve precision and tighten the bounds further by presenting results based on the additivity restriction described earlier (see equation (14)). The block shows the bounds and the thin line shows the 95% confidence interval around the unidentified parameter.

On the basis of these results educational differentials increased by between 0.081 (*6.0%*) and 0.225 (*6.3%*) log points for 25 year olds. For 40 year olds the increase lies between 0.099 (*0.073*) and 0.186 (*0.083*) log points. When we impose the additivity restriction for 25 year olds the increase is estimated to be between 0.136 (*0.043*) and 0.170 (*0.044*) log points and much more precisely estimated. The additivity restriction for 40 year olds is invalid as can be seen by the fact that the restricted estimate lies outside the less restrictive bounds and this may explain the low estimated change in the differential.

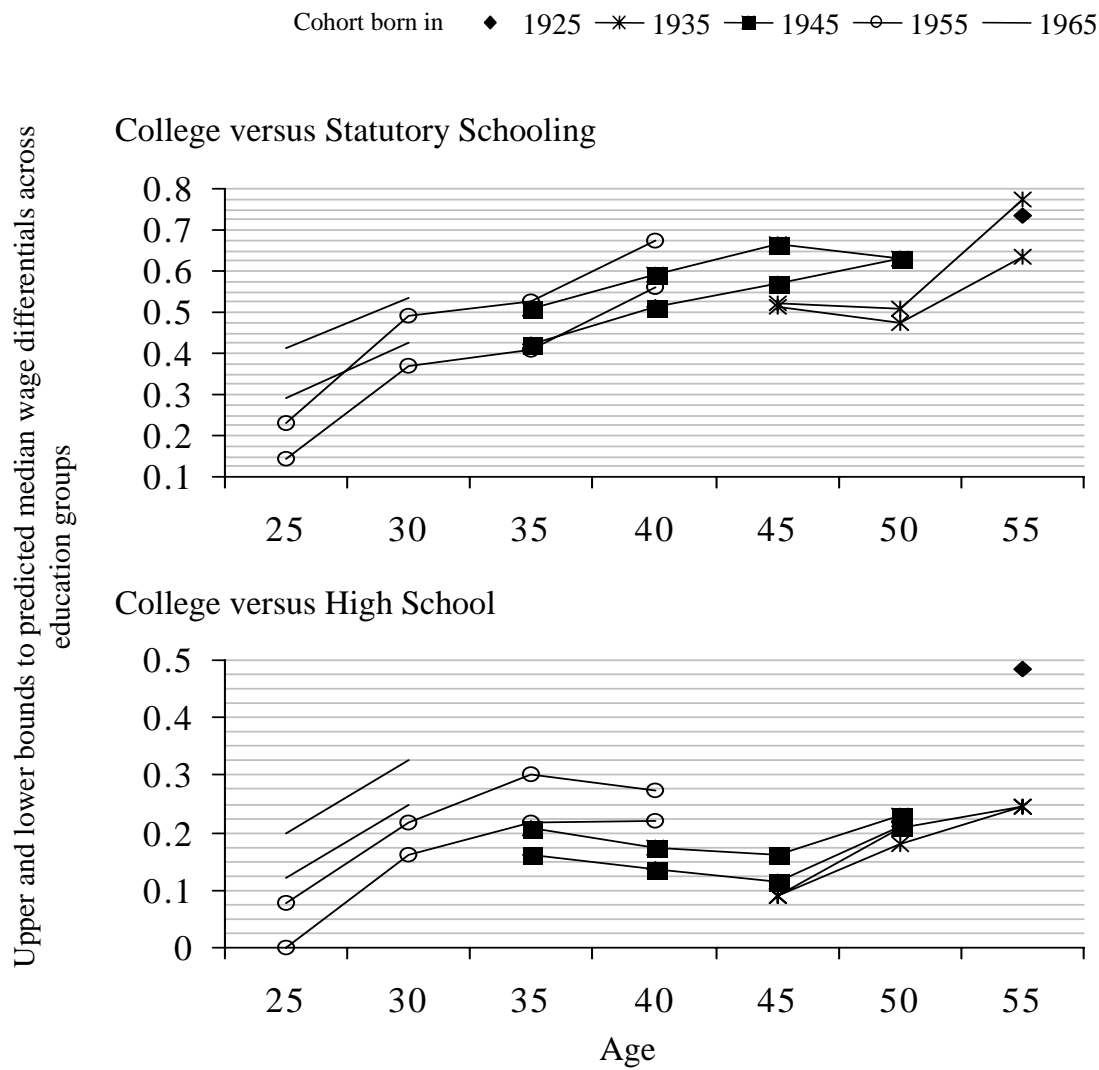


Figure 11: Bounds to the differences in median log wages between education groups

### Monotonicity plus median restrictions

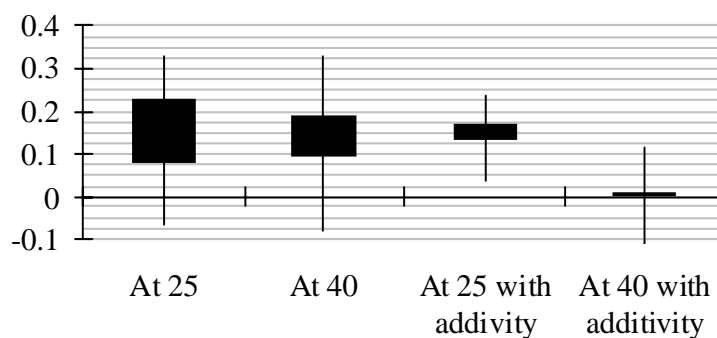


Figure 12: Bounds to the change in the educational differentials between college and high school for men at 25 and 40 between the years 1978 and 1998

For women all the estimated changes span zero, meaning that we cannot reject the possibility that the return to education has fallen or remained the same. Note that even the changes amongst women observed working are small and sometimes even negative so this failure to find a clear result should not be a surprise.

#### 4.5.2 Gender wage differentials

How do the wages of women compare to those of men? How have the relativities between men and women changed over time? Compositional issues are crucial here, selection process into work may be both very different for women compared to men and have changed with the large changes in employment rates. This problem is very similar to that faced by those interested in understanding the convergence of black and white observed wages in the 1960s and 1970s (see Butler and Heckman, 1977). Brown (1984) and Smith and Welch (1986) adopt different strategies to deal with this that can both be considered as special cases of our approach. The worst case bounds to the change in the differentials are uninformative because of the lower employment rates



for women, particularly the statutory schooling group. Figure 13 presents results based on the combination of the monotonicity, the median and the additivity restriction. Here we have also made the assumption that changes in differentials are the same for all those with less than college education, thus we can look together at the statutory schooling and the high school group. For this group the male/female differential declined from 1978 to 1998 by between 0.24 and 0.29 log points and the confidence interval implies that the difference is significant. The change observed between working men and women was about 0.21 log points. This suggests that composition effects conceal part of the improvement in the labour market position of women. The other declines are not significant, despite the fact that we get close to point estimates (SS at 40 and college at 40).

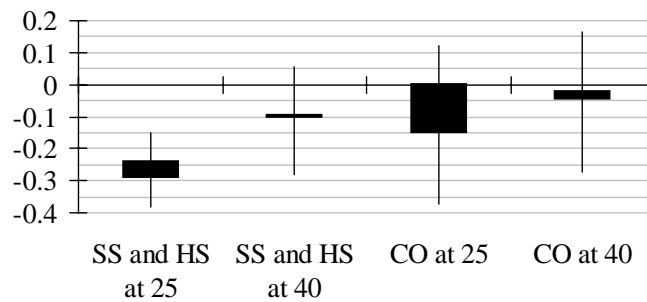


Figure 13: Bounds to changes to gender wage differentials by age and education groups between 1978 and 1998

## 5 Conclusions

This paper follows on from the large literature on wage inequality by investigating the possible impact of non random selection into employment. Our aim was to achieve this without making

unnecessarily strong behavioral and arbitrary statistical assumptions. This aim leads us to construct bounds to the wage distribution and its changes and we present both worst case bounds (i.e. those derived without any economic or statistical assumption) and also those tightened by restrictions that have a direct and easy economic interpretation. The methods are implemented using the UK Family Expenditure Survey 1978–2000.

While the worst case bounds are sometimes useful, for example in examining the wage growth across and within cohorts, they are in many other cases too wide to provide definitive answers. Our preferred model thus invokes two restrictions to tighten these bounds; the first is that wages of workers at the median are higher than the potential wage of non-workers at the median; the second is that potential wages do not decline with increases in potential welfare income out of work.

The main conclusions of existing studies of the distribution of wages in the UK/Britain, namely that inequality amongst men has risen both overall and within groups, that differentials between different education groups have risen (at least for men) and that differentials have narrowed between some groups of men and women are shown to be robust to selection; the lowest bound to the change in the interquartile range of male wages shows it increasing by almost a half from 0.437 in 1978/9 to 0.631 in 1998/2000, educational differentials amongst younger (under 40) men have increased by at least 0.14 log points and younger women with less than college level education saw their wages rises at least 24% faster than their male counterparts. These results are robust to selectivity bias subject of course to our assumptions.

## **A Monte Carlo simulations of the testing procedure**

We conducted Monte Carlo simulations to examine the power and size properties of our test. The null hypothesis is that the bounds are equal and the alternative is that they cross. In this appendix we describe the model used in the simulation.

The model under the null hypothesis is

$$\begin{aligned}
W &= \varepsilon_1 & (17) \\
Z &= I(\varepsilon_2 > 0) \\
E &= [I(W < 0.75) + I(W \geq 0.75)I(\varepsilon_3 > 0)] Z \\
&\quad + [I(W > -0.75) + I(W \leq -0.75)I(\varepsilon_3 > 0)] (1 - Z)
\end{aligned}$$

where  $\varepsilon_1$ ,  $\varepsilon_2$  and  $\varepsilon_3$  are independent standard normal random variables. Under the null hypothesis  $W$  and  $Z$  are independent. As an alternative we consider

$$W = \varepsilon_1 - \frac{1}{4}Z. \quad (18)$$

The structure of the determination of employment shown in equation (17) will mean that the bounds touch in the population over a positive range of wages which is our null. To see this note

$$\max_z F(w|E = 1, z) \Pr(E = 1|z) = \min_z [F(w|E = 1, z) \Pr(E = 1) + 1 - \Pr(E = 1|z)].$$

Given that  $F(w|z) = F(w)$ , this expression is identical to

$$\max_z F(w) \Pr(E = 1|z, W < w) = \min_z [F(w) \Pr(E = 1|z, W < w) + 1 - \Pr(E = 1|z)].$$

This can be rearranged to give:

$$\max_z F(w) \Pr(E = 1|z, W < w) = 1 - \max_z \{[1 - F(w)] \Pr(E = 1|z, W \geq w)\}.$$

The above expression can be true if and only if the following is true

$$\max_z \Pr(E = 1|z, W < w) = \max_z \Pr(E = 1|z, W \geq w) = 1. \quad (19)$$

$\Pr(E = 1|z) = 1$  is a sufficient but not a necessary condition for equation (19) to hold. The bounds will be equal when there is a value of  $z$  for which all those with wages *below*  $w$  work and a value of  $z$  for which all those with wages *above* the same  $w$  work. These values of  $z$  may not be the same. Both bounds reduce to an unconditional CDF.

Equation (17) and our Monte Carlo simulations forces condition (19) to hold for a large range of wages: Consider any value  $w$  between  $-0.75$  and  $0.75$ . When  $Z = 1$  and  $W < w$ , everyone works. When  $Z = 0$  and  $W > w$ , everyone works as well. So the upper and lower bounds for the model coincide for any  $w$  between  $-0.75$  and  $0.75$ . Figure 14 below shows how the estimated distribution functions obtained from imposing the exclusion restriction look like in the population.

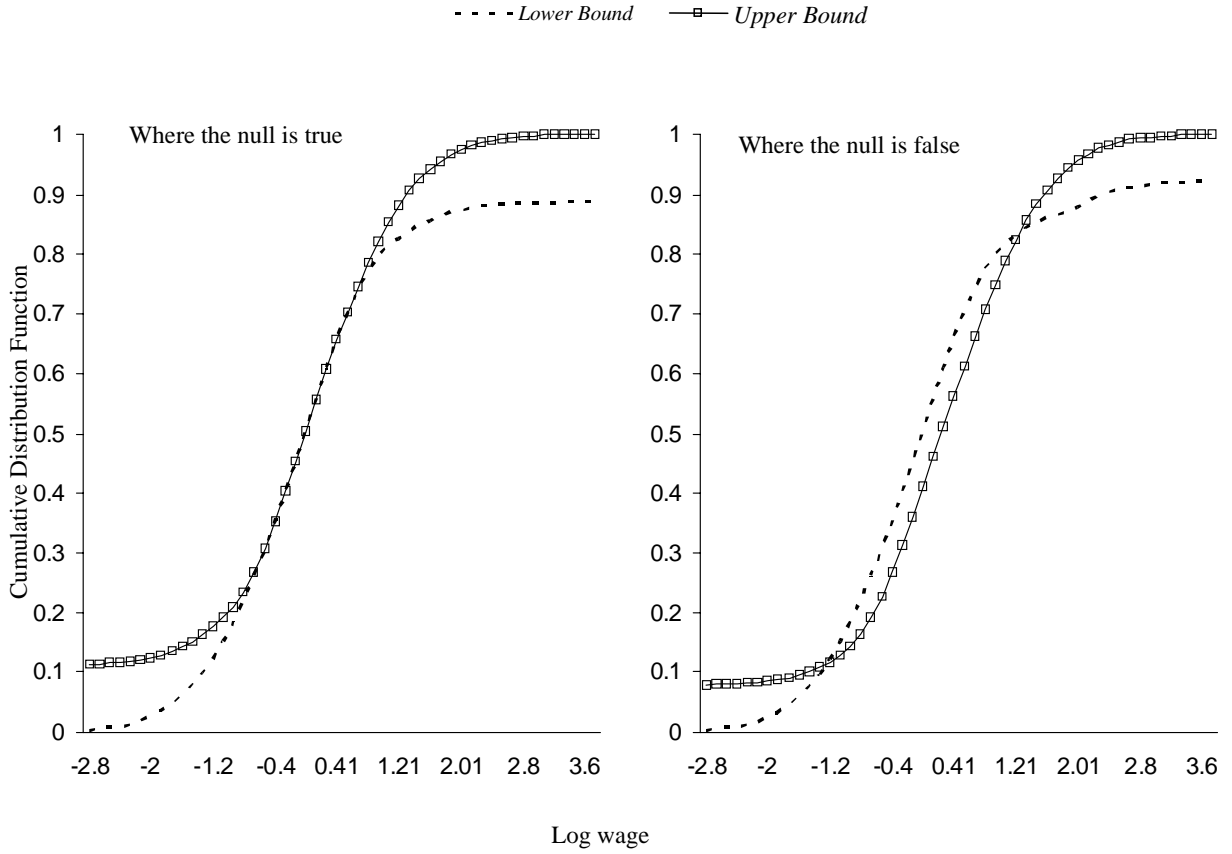


Figure 14: Estimated population distribution functions from simulated data

We carried out two Monte Carlo simulations; one with 500 observations and one with a 1000. For each of these two simulations we generated 1000 random samples from each of the two models expressed in equations (18) and (17). Within each replication we used the bootstrap with 114 draws to compute critical values for a nominal size of 5% for the test statistic in (15), as described in the main body of the paper. We also considered a version of the test where the  $\max_w$  (maximum over wages) is replaced by a  $\Sigma_w$

The results are shown in Table 1 below

Table 1: Power and size of our test

	Max		Sum	
	<i>N=500</i>	<i>N=1000</i>	<i>N=500</i>	<i>N=1000</i>
$H_o$ is true	0.048	0.049	0.040	0.031
$H_o$ is false	0.762	0.960	0.756	0.963

Note: rejection probabilities given

Table 1 shows first that both tests have good size properties with the actual size being very close to the nominal one, particularly for the *max* version that we use. Both tests have similarly good power properties, with the power being a bit better in the version we have adopted at the smaller sample size.

## B Using independence restrictions

Many empirical studies impose independence of the instrument used in correcting for selection from the unobservables in wages. This independence assumption is reflected in the fact that the selection model is single index and that the coefficient of the selection correction term(s) do not depend on the instrument. In this section we explore how such an independence assumption, but without any other assumptions can help tighten the bounds to returns to education or other characteristics  $X$ .

Suppose we partition the vector of observables into the sub-vectors  $X^1$  and  $X^2$  and suppose that the dependent variable can be written as

$$W = m(X^1, X^2) + \varepsilon$$

where  $F(\varepsilon|X^1, X^2) = F(\varepsilon|X^1)$  In this case none of the quantiles of  $\varepsilon$  depend on  $X^2$ . Hence we can write the  $q$ th quantile of  $W$  as

$$w^q(X^1, X^2) = m(X^1, X^2) + g(q, X^1).$$

In this context the impact of changing  $X^2$  (say education) is easily defined. Moreover the independence restriction can be used to obtain a tight bound for such a return. First note that the impact of changing  $X^2$  from  $B$  to  $A$  on the dependent variable, defined by  $w^q(X^1, A) - w^q(X^1, B) = \Delta_{X^2} m(X^1, X^2)$  does not depend on  $q$ .<sup>13</sup> Then under these assumptions the tightest bound on the return  $\Delta_{X^2} m(X^1, X^2)$  can be obtained by searching across quantiles. Thus the tightest bound takes the form

$$\max_q \left\{ w^{q(l)}(X^1, A) - w^{q(u)}(X^1, B) \right\} \leq \Delta_{X^2} m(X^1, X^2) \leq \min_q \left\{ w^{q(u)}(X^1, A) - w^{q(l)}(X^1, B) \right\}.$$

If the independence assumption is invalid the bounds may cross.

## References

- [1] Ahn, H. and J.L. Powell (1993) "Semiparametric Estimation of Censored Selection Models with a Nonparametric Selection Mechanism", *Journal of Econometrics*, 58, 3-29.
- [2] Becker G (1964) *Human capital; A Theoretical and Empirical Analysis, with Special Reference to Education*, New York, National Bureau of Economic Research
- [3] Blau Francine (1998) 'Trends in the Well-being of American Women, 1970-1995' *Journal of Economic Literature*, Mar 1998, Vol.XXXVI, No.1, pp.112-165
- [4] Blau Francine and Laurence Kahn (1997) 'Swimming Upstream: Trends in the Gender Wage Differential in the 1980's' *Journal of Labor Economics*, Jan 1997, Vol.15, No.1(1), pp.1-42
- [5] Blundell R, H Reed and T Stoker (2003) 'Interpreting Aggregate Wage Growth: The Role of Labor Market Participation' *American Economic Review*, Sep 2003, Vol.93, No.4, pp.1114-1131

---

<sup>13</sup>The analysis could also be carried out in terms of a continuous variable.

- [6] Brown C (1984) ‘Black-White Earnings Ratios Since the Civil Rights Act of 1964: The Importance of Labor Market Dropouts’ *The Quarterly Journal of Economics*, Vol. 99, No. 1. (Feb., 1984), pp. 31-44
- [7] Buchinsky M (1994) ‘Changes in the U.S. Wage Structure 1963-1987: Application of Quantile Regression’ *Econometrica*, Vol. 62, No. 2. (Mar., 1994), pp. 405-458
- [8] Buchinsky, M. (1998), The Dynamics of Changes in the Female Wage Distribution in the USA: A Quantile Regression Approach, *Journal of Applied Econometrics*, 13, 130.
- [9] Butler R and J Heckman (1977) “The Governments Impact on the Labour Market Status of Black Americans, a Critical Review” in L J Hausman et al. (eds) *Equal Rights and Industrial Relations* (Madison WI: Industrial Relations Research Association 1977) pp 235-81
- [10] Choi, K. (1990) “The Semiparametric Estimation of the Sample Selection Model Using Series Expansion and the Propensity Score”, manuscript, University of Chicago
- [11] Cosslett, S.R. (1991) “Distribution-Free Estimator of a Regression Model with Sample Selectivity”, in: W.A. Barnett, J.L. Powell and G. Tauchen, eds., *Nonparametric and Semiparametric Methods in Econometrics and Statistics*. Cambridge: Cambridge University Press.
- [12] Gosling A S Machin and C Meghir (2000) The Changing Distribution of Male Wages in the UK, *Review of Economic Studies* Oct 2000, Vol.67(4), No.233, pp.635- 666
- [13] Harkness S ‘The Gender Earnings Gap: Evidence from the UK’ *Fiscal Studies*, May 1996, Vol.17, No.2
- [14] Heckman J (1974) “Shadow Prices, Market Wages and Labour Supply” *Econometrica* 42 pp 679-694
- [15] Heckman (1979) “Sample Selection Bias as a Specification Error” *Econometrica*, Vol. 47, No. 1. (Jan), pp. 153-162.



- [16] Heckman J (1990) ‘Varieties of Selection Bias’ *The American Economic Review*, Vol. 80, No. 2, Papers and Proceedings of the Hundred and Second Annual Meeting of the American Economic Association. (May), pp. 313-318.
- [17] Heckman J, R. LaLonde and J. Smith (1999) “The Economics and Econometrics of Active Labour Market Programmes” in O Ashenfelter and D Card eds *Handbook of Labour Economics* Vol 3a North Holland
- [18] Heckman J and G Sedlacek (1985) “Heterogeneity, Aggregation and Market Wage Functions; An Empirical Model of Self Selection into the Labor Market” *Journal of Political Economy* 98 (6) pp 1077-1125
- [19] Heckman J J Smith and N Clements (1997) “Making the Most Out of Programme Evaluations and Social Experiments: Accounting for Heterogeneity in Programme Impacts” *The Review of Economic Studies*, Vol. 64, No. 4, Special Issue: Evaluation of Training and Other Social Programmes. (Oct), pp. 487-535.
- [20] Heckman J and E Vytlacil (2001) “Policy Relevant Treatment Effects,” *American Economic Review*, Papers and Proceedings, May 2001, 91(2): 107-111
- [21] Heckman J and E Vytlacil (2004) “Structural Equations, Treatment Effects and Econometric Policy Evaluation” *Econometrica*, forthcoming
- [22] Ichimura, H. and L.-F. Lee (1991) “Semiparametric Least Squares Estimation of Multiple Index Models: Single Equation Estimation”, in W.A. Barnett, J.L. Powell and G. Tauchen, eds., *Nonparametric and Semiparametric Methods in Econometrics and Statistics*. Cambridge: Cambridge University Press.
- [23] Imbens, G. and C. Manski (2004) “Confidence Intervals for Partially Identified Parameters” *Econometrica*, forthcoming

- [24] Juhn C, Murphy K and Pierce B (1993) “Wage Inequality and the Rise in Returns to Skill” *Journal of Political Economy* 108 pp 410-42
- [25] Koenker, Roger and Gilbert Bassett Jr. (1978) “Regression Quantiles”, *Econometrica*, 46, 33-50.
- [26] Manski C (1994) “The Selection Problem,” in *Advances In Econometrics, Sixth World Congress* Vol 1., Editor C. Sims, Cambridge University Press.
- [27] Manski C and J Pepper (2000) “Monotone Instrumental Variables: With Application to the Returns to Schooling,” *Econometrica*, Vol. 68, No. 4. (Jul., 2000), pp. 997-1010. working paper
- [28] Neal (2004) “The Measured Black-White Wage Gap Among Women is too Small” *Journal of Political Economy*, Feb 2004, Vol.112, No.1(2), pp.1- 28 JPE
- [29] Newey, W.K. (1988) “Two-Step Series Estimation of Sample Selection Models”, Princeton University, manuscript.
- [30] Powell, J.L. (1987) “Semiparametric Estimation of Bivariate Latent Variable Models”, Social Systems Research Institute, University of Wisconsin-Madison, Working Paper No. 8704.
- [31] Smith J and Welch F (1986) ‘*Closing the Gap; Forty Years of Economic Progress for Blacks*’ Rand Corporation, February