

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

Richard Blundell, Amanda Gosling, Hidehiko Ichimura and Costas Meghir¹

August 2003

¹Richard Blundell, Hidehiko Ichimura, and Costas Meghir are at the Department of Economics, University College London and at the Institute for Fiscal Studies, Amanda Gosling is at the Department of Economics, University of Essex and also at the Institute for Fiscal Studies.

Abstract

This paper presents estimates of the changing distribution of wages that are robust to possible selection effects. We find convincing evidence of an increase in overall inequality, changes in the “return” to education and increases in inequality within age and education groups. On the other hand we find that the increase in the relative wages of women may have been driven by selection.

1 Introduction and motivation

There has been a large literature on the way the distribution of observed wages has evolved over the past twenty to thirty years. Some economies, such as the US and the UK have seen large and unprecedented increases in inequality among the wages of workers. This is illustrated in Figure 1 where we show the way that the interquartile range of male and female log hourly wages has evolved for those who work. These increases in inequality have been associated with increased returns to education, cohort effects and increases in the returns to unobserved skill.¹ A variety of interpretations have been 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 can be attributed to permanent differences across cohorts and in changes in the returns to education.

Another aspect of the debate about changes in the distribution of wages concerns the relationship of male and female wages. Generally, both in the US and the UK the consensus is that female wages among workers have increased and converged with male wages.

However, in parallel with these momentous changes in the distribution of observed wages, the labour market participation rates for males and females have changed in dramatic ways. Female participation has tended to increase, particularly among women married to employed men. Male participation on the other hand has decreased to a very large degree. (see Figure 2).

The decline in male participation is not confined to older men, just reflecting the increase in early retirement. As illustrated in Figure 3 the decline over time occurs at all ages, although it is even more pronounced for men close to 55. Female employment on the other hand has risen (see same figures) and this has been particularly true for women younger than 35. Looking at figure 4 we also see that the change in employment has been heavily skill biased. Although male employment has declined for all education groups the highest decline has been for the least skilled group. Moreover for women the unskilled group has shown a slight decline, while most of the increase can be accounted for by the increase in

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

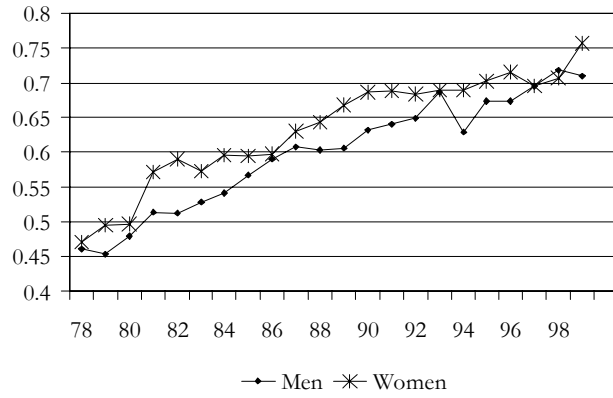


Figure 1: Interquartile range for male and female hourly log wages

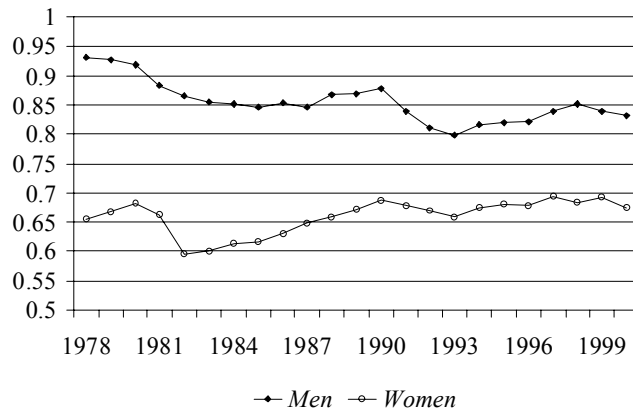


Figure 2: Labour Market Participation

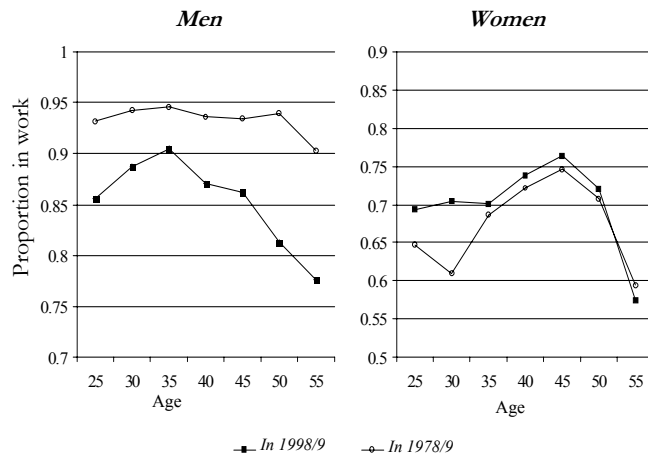


Figure 3: Participation rates by age for 1978 and for 1998

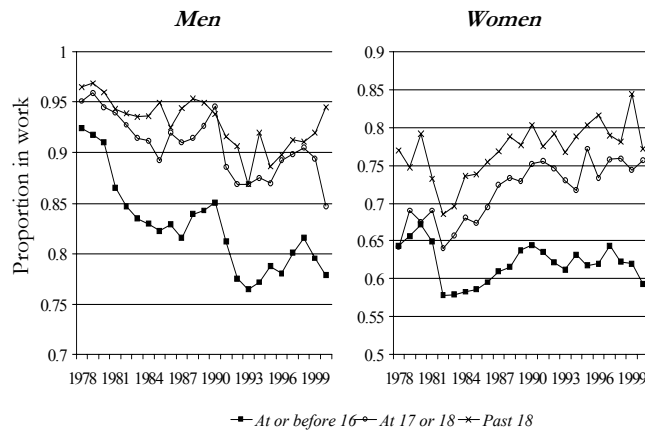


Figure 4: Participation rates by education level

employment of women with more than the lowest level of education.

To the extent that changes in participation are related to wages, part of the changes in the distribution may have been induced by changes in the composition of those employed in terms of their unobserved characteristics. Moreover, since the changes in participation are quite complex it is not at all clear how the composition of those employed has changed in terms of unobservables.

This point has a bearing on the interpretation we give to 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. In addition it is not obvious how composition is changing. For some (perhaps for young men), the dominant factor in the participation decision will be the hourly wage rate which is driven both by observed and unobserved productivity related characteristics. For others, one of the crucial factors will be accumulated wealth (say for those close to retirement). The combination of these two factors implies that we are not necessarily able to state whether it is predominantly the less or the more productive who have been leaving the labour force.

At least for analysing means, this selection problem is well understood in the literature and it is usually dealt with either in a parametric or in a nonparameteric framework (with or without exclusion restrictions, see Heckman, 1979 Heckman and Sedlacek etc.). A recent example which shows how important such selections issues can be is the paper by Blundell, Reed and Stoker (2003). They use changes in the policy framework for out of work benefits, exploiting the differences generated by differences 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 when we account for selection effects.

Sometimes the assumptions required to correct for selection may not be considered satisfactory even more so when we are dealing with the entire distribution of wages. We often require a combination of exclusion restrictions and distributional assumptions or assumptions allowing identification based on individuals with probability of participating close to 1 (identification at infinity). These may reduce the credibility of the results. Manski (1989)

introduced the idea of using bounds in order to analyse selection problems in an “assumption free” context. However, it often proved to be the case that the worst case bounds were too uninformative to be useful. The use of Frechet bounds by Heckman, LaLonde and Smith (1997) for joint distributions have underscored this point.

Our aim in this paper is to analyse the changes in the distribution of wages, with particular emphasis on the returns to education and on male/female wage differentials. The data source we use is the UK family Expenditure Survey from 1978 to 2002. We develop a bounds framework in which we progressively add stronger restrictions motivated directly from economic theory. We show how to derive bounds under these assumptions and we then use the methods for our empirical exercise. All results are compared to the worst case (“assumption free”) bounds. In addition for some cases we develop a way of testing whether there is evidence that our assumptions are violated. In this way we can construct the bounds that are most credible and at the same time as informative as possible. In many instances the bounds lead to actual point estimates.

In terms of substantive results, we show that inequality must have increased over this time period and that composition changes are not capable of accounting for the observed changes in the UK. We also find that wage differentials have improved for women but not for all age and education groups. Educational differentials have increased for men quite substantially. Finally, our framework confirms that wages grow over most of the lifecycle, particularly for the highly educated. Thus wages do not decline for older people.

Our paper builds on the growing literature on bounds. Apart from the references already mentioned important papers include Manski (1990) who bounds treatment effects both in general and under the assumption that they do not depend on certain observables. Manski (1994) and Manski and Pepper (2000) who introduce the idea of tightening the bounds for treatment effects using either exclusion restrictions or monotone instrumental variables; finally Heckman and Vytlacil (2001) who use index type restrictions together with exclusion restrictions to tighten the bounds on treatment effects. Our paper uses and extends some of these ideas to bounding any quantile of a distribution. We do not attempt to bound means since we do not wish to restrict the range (support) of wages

2 Worst case bounds and Bounds with restrictions.

In this section we describe our approach in general terms. This has applicability to all selection problems, although the focus of the current paper is on understanding the evolution of the wage distribution and how measures of interest are affected by inequality. As can be expected the interpretation of our results depends on the underlying model of wage determination and we discuss this later in the paper.

The approach that is often followed in the literature is to use a parametric selection model (usually for the mean). This depends on distributional assumptions and/or exclusion restrictions. In practice these are often hard to justify from economic theory. The aim of this paper is to evaluate what can be said without any (or only minimal) restrictions.

Let Y and X denote the dependent variable and the conditioning variable, respectively. In our case the dependent variable should be taken to be the log hourly wage and X should be understood to include gender, age, education and possibly other observable characteristics. When a realisation of $Y = y$ is observed, the indicator variable I equals 1 and when y is not observed, I equals 0. In our case I indicates whether the person is working or not. The probability of $I = 1$ given x is written as $P(X)$. In our analysis this is the participation probability for individuals with characteristics x . We write the conditional distribution of Y given X by $F(Y|X)$, the conditional distribution of Y given X and $I = 1$ by $F(Y|X, I = 1)$, and the conditional distribution of Y given X and $I = 0$ by $F(Y|X, I = 0)$. While $F(Y|X)$, the object of interest, is not observed (because of non-random selection into work) we can write

$$F(Y|X) = F(Y|X, I = 1) P(X) + F(Y|X, I = 0) [1 - P(X)] \quad (1)$$

Given that the data identify $F(Y|X, I = 1)$ and $P(X)$ the problem can be respecified as one in which only the distribution of wages among those not working, $F(Y|X, I = 0)$, is unknown. This should be understood as the distribution of wages rejected by those not taking up employment. The distribution $F(Y|X)$ is an equilibrium distribution at a point in time and would change if a large proportion of those out of work decided to work. Since this is not conditioned on participation for ease of exposition we refer to this distribution as

the unconditional distribution of wages.

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

$$0 \leq F(Y|X, I = 0) \leq 1.$$

is substituted into equation (1), the bounds to the cumulative distribution function can be derived as in equation (2) below²

$$F(Y|X, I = 1) P(x) \leq F(Y|X) \leq F(Y|X, I = 1) P(X) + [1 - P(X)]. \quad (2)$$

As cumulative distribution functions are monotonic the bounds to the conditional quantiles can then be easily estimated (see below).

The basic idea underlying these worst case bounds is quite simple: it must be the case that assuming that all non-workers are less productive than workers will identify the lowest possible value that the quantile of the unconditional distribution $F(Y|X)$ could take. Assuming the opposite would identify the highest possible value that the unknown quantiles could ever take. The truth must either lie at one of these two extremes or somewhere in the middle.

Finally, note that the proportion of individuals in work, $P(X)$ will determine which quantiles can be estimated. Upper bound quantiles can only be identified up to $P(X)$ and lower bound quantiles can only be identified down to $1 - P(X)$. For women, during most of our time period female participation was below 60% and so in practice we can really only estimate the bounds around the median.

²Proposition 3, p.152.

2.1 Estimating bounds to within group inequality

Our measure of inequality will be mainly the interquartile range. In general, we wish to bound differences between any two quantiles conditional on X , $(y^{q_2}|_x)$, such as

$$D(X) = y^{q_2}|_x - y^{q_1}|_x. \quad (3)$$

In doing this we can exploit the fact that distribution functions are monotonic. In particular we need to ensure that the bounds for $D(X)$ are consistent with a monotonically increasing distribution of wages for non-workers (which we do not observe). To see how this affects the bounds to $D(X)$ first note that

$$F(Y|X, I = 0) = \frac{F(Y|X) - P(X)F(Y|X, I = 1)}{P(X)}$$

Assuming differentiability for simplicity, to ensure that $\frac{\partial F(Y|X, I=0)}{\partial y} > 0$ the above expression implies that

$$\frac{\partial F(Y|X)}{\partial y} > P(X) \frac{\partial F(Y|X, I = 1)}{\partial y}$$

Now consider the upper bound of $D(X)$, say $D^u(X)$. This will be $\max D(X)$ such that a. $D(X) \geq 0$, $y^{q_2}|_x$ and $y^{q_1}|_x$ lie within their respective bounds and such that the pair of $y^{q_2}|_x$ and $y^{q_1}|_x$ lie on a curve with the same gradient as $P(X) \frac{\partial F(Y|X, I=1)}{\partial y}$ or steeper. Similarly for the lower bound.

2.2 Imposing restrictions to tighten the bounds

With the labour market participation rates in the range we observe them in our data, worst case bounds can be informative about certain aspects of wages, such as life-cycle growth. However, in many cases they are uninformative other than for groups where the participation probability $P(X)$ is sufficiently high. In addition the worst case bounds cannot in themselves reveal the impact of selection on the observed quantities. They can only show the extent to

which selection can affect the results. To measure the extent to which selection is biasing the results we need further structure, which we will introduce below. One may take the view that this is what the data tells us and that's that. Alternatively, we can fruitfully combine theoretical ideas with information from the data to arrive at firmer (albeit conditional) conclusions. Some of these restrictions will have testable implications, but others will not.

2.2.1 The median restriction

The first assumption we consider reflects the idea that the probability of someone working is higher the higher their hourly wage. We model this positive selection into work by assuming that the median wage of the distribution of hourly wages for workers is higher than the median wage of the distribution for non-workers. This is equivalent to assuming that the probability of working is higher if ones wage is above the median of the unconditional distribution $F(Y|X)$ than it is if ones wage is below the median.

More generally, we consider restrictions that change the bounding function specified in equation (2) to the form

$$LB = F(Y|X, I = 1) P(X) + G(Y|X) [1 - P(X)].$$

where $G(Y|X)$ is an (incomplete) conditional CDF. Denoting the median wage of the working population conditional on observables X by $q_{0.5}^1(X)$, then the lower bound function for this particular case becomes

$$G(Y|X) = \begin{cases} 0 & \text{if } y < q_{0.5}^1(x), \\ 0.5 & \text{if } y \geq q_{0.5}^1(x). \end{cases}$$

Thus the lower bound function to the distribution $F(Y|X)$ remains unchanged up until the median wage for workers. At that point the lower bound improves to become $F(Y|X, I = 1) P(X) + 0.5 [1 - P(X)]$. Hence, the new restriction is informative about any quantile above $0.5P(X)$.

The median restriction is an assumption about positive selection into the labour market.

Although one may be used to expressing such assumptions with respect to means, in the context of censoring this is not very helpful since means are unbounded, unless we restrict the support of the distribution of wages.

To gain some further insight about this assumption we can rewrite the observed distribution functions as follows

$$F(Y|X, I = 1) = \frac{\Pr(I = 1|x, Y < y)}{P(X)} F(Y|X)$$

Thus $F(Y|X, I = 1) \leq F(Y|X)$ implies

$$\Pr(I = 1|x, Y < y) \leq P(X)$$

If workers adopt a reservation wage strategy and move into work once their wage moves above a certain threshold, then holding everything else constant, increases in wages will result in increases in the probability of work. As the algebra above shows, this will be consistent with the median restriction.

There are various reasons why the assumption of positive selection could be false. First, if wages are correlated over time, higher waged people are more likely to have accumulated assets which may increase their reservation wages. Second if high waged women are matched with high waged men, then the out of work income of women could be increasing in their potential earnings. Lastly as higher waged women tend to delay rather than avoid childbirth, it could be the higher waged women in older age groups who have pre-school children. We thus suspect that the median restriction could be less plausible for women and for those over 50. In the empirical section we present some circumstantial evidence that strongly supports the median restriction. However this restriction is not directly testable. The remaining restrictions we introduce do however have testable implications.

2.2.2 Using determinants of participation to tighten the bounds.

This next sections examines how economic restrictions about the relationships between wages and another set of variables (Z) can be used to obtain tighter bounds to the estimated quantiles. One example of this would be an exclusion restriction. We show, however, that other weaker assumptions may also be employed. We also show how these restrictions can be rejected empirically.

2.2.3 An exclusion restriction.

Manski (1994) shows that if y is independent of Z conditional on x i.e.

$$F(Y|X, Z) = F(Y|X) \forall y, x, Z$$

then

$$\max_Z q^l(x, Z) \leq q(X) \leq \min_Z q^u(x, Z)$$

where $q(X)$ denotes the q^{th} quantile of wages conditional on x , and q^l denotes the lower bound to the q^{th} quantile and q^u the upper bound; in all cases these are conditional on both x and Z (the instrument). Implementation is thus simply a matter of finding the bounds to the quantiles of wages conditional on x and Z and then searching across Z within x to find the lowest upper bound and the highest lower bound.

Tightening the bounds requires that the instrument has an effect on the probability of participation, which is the usual rank condition of identification.

It might be thought that the tightest bounds will be found at the value of Z where the proportion of those with $I = 1$ is highest. This is not necessarily the case. To illustrate this note that we can rewrite the lower bound to the estimated distribution function ($F(y|I = 1, X, Z) \Pr(I = 1|x, Z)$) when y is independent of Z as

$$F(Y|X) \Pr(I = 1|Y < y, x, Z)$$

where the second term is the probability of participation for those with wages below some threshold y . Take two values of instrument $Z = z_1$ and $Z = z_2$ such that $\Pr(I = 1|x, Z = z_2) > \Pr(I = 1|x, Z = z_1)$. The difference in the lower bounds at these two points will be

$$F(Y|X) [\Pr(I = 1|Y < y, x, Z = z_2) - \Pr(I = 1|Y < y, x, Z = z_1)]$$

This will be positive when $\Pr(I = 1|Y < y, x, Z = z_2) > \Pr(I = 1|Y < y, x, Z = z_1)$. However $\Pr(I = 1|x, Z = z_2) > \Pr(I = 1|x, Z = z_1)$ is neither a necessary or a sufficient condition for this to be true.

There is nothing in the definition of the bounds with an exclusion restriction that forces the min of the upper bounds to lie above the max of the lower bounds. Thus, in cases where y is not in fact independent of Z it is perfectly possible for the bounds to cross and the upper bound for some values of Z to lie below the lower bound for others. Later we construct a test of the null hypothesis that the bounds do not cross. If we reject the hypothesis this is evidence against the exclusion restriction.

2.2.4 Weakening the exclusion restriction: Monotonicity.

Strong exclusion restrictions of the type discussed above may not always be credible. We might, however, be prepared to assume the direction of the relationship between y and Z . This idea is very similar to the monotone instrument variable (MIV) for the mean proposed by Manski and Pepper (2000) and the following analyses borrows from their exposition. The restriction then becomes:

$$F(Y|X, Z \geq z_1) \leq F(Y|X, Z \leq z_1) \quad \forall y, x \tag{4}$$

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 . To exploit this we can find tightest bounds over the support of Z and then integrate out Z . To illustrate suppose Z is discrete. For a value of $Z = z_1$ the best lower bound is the largest lower bound

over the values of Z such that $Z \geq z_1$. This is given by

$$F(Y|X, Z = z_1) \geq F^l(Y|X, Z = z_1) \equiv \max_{Z > z_1} \{F(Y|X, Z, I = 1)P(x, Z)\}.$$

Similarly we can obtain a best upper bound at $Z = z_1$ by choosing the smallest possible upper bound over the support of $Z \leq z_1$, i.e.

$$F(Y|X, Z = z_1) \leq F^u(Y|X, Z = z_1) \equiv \min_{Z < z_1} F(Y|X, Z, I = 1)P(x, Z) + 1 - P(x, Z)$$

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

$$E [F^l(Y|X, z_1)|x] \leq F(Y|X) \leq E [F^u(Y|X, z_1)|x]$$

As with the exclusion restriction here again there are conditions that ensure the bounds we obtain with monotonicity will be tighter than the worst case bounds.

Bounds will be tighter if either the lower bound becomes higher or the upper bound lower. Since we can never do worse than the worst case bounds, the lower bound will improve if we can get an improvement at least at any one value of the instrument. To see what this entails for a particular point of the wage distribution y , take two values of the instrument $Z = z_1$, $Z = z_2$ ($z_1 < z_2$). If the lower bound is to be better than the worst case bounds it must be that

$$F(y|I = 1, X, Z = z_1) \Pr(I = 1|x, Z = z_1) \leq F(y|I = 1, X, Z = z_2) \Pr(I = 1|x, Z = z_2).$$

Define $F_i = F(Y|X, Z = z_i)$ and $P_i^c = \Pr(I = 1|Y < y, x, Z = z_i)$. Then the condition for improving the bound requires that

$$\frac{F_1}{F_2} \leq \frac{P_2^c}{P_1^c}$$

Note that under the monotonicity assumption $\frac{F_1}{F_2} \geq 1$, which means that if the lower bound

is to increase, a higher value of the instrument will have to be associated with a higher probability of participation of those with wage below y . More specifically the effect of changing Z from z_1 to z_2 has to increase the probability of participating of those with wages below y more than it decreases the probability of having a wage below y . In terms of observable quantities the conditions requires that

$$\frac{F_1^c}{F_2^c} \leq \frac{P_2}{P_1}$$

where $F_i^c = F(y|I = 1, X, Z = z_i)$ and $P_i = \Pr(I = 1|x, Z = z_i)$.

As we discuss below our instrument is out of work income. An increase of the instrument such as ours is likely to reduce participation, not increase it; and since our monotonicity assumption states that individuals who can expect higher out of work income are likely to be higher wage individuals the instrument we use is unlikely to improve the lower bound.

An improvement in the upper bound will happen if for some $z_2 \geq z_1$ we get that

$$F_1 P_2^c \left[\frac{F_2}{F_1} - \frac{P_1^c}{P_2^c} \right] \leq P_2 - P_1$$

where the notation is the same as above. If increases in the instrument are associated with decreases in the probability of participation then $P_2 - P_1 \leq 0$, which then requires that $\left[\frac{F_2}{F_1} - \frac{P_1^c}{P_2^c} \right]$ the effect of increasing the instrument decreases the participation of those with wages below y by sufficiently more than it decreases the proportion of those with wages below y . Thus the monotonicity restriction may improve either the upper bound or the lower bound or both. The condition has an observable counterpart which is

$$F_1^c P_2 \left[\frac{F_2^c}{F_1^c} - \frac{P_1}{P_2} \right] \leq P_2 - P_1$$

and can be checked in the data. The bounds can be tightened further by combining them with the median restriction.

Finally in the appendix we also present ways of tightening the bounds by assuming the

statistical independence of some of the observables in the wage equation with the unobservables. This is interesting because such independence assumptions are often made either implicitly or explicitly in the empirical literature. However, we do not pursue this here in the empirical analysis because this assumption is not well motivated from economic theory and in most cases it can easily be rejected, because it leads to the bounds crossing.

2.2.5 The relationship of our restrictions to those of the parametric approach

The median restriction we impose leads to a model that does not nest the standard selection model because it imposes some form of positive selection which is not implied by the selection model. In this sense the results are not directly comparable. The model with the exclusion restrictions however nests directly any nonparametric selection model that uses such restrictions. It is more general in the sense that it admits bounds whenever point estimates are not obtainable. Finally the model that exploits the monotonicity restriction also nests all models that use exclusions. In addition, we do not impose any functional form restrictions for the wage equation. Thus our estimates are more robust than the ones based on the standard approaches, at least when we do not use the median restriction.

3 Estimation and Inference

3.0.6 Estimation

Although much of the discussion has been in terms of bounding the distribution of wages in practice we will be working with the quantiles of the distribution. Thus we are interested in estimating the bounds to these. The upper bound to the distribution provides the lower bound to a quantile of interest and vice versa.

To estimate the worst case bounds for the q th quantile of the wage distribution, conditional on x we need to estimate the probability of work, conditional on X ($P(X) \equiv \Pr(I = 1|x)$). Given this we can use the following relationship to estimate the upper and lower

bounds using the formula below:

$$F^{-1}\left(\frac{q-1+P(X)}{P(X)}|I=1, X\right) \leq F^{-1}(q|X) \leq F^{-1}\left(\frac{q}{P(X)}|I=1, X\right). \quad (5)$$

Thus the estimation of the lower and upper bound to the q th quantile is a quantile regression problem with suitably defined quantiles. Note by the way that we can only bound quantiles such that $q > (1 - P(X))$. Thus we can only bound the 25th percentile if at least 75% of individuals are working.

Generally, the main difficulty, if one is to remain non-parametric³, is the dimensionality of x . In our case we take X to be discrete and it includes, gender, age, year and education. The principal practical concern is the cell size, particularly for higher education groups among older individuals.

To avoid problems with small cell sizes we use seven age groups (23-27, 28-32, 53.-57).⁴ We label the age groups 25, 30, 35 etc. We define three education groups based on the age individuals reported leaving full time education (16 or less - the unskilled, 17 or 18 - the high school graduates, above 18 - the College Graduates). Finally we use ten year groups (78-79..98-99). The probability of work $P(X)$ was estimated as the (weighted) proportion in work for each cell. We then use these probabilities to estimate the bounds of quantiles in each cell separately.

Estimation using the median restriction. When we impose the median restriction the lower bounds to the quantile are not affected. The upper bound of all quantiles of the unconditional distribution above and including the $0.5P(X)$ decline. To give further insight note that the median of the observed distribution becomes the upper bound of the median to the unconditional distribution, which reflects the assumption that some positive selection is imposed.

³Note that even if the underlying relationship between the quantiles of wages and X were linear, there is nothing to say that the bounds should have the same functional form, thus we need to be as nonparametric as possible in our estimation procedure.

⁴As these age groups are quite broad, we assigned a weight for each age group to reach a maximum for workers in the middle of the age category (25,30...). Effectively, then, this procedure is equivalent to a non-parametric regression with a bandwidth of 3 and triangular non overlapping kernels.

Thus to estimate the constrained upper bound for all quantiles $q \geq 0.5P(X)$ we look for the $\frac{q-0.5(1-P(X))}{P(X)}$ quantile of the observed distribution in each data cell defined by a value of $X = x$.

Estimation under the exclusion restriction. To estimate the bounds under the exclusion restriction we first need to estimate the bounds conditional on the instrument z , in the same way that was described for the worst case bounds above. Although our instrument is effectively continuous we discretise it. Following this we estimate the upper and lower bounds for each discrete value of the instrument. The estimated upper bound then is the minimal upper bound over all values of Z , while the lower bound is the maximal lower bound over Z . This procedure is carried out separately in each cell defined by the conditioning variables x .

Estimating bounds under the exclusion and the monotonicity restrictions. When using the monotonicity restriction we need to go through a few more steps. Lets normalise the instrument so that an increase in its value implies no decrease in any quantile. Then at each value of Z estimate the lower and upper worst case bound. Following this at each point of z_1 we search the upper and lower bounds for values of $Z \geq z_1$ and pick the minimal and maximal values respectively. This provides us with the best bounds conditional on each value of Z .

The bounds to the distribution of wages conditional only on $X = x$ are then estimated in the following way. First find for each value of Z the upper ($UB_{z_i|x}^y$) and lower ($LB_{z_i|x}^y$) bounds that are consistent with the assumption that $F(Y|X, Z)$ is never increasing in Z and then average over the sample to get bounds to $F(Y|X)$

$$\sum_{i=1}^N \left[LB_{z_i|x}^y \Pr(Z = z_i|x) \right] \leq F(Y|X) \leq \sum_{i=1}^N \left[UB_{z_i|x}^y \Pr(Z = z_i|x) \right]$$

Once upper and lower bounds for the distribution function conditional on only X are obtained, it is then easy to use it to estimate various quantiles.

Combining restrictions. We will be presenting results that combine the median restriction with monotonicity. In this case, instead of estimating worst case bounds conditional on Z we will be estimating bounds that respect the median restriction at each Z . Given these the remaining steps used to impose monotonicity and average over all bounds on the support of Z are identical.

3.0.7 Confidence intervals for bounds.

There are different ways to construct a confidence interval. What we need is an interval I_α such that

$$\Pr \{ [w_{low}^q, w_{up}^q] \in I_\alpha \} \leq \alpha$$

but there are many such intervals. There are at least three different criteria we could use: Equal distance interval, equal probability interval, and the shortest distance interval. In the standard normal theory with a point estimate, these three concepts coincide but in our case they do not. In particular, the asymptotic variances of the two bounds are different and hence the equal distance approach produces wide interval with much smaller coverage probability than the nominal probability α . Below we construct equal probability intervals.

To construct our interval we repeated our estimation procedure on 200 bootstrapped samples of the data⁵ We then defined the 95% confidence interval as that which we obtained after the following steps:

- Define $w_{up}^q[200]$ as the maximum of the upper bound and $w_{low}^q[1]$ as the minimum of the lower bound in a given cell across replications
- Take the range defined by the pair $w_{up}^q[199]$ and $w_{low}^q[2]$ and find the proportion of replications in which both the upper and lower bounds lie within it
- If this proportion is greater than 0.95 then go on to the range defined by the pair $w_{up}^q[198]$ and $w_{low}^q[3]$, repeat until the proportion goes under 0.95.

⁵In fact when we bootstrapped the results for overall inequality there were 1000 replications but the estimated confidence intervals did not differ that much when the smaller number of replications was considered.

For differentials which are obtained from values which are not independent (say a within group differential) we have to conduct this analysis on the differential rather than the value of each quantile

This procedure obtains tighter estimates than if we started at the point estimates and then added (subtracted) a fixed number from the upper (lower) bounds. This is because the precision with which the upper bound is estimated is not the same as that with which the lower bound is estimated, particularly for quantiles other than the median.

3.0.8 Testing that bounds do not cross

When we impose the monotonicity or the exclusion restrictions bounds may cross if the restrictions are invalid. Of course in any finite sample they may also cross even if the restrictions are valid and hence a crossing should not be immediately interpreted as a rejection of the restriction. Thus we have developed a statistical test for the hypothesis that they do not cross. We construct the test in the following way

4 Empirical 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 education⁶ This gave us a sample of 187,467 individuals in total. We defined individuals to be in “work” (i.e. $I = 1$) if they were either employed, whether full time or part time or were self employed over the last week⁷. Hourly wages were defined as usual weekly earnings divided by usual

⁶In the UK university education is completed for most students by the age of 22

⁷We treated the self-employed as workers because excluding them from the analysis would result in another source of selection bias. We believe that the assumption that self-employment is just another form of work, albeit one with a different tax treatment is an acceptable one. We do however treat the wages of the self employed as missing at random.

weekly hours (inclusive of overtime) and were deflated by the consumer all items quarterly retail price index .

Constructing out of work income Apart from the median restriction we will also use the exclusion restriction and the monotonicity restriction. This requires an instrument which we choose to be potential (simulated) income out of work. i.e. the income that the individual would have if she/he were to be out of work (see Blundell, Reed and Stoker (2003)).

5 Results

We now present results in turn for changes in inequality, for wage growth over the life-cycle and across cohorts, and for changes between groups, in particular between education groups and between men and women. In all our results we will present estimates based on different assumptions, starting from the worst case bounds. But before we do this we present some evidence on the plausibility of our assumption on positive selection (the median restriction) and a set of hypothesis tests relating to whether the bounds cross or not when we impose monotonicity or exclusion restrictions.

5.0.1 The plausibility of 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.

We argued that this restriction may be particularly problematic for older individuals where the wealth effect may dominate the substitution effect in the decision to continue working. It may also be suspect for women because high productivity women are likely to be married with high productivity men. These same women may choose not to work through a wealth effect.

The graph below (figure 5) uses longitudinal data from the British Household Panel Survey 1991-9 (BHPS) to show, that such effects may not be that important. Here we

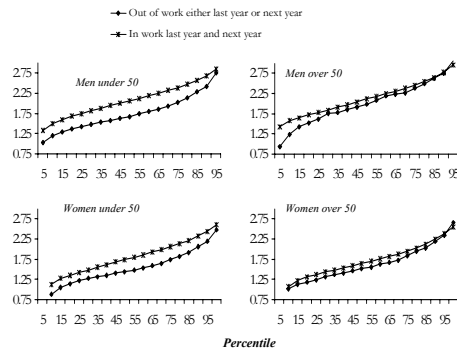


Figure 5: Percentiles of residual wages for workers with and without unemployment spells - Source British Household Panel Survey.

regressed log wages for each year of the panel 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 unemployed either last year or will be unemployed next year and into those who were not. Figure 5 shows that the distribution of residual wages of those in work lies below the distribution of those who have been out of work, even controlling for factors such as age and education which are important determinants of unemployment. Hence the median restriction we impose may not be that unreasonable. In fact this graph, taken at face value would provide support for the stochastic dominance assumption.

5.0.2 Testing the validity of the restrictions.

The exclusion restriction and the monotonicity restriction can, if invalid, lead to the bounds crossing. We thus test the null hypothesis that they do not cross. If we reject, this is evidence against the restriction. However if the bounds do not cross it is still possible that the underlying restrictions are invalid. Thus although the test is powerful with respect to the alternative that the bounds cross, it is not necessarily powerful with respect to the restriction we impose.

Table ?? includes the p-values of the test against a number of hypotheses for the three education groups and for men and women. We consider the monotonicity restriction on its own as well as combined with the median restriction. We then consider the exclusion

restriction also on its own and combined with the median. As far as women are concerned there is no evidence of the bounds crossing for any of the education groups. For men the monotonicity restriction does not lead to bounds crossing even when combined with the median restriction. However the exclusion restriction does lead to significant bound crossing for all but the highest education group. When combined with the median we also get a rejection for the highest education group.

Thus in summary, while the monotonicity restriction does not lead to bounds crossing the exclusion restriction is not acceptable for men. We present results with a number of alternative restrictions. However these test results need to be kept in mind when interpreting the results.

5.1 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 * thus plots the upper and lower bound to the 75th-25th percentile differential from 1978 to 2000 for the male wage distribution. Since the participation rate for women is often quite low we cannot produce the same figure for them

The central line shows, for comparison, what has happened to wage inequality amongst workers and the dotted lines give 95% confidence intervals for the upper and lower bounds (see below). 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 higher bound at the beginning of the period. Figure 1 thus shows strong evidence of an increase in inequality. The lower bound in 1998 is higher than the highest bound in 1978, suggesting that inequality must have risen and a comparison of the limits of the bootstrapped confidence intervals suggest that this change is significant implying that these bounds are quite precisely estimated and that there is less than a 5% chance that inequality was the same in 1999 as in 1978.⁸

The bounds above were estimated by placing no restrictions on the data. We have shown earlier that imposing restrictions on the relationship of the wage distribution for workers and

⁸see the appendix below for the description of the bootstrapping procedure

non-workers can lead to tighter bounds. Thus we now impose the restriction that the median wage of workers is at least as high as the median wage for those out of work at any point in time and for any group defined by x . We then follow this up by deriving bounds assuming stochastic dominance, i.e. that all quantiles of the distribution of wages for non-workers are lower than those of workers. As described earlier panel data suggested these assumptions were reasonable overall.

The first panel of figure ** is simply the earlier bounds shown in figure *. The next panel shows the bounds that result from imposing the median restriction; the third panel shows the bounds obtained by exploiting the assumption of stochastic dominance. The 95% confidence intervals for these bounds are reported in Figure ***. The results are particularly strong when we impose stochastic dominance. In this case we can establish that inequality rose *vis a vis* 1978 in all periods after 1986. Taken together, these results give strong evidence for an increase in overall inequality between 1978 and 1999, and that the observed changes cannot be just an artifact of selection.

5.1.1 Within group inequality.

A feature of the increase in inequality in Britain (as well as the US) has been the large increase in within group inequality. Looking at the point estimates, as Figures * and * show, we can confirm that this increase cannot have been driven by selection. Looking first at figure *, which documents the bounds to the changes in inequality (the 75th-25th differential) for those men leaving school at or before 16, inequality has risen for some if not all age groups. The first panel shows that even when we place no restrictions on the data, inequality amongst younger ages has risen, the next two panels suggest that if we make the assumption that those in work have higher potential wages than those out of work, inequality must have risen for all age groups. Figure * shows the changes for the higher education group (those leaving school after 18). Note first that the bounds to the changes are much sharper as employment rates are much higher for this group and the observed changes have been relatively smaller. For this group, there is relatively little qualitative gain in making either the median restriction or assuming stochastic dominance, all panels show inequality rising amongst younger cohorts,

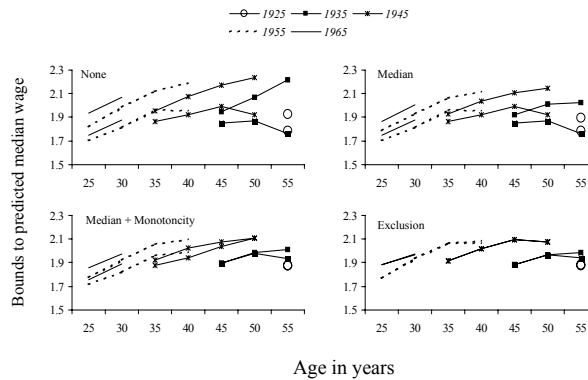


Figure 6: Bounds to median wages over the life-cycle - Unskilled

little changes amongst those in their 40s and rising for older workers.

The picture of an unambiguous increase in within group inequality becomes less clear when we look at the bootstrapped confidence intervals (shown in figure *). The point estimates suggest that the observed increase cannot be driven by selection but the bootstrapped confidence intervals show that any changes are not significant at the 5% level. For 30 year olds, they are, however, significant at the 20% level for the lower educated group and at 6% for the higher educated group. We are thus relatively confident that our data do show increases in within group inequality.

5.2 The determination of median wages.

5.2.1 Wage growth over the life cycle and across cohorts.

Before moving onto the way that within group inequality and education/gender wage differentials have evolved we provide an account of how wages have grown over cohorts and age.

Figures 6 and 7 show bounds for the median wage by cohort and age for unskilled men and College graduates respectively. Each pair of lines shows the bounds to the median wage for a cohort across different ages. Each of the four panels shows the results under different

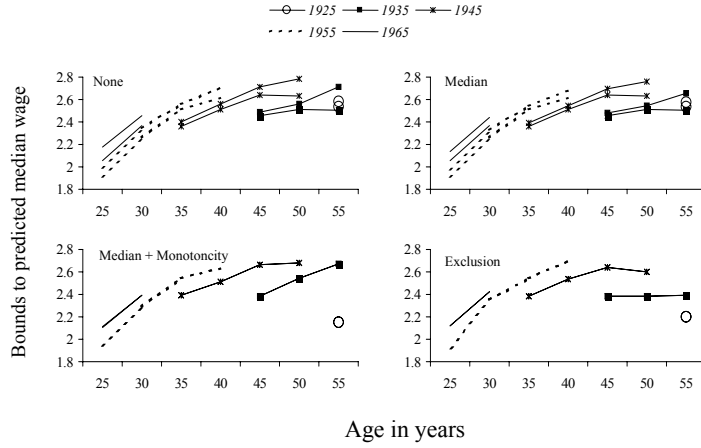


Figure 7: Bounds to median wages over the lifecycle - Men College Graduates

restrictions.

For unskilled men who completed their full time education by the age of 16 the bounds are wider as we would expect given the lower participation rates. When we consider the worst case bounds, we can only be sure that there has been lifecycle wage growth for 1955 cohort. When we impose the median restriction we can conclude also that wage have grown with age for the youngest 1965 cohort. Adding the monotonicity restriction leads to the implication that wages grow over the lifecycle for all cohorts. The result is similar with the exclusion restriction. However we should bear in mind that this is rejected.

Turning now to growth across cohorts we can obtain the minimal growth by comparing the upper bound to the median wage of an older cohort to the lower bound of the median for a younger one at a particular age.

For the unskilled group the picture is quite mixed when we just look at the worst case bounds, but mostly we can conclude that wages do not decline across cohorts. However, with the median and the monotonicity restriction we can see clear growth between the 1935 and the 1945 cohort and at least no decline between the 1945 and 1955 cohorts; however we cannot be certain that such growth continued between the 1955 and 1965 cohorts, unless we are willing to believe in the exclusion restrictions.

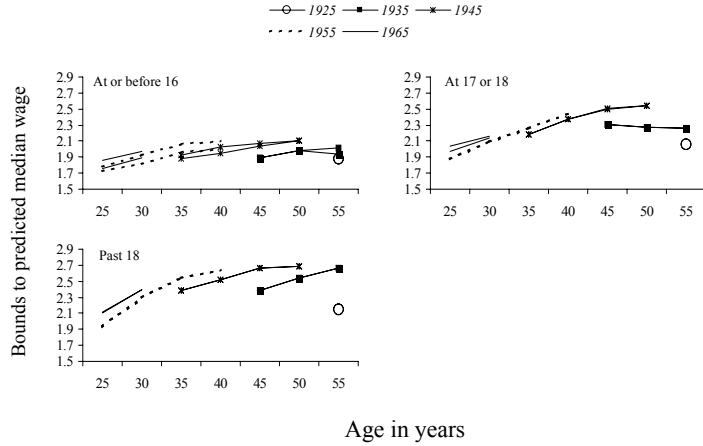


Figure 8: Life-cycle profiles by education group men. Monotonicity and median restriction

For the College graduates all bounds are very tight and show unambiguously growth of wages over the lifecycle. The bounds are wider for the older individuals of the older cohorts reflecting the drop in participation. Nevertheless, even there, their maximum width is no more than 0.2 (20%). The worst case bounds and the ones with the median restriction indicate that wages for the 1935 cohort may have flattened out after 45 years of age. Imposing the median restriction we also see growth there. With the median and the monotonicity restriction we actually obtain point estimates and they all clearly show growth, even for the oldest group. The exclusion restriction on its own, without the median restriction leads to completely flat wages. we should bear in mind though that the exclusion restriction is rejected.

The other characteristic of these results is that they show unambiguously growth of wages across cohort. This is true even with the worst case bounds, but under the median and the monotonicity restriction we can get point estimates of this growth. Such cohort growth underlies partly the increase in inequality as illustrated by Gosling, Machin and Meghir (2000).

To obtain a summary picture for men based on the monotonicity restriction combined with the median we put together in figure 8 the previous results and now include the middle

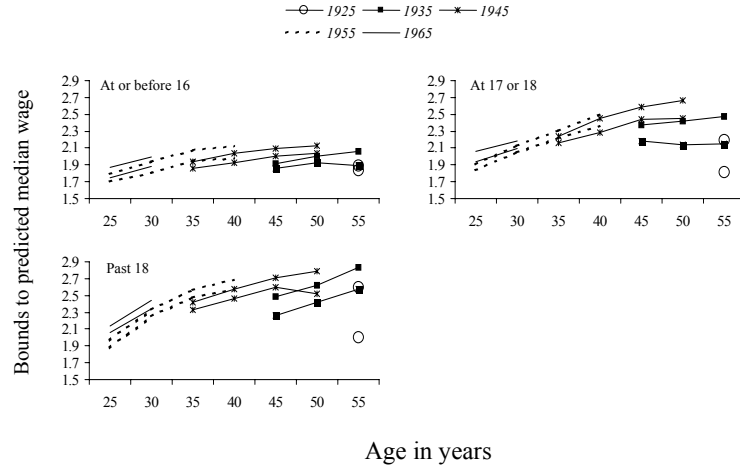


Figure 9:

education group (high school graduates) The results are based on the monotonicity restriction combined with the median. From this we can clearly see that while all education groups enjoy lifecycle growth of wages this growth is higher for the higher education groups. Moreover in all cases there is growth (or no decline) of wages across cohorts. Finally in figure 9 we present 95% pointwise confidence intervals for the graphs in figure 8.

In figure ?? we present a similar figure for women followed by the confidence intervals in figure ?? For the unskilled women (whose employment has declined considerably recently) we cannot say much either about life-cycle growth or about inter cohort growth when we consider the 1965 and 1955 cohorts. For the older cohorts we observe low but positive growth for over age and cohorts. For the other two education groups growth s unambiguously positive and higher for the higher education group. The results are however less precise than those for men and some of the growth figures are not statistically significant (although most are).

5.2.2 Returns to education

We now consider the returns to education and the extent to which they have changed. To obtain a return to education we compare median wages of the highest ($ed = 1$) to the

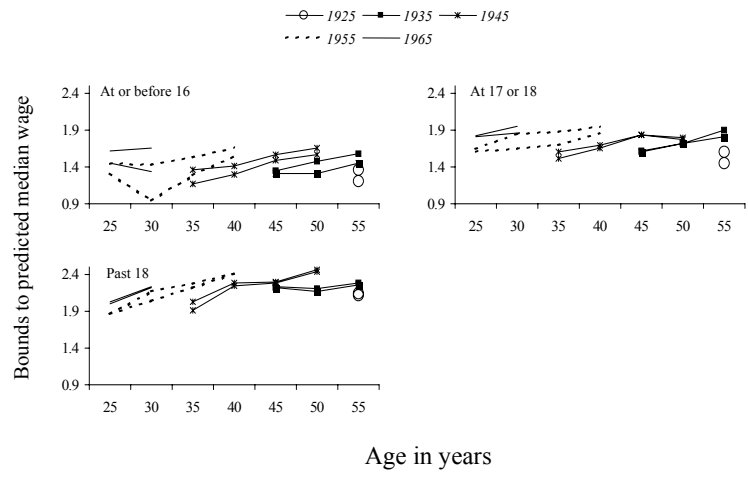


Figure 10:

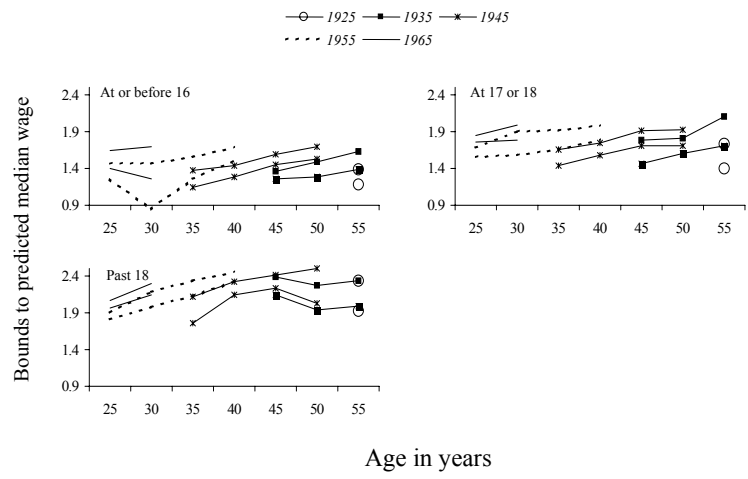


Figure 11:

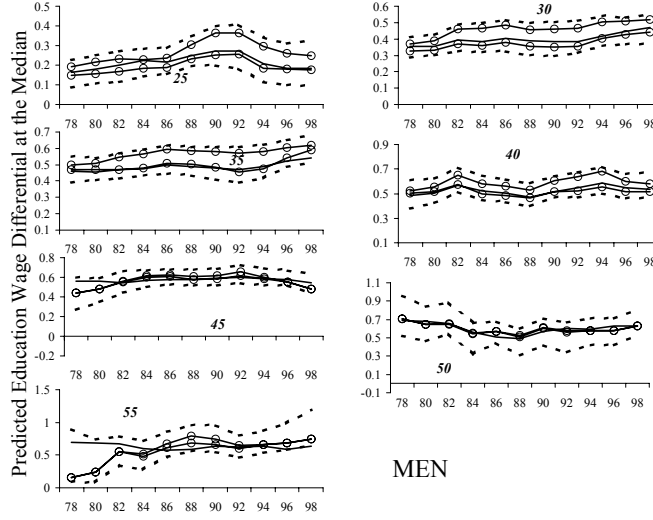


Figure 12: Returns to College for men by age and year - Monotonicity and Median restriction lowest ($ed = 0$) education group. The usual measure based on comparing the means of log wages is not available since we do not bound means. Moreover we should point out that the underlying model of wages may be non-separable in observables and unobservables. This would give rise to a distribution of returns to education. Bounding such a distribution is beyond our scope here.

When we examine differentials across groups we have to allow for the fact that workers may be positively selected in one group and negatively selected in another. Thus the bounds to wage differentials (D) are set as :

$$w_{low}^q(x, ed = 1) - w_{up}^q(x, ed = 0) \leq D \leq w_{up}^q(x, ed = 1) - w_{low}^q(x, ed = 0)$$

where $w_{low}^q(x, ed = i)$ is the lower bound of the q^{th} quantile of wages conditional on X and educational level i . Similarly $w_{up}^q(x, ed = i)$ is the upper bound.

Figure ?? shows the educational differentials for men at different ages plotted against time. The single solid line is the observed differential, the lines with the circle are the bounds we obtain to the differential based on combining the median and the monotonicity restriction and the broken lines are the 95% confidence intervals for the bounds.

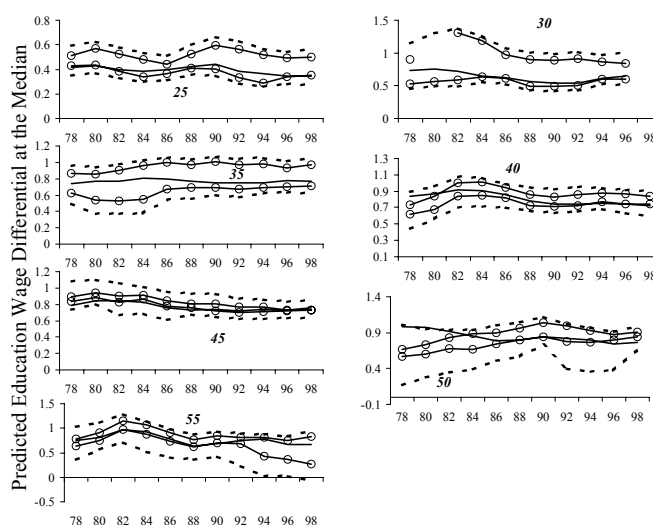


Figure 13: Returns to College for women by age and year - Monotonicity and Median restriction

In most case the observed educational differentials is within the bounds - something not guaranteed once we impose our restrictions.⁹ There are however exceptions; these are for the 45 and the 55 age groups in the early years where the bounds imply much lower returns to education than the observed estimated. However, as the confidence intervals indicate the differences are not significant.

When we consider the change of the differentials over time we see that for the youngest group they increased substantially up to about 1992 and fell after that point back to where they were in the late 70s. However, they increased continuously for 30 and 35 year olds, but remained more or less constant for older groups, with the exception of 55 year olds where the point estimates of the bounds suggest large increases over time.

Because of the low participation rates for unskilled women the bounds to the educational differentials are quite broad, particularly for the under 40 group (figure 13). For the remaining age groups there is no single clear picture as to what occurred to educational differentials. For some age groups they seem to have increased (40,50), while for others they have decreased or remained constant.

⁹Obviously the worst case bounds always contain what we actually observe.

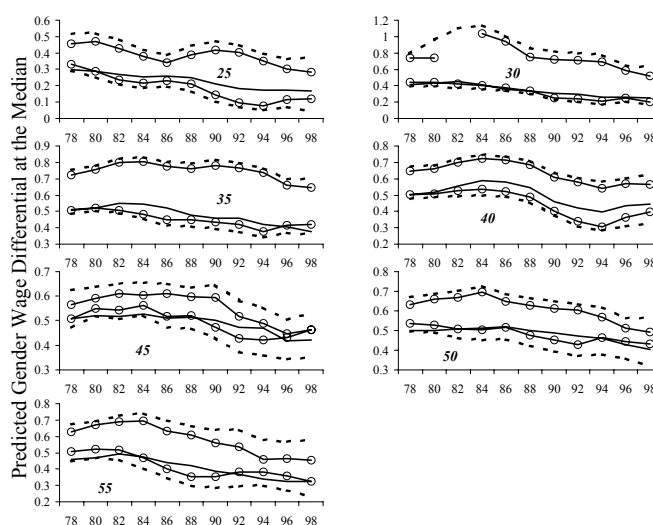


Figure 14: log of male to female hourly wages by age and year. Unskilled workers - Median and monotonicity restriction

5.2.3 Gender wage differentials

A key issue is how do the wages of women compare to those of men and are women's wages improving generally. This is a particularly difficult question exactly because of compositional questions: Which type of men are working and which type of women and how is this changing?

In figure 14 we present the difference of log male wages to log female wages for unskilled workers. The bounds to the differentials are quite wide because of the low participation rates of both unskilled men and unskilled women. However there are indications that if anything the differentials have decreased since 1978 (i.e. wages of women grew relative to those of men).

When we turn to the College graduates in figure 15 we see that differentials fell for all but 45 and 55 year olds where they seem to have increased (although probably not significantly so). In some cases the precision is low because of the relatively low sample sizes among higher education workers. Nevertheless interesting conclusions can be drawn.

In a number of cases selection effects seem to be very important. For 30, 35, 40 and 50 year olds the bounds imply higher differentials than the ones calculated directly from

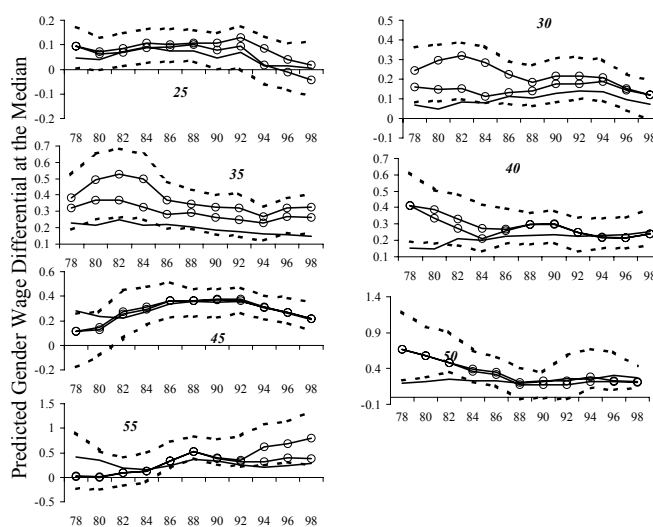


Figure 15: log of male to female hourly wages by age and year. College Graduates - Median and monotonicity restriction

observed workers' wages. In other words the productivity of women relative to men is higher among workers than it is in the overall population. Moreover, in some of these cases the observed differentials record a slight worsening position for women, when the bounds imply an improvement.

6 Conclusions

In this paper we have asked a very simple question: What can be learned about the evolution of economic opportunity (as reflected by the wage) without making strong behavioral assumptions or relying on any narrowly defined economic model. In particular we focus on the question of whether observed changes could be an artifact of changes in composition induced by the momentous changes in male and female participation over the last 25 years or so. We show in fact that in some cases we can get quite definite conclusions on how things have changed. However, we also show that many conclusions taken for granted (e.g. male/female differentials have declined) are not robust and critically depend on what one assumes about the employment model one has in mind.

We have explored ways to tighten Manski’s worst case bounds and implemented the methods using the UK Family Expenditure Survey to explore the changes in wage distribution accounting for the compositional effect explicitly. We showed that while the worst case bound is sometimes useful, for example in examining the wage growth over cohorts, that alone in many other cases is not sufficient to provide definite results. We proposed and examined the effect of imposing the restriction that wages of workers at the median are higher than the potential wage of non-workers at the median and demonstrated that the condition is informative as described above.

In this version of the paper we have not presented results for bounding the distribution of returns to education when errors are not separable. Neither have we presented such bounding methods for the case where the unobservables (whether separable or not) are not independent of education.

.1 Appendix: 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 on its own, without any other assumptions (exclusion restrictions distributional 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 wages can be written as

$$y_i = m(x_i^1, x_i^2) + \varepsilon_i$$

where $F(\varepsilon|x_1, x_2) = F(\varepsilon|x_1)$ In this case none of the quantiles of ε depend on x_2 . Hence we can write the q th quantile of y as

$$y^q(x_i^1, x_i^2) = m(x_i^1, x_i^2) + g(q, x_{i1})$$

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 x_2 on wages, defined by $y^q(x_i^1, x_i^2 = A) - y^q(x_{i1}, x_{2i} = B) = \Delta_{x_2} m(x_{i1}, x_{2i})$

does not depend on q .¹⁰ Then under these assumptions the tightest bound on the return $\Delta_{x_2}m(x_{1i}, x_{2i})$ can be obtained by searching across quantiles. Thus we have that the tightest bound takes the form

$$\begin{aligned} & \max_q \{y^{qLower}(x_{1i}, x_{2i} = A) - y^{qHigher}(x_{1i}, x_{2i} = B)\} \\ & \leq \Delta_{x_2}m(x_{1i}, x_{2i}) \leq \\ & \min_q \{y^{qHigher}(x_{1i}, x_{2i} = A) - y^{qLower}(x_{1i}, x_{2i} = B)\} \end{aligned}$$

where “Higher” and “Lower” refer to the upper and lower bound functions on the quantiles.

References

- [1] Blundell, R, H. Reed, and T. Stoker (2003), ‘Interpreting Aggregate Wage Growth: The Role of Labour Market Participation’, *American Economic Review*, Vol.93, No.4. September, 114-1131.
- [2] 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
- [3] Heckman J (1974) “Shadow prices, market wages and labour supply” *Econometrica* 42 pp 679-694
- [4] Heckman J LaLonde R and Smith J (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
- [5] Heckman J and Honoré B (1990) “The empirical content of the Roy model” *Econometrica* 58(5) pp 1121-1149
- [6] Heckman, J.J. and E. Vytlacil (2001) Instrumental Variables, Selection Models, and Tight Bounds on the Average Treatment Effect,” with , in *Econometric Evaluations of*

¹⁰The analysis could also be carried out in terms of a continuous variable

Active Labor Market Policies in Europe, M. Lechner and F. Pfeiffer (editors), (Heidelberg, Germany: Physica Verlag), 2001, pp. 1-15.

- [7] Heckman J and Sedlacek G (1985) "Heterogeneity, aggregation and market wage functions; an empirical model of self selection into the labour market" *Journal of Political Economy* 98 (6) pp 1077-1125
- [8] 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
- [9] Koenker, Roger and Gilbert Bassett Jr. (1978) "Regression Quantiles", *Econometrica*, 46, 33-50.
- [10] Manski, C. (1989) Anatomy of the Selection Problem (in New Developments in Statistical Techniques) *The Journal of Human Resources*, Vol. 24, No. 3. (Summer, 1989), pp. 343-360
- [11] Manski, C. (1990) Nonparametric Bounds on Treatment Effects (in Selectivity Bias: New Developments) *The American Economic Review*, Vol. 80, No. 2, Papers and Proceedings of the Hundred and Second Annual Meeting of the American Economic Association. (May, 1990), pp. 319-323.
- [12] Manski C (1994) "The Selection Problem," in *Advances In Econometrics, Sixth World Congress* Vol 1., Editor C. Sims, Cambridge University Press.
- [13] Manski C and J Pepper (2000) "Monotone Instrumental Variables: with Application to the Returns to Schooling," *Econometrica*
- [14] Schmitt J (1994) "The changing structure of male earnings in Britain 1974-1988" in R Freeman and L Katz (eds) *Differences and Changes in Wage Structures* Chicago, University of Chicago Press