Centre for Longitudinal Studies

CLS Cohort Studies

Working Paper 2009/2

Trends in the relative wage opportunities of women and men across three British generations

Jenny Neuburger
Diana Kuh
Heather Joshi

December 2009

Trends in the relative wage opportunities of women and men across three British generations

Jenny Neuburger
London School of Hygiene & Tropical Medicine, University of London

Diana Kuh
National Survey of Health and Development
Unit for Lifelong Health and Ageing
Medical Research Council

&

Heather Joshi
Institute of Education, University of London

December 2009

First published in December 2009 by the
Centre for Longitudinal Studies
Institute of Education, University of London
20 Bedford Way
London WC1H 0AL
www.cls.ioe.ac.uk

ISBN: 978-1-906929-10-7

The Centre for Longitudinal Studies (CLS) is an ESRC Resource Centre based at the Institution of Education. It provides support and facilities for those using the three internationally-renowned birth cohort studies: the National Child Development Study (1958), the 1970 British Cohort Study and the Millennium Cohort Study (2000). CLS conducts research using the birth cohort study data, with a special interest in family life and parenting, family economics, youth life course transitions and basic skills.

This document is available in alternative formats.
Please contact the Centre for Longitudinal Studies.
tel: +44 (0)20 7612 6875
email: info@cls.ioe.ac.uk

CONTENTS

Acknowledgements

The funding for this work was provided via an ESRC studentship and the work was linked to the ESRC funded Gender Equalities Network (GeNet). Thanks are due to everyone at the Centre for Longitudinal Studies and the MRC National Survey of Health and Development who are engaged in running the cohort studies and to cohort members themselves. Lorraine Dearden and Anna Vignoles provided useful comments.

Abstract

Since the 1970s in Britain, women's hourly wages have increased, in real terms and relative to men's wages. The observed increase may differ from trends in wage opportunities for the whole population though, since the proportion of women in work has simultaneously increased and, with it, the relative characteristics of the workforce have changed. We have analysed trends across three British generations, covering the period 1972-2004. We use detailed, longitudinal data from the birth cohort studies to impute potential wages for non-employed individuals. Our results suggest that observed wage trends understate the full increase in women's wage opportunities over this period.

1 Introduction

Between 1970 and 2008 in Britain, women's mean hourly earnings increased from 60 per cent to 75 per cent of men's (Figure 1). Over the same period, the proportion of women employed rose from 60 per cent to 70 per cent, while employment of men fell from 90 to 80 per cent (Figure 2). The simultaneous changes in relative wages and employment rates raise difficulties in comparing the relative labour market opportunities of women and men over this period.

Figure 1: Ratio of women's to men's mean hourly earnings in full-time work. Source: New Earnings Survey 1970-2002, Annual Survey of Hours and Earnings 2003-2007, Office for National Statistics

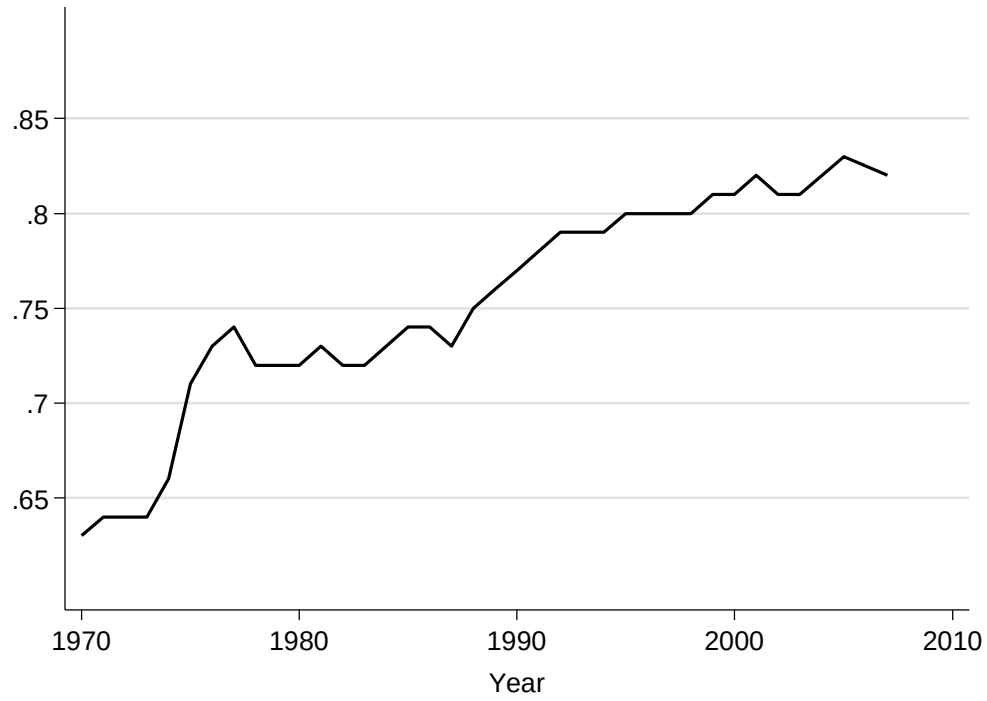
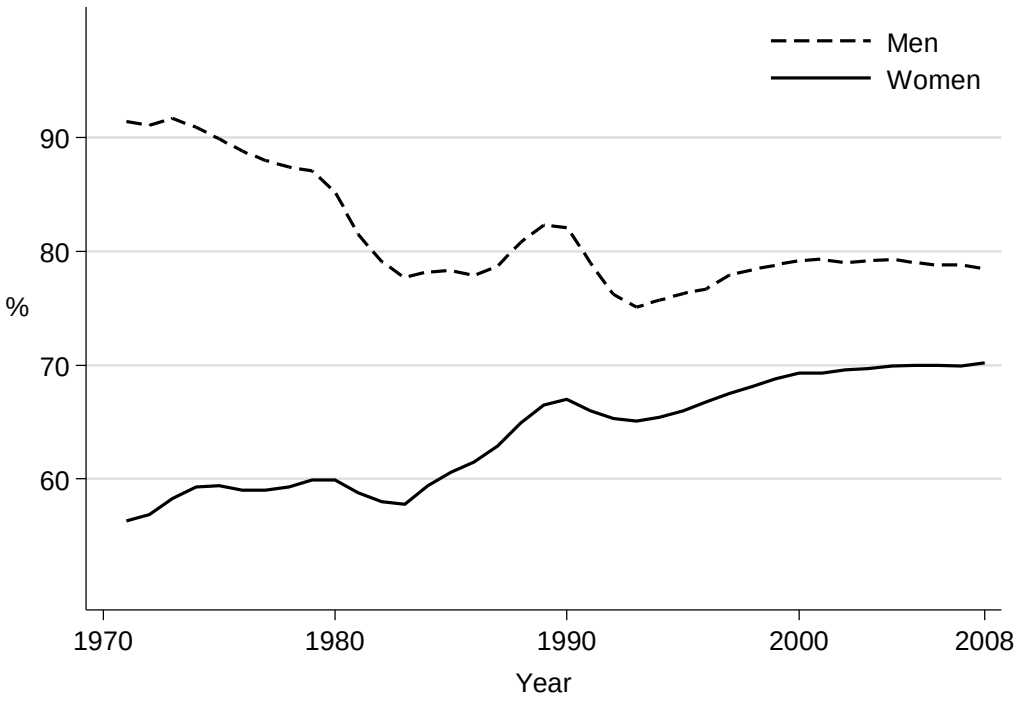


Figure 2: Percentage of women and men in employment (women aged 16-59, men aged 16-64). Source: Labour Force Survey, Office for National Statistics



Focusing on trends in earnings for the working population is problematic in that changes in potential earnings and employment rates are interdependent. Earnings are observed only for individuals who want and find work, whilst employment decisions are partly based on the likely wage. Consequently, an improvement in wage opportunities for women could lead to an increase in employment rates, as well as in observed wages. The observed increase in wages may though lag behind the increase in wage opportunities, if women with relatively low wages are drawn into the workforce. Alternatively, changes in social norms around childcare, maternity leave, the distribution of wages and family structures may contribute to a variety of changes in women's selection into employment, with indeterminate impacts on observed relative to potential wages. Other studies (reviewed below) have found evidence consistent with employment selection effects concealing some of the improvement in women's wage opportunities over the last three decades in the UK and the US (e.g. Blundell, Gosling, Ichimura and Meghir, 2007; Blau and Kahn, 2006).

This article aims to reconstruct trends in wage opportunities, using detailed longitudinal data from the British Birth Cohort Studies, described below, to impute wage offers for non-working individuals. Wage and employment data from the 1946, 1958 and 1970 cohort studies are brought together to analyse trends within and across three generations. The detailed information collected on job histories and family circumstances yields plausible estimates of potential wages for non-working individuals. We impute potential wages for non-workers by borrowing observed values from workers with similar histories.

We find that the observed cross-generational increase in younger women's average wage is less than the increase in estimated wage opportunities. This is because a larger fraction of women with relatively low wage potential were not employed in 1972, compared to later years. This is consistent with the finding cited above (Blundell et al., 2007). We also find some evidence that selective withdrawal and re-entry into the labour market around childbearing masks some of the decline in women's earning potential after having children.

The next section discusses trends in employment and hourly earnings for the years covered by our analysis. Section III reviews previous work on selectivity biases in wage analyses. Section IV describes the datasets used in our analysis and discusses the method used to impute potential wages for non-working individuals. Section V presents our analysis of employment selection and estimates of changes in women's and men's relative wage opportunities. Section VI draws out the key conclusions of this study.

2 Cross-cohort trends in employment and earnings

The change in women's employment and pay across the cohorts is striking. In 1972, just under half of women from the 1946 cohort were in work at the age of twenty-six. Nearly quarter of a century later, more than three-quarters of women from the 1970

cohort were in work at the same age. Across these samples, women's median hourly earnings increased from 70 per cent to 90 per cent of men's earnings.

The proportion of women in work, in their twenties and thirties, has increased across each successive cohort. This is illustrated in Figure 3.[1] The cross-cohort increase in women's rates of employment is composed of women having fewer or no children; and of mothers being more likely to work when they do have young children.

Figure 3: Percentage of women in work at each age, by cohort (including self-employees)

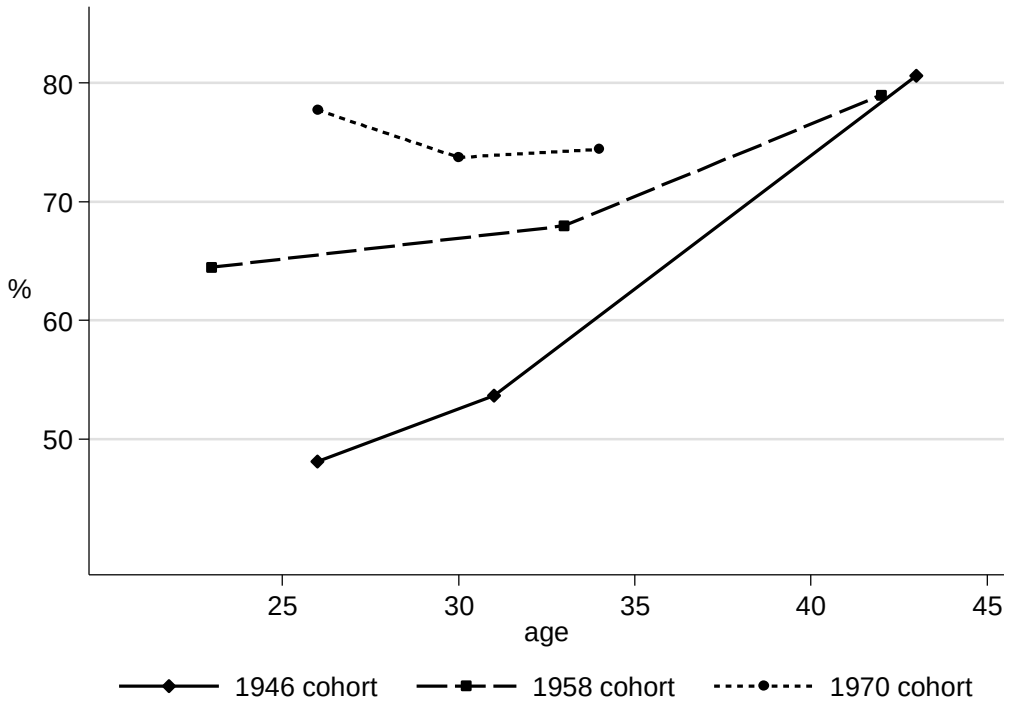
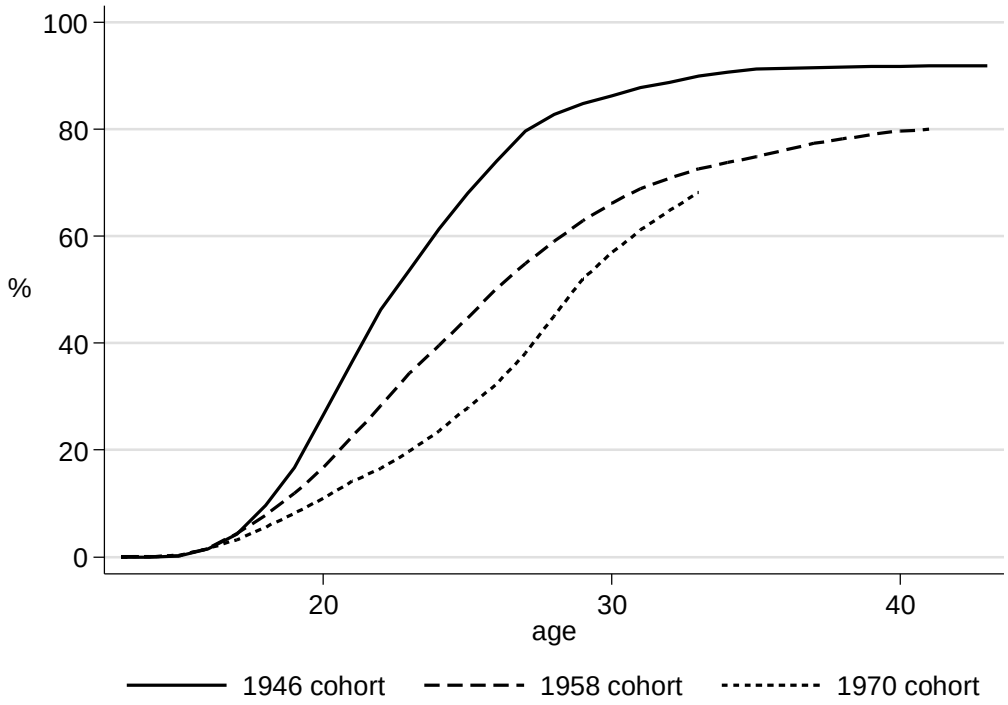


Figure 4 shows the percentage of women in each cohort who had become mothers by each age for the three cohorts.[2] Across the cohorts, years of entry into motherhood have become later and more spread out. By age 26, more than 60 per cent of women in the 1946 cohort had become mothers, compared to less than a quarter of women in the 1970 cohort at the same age. De Cooman, Ermisch and Joshi (1987) drew attention to the negative aggregate impact on births of the increase in women's real earnings over the 1960s and 1970s. Hobcraft (1996) identified the role of increased education, particularly the raising of the school leaving age, of improved contraception, changes in housing, the increased instability of partnerships and the emergence of a culture of 'assertive individualism'. The postponement and decrease in births has been greater amongst highly-qualified

[1] The figures for cohort trends in this section are based on data from the cross-sectional samples of all cohort members interviewed at a particular age. Figures for the 1946 data, which are based on a stratified sample, are weighted to give population estimates.
[2] The figure uses birth history data collected from all mothers and is based on the sample interviewed at age 43 for the 1946 cohort, age 42 for the 1958 cohort and age 34 for the 1970 cohort.

women (Joshi, 2002; Jenkins, Killingsworth and Joshi, 2008; Kneale and Joshi, 2008).

Figure 4: Percentage of women who have had a first child by each age, by cohort (live births only)

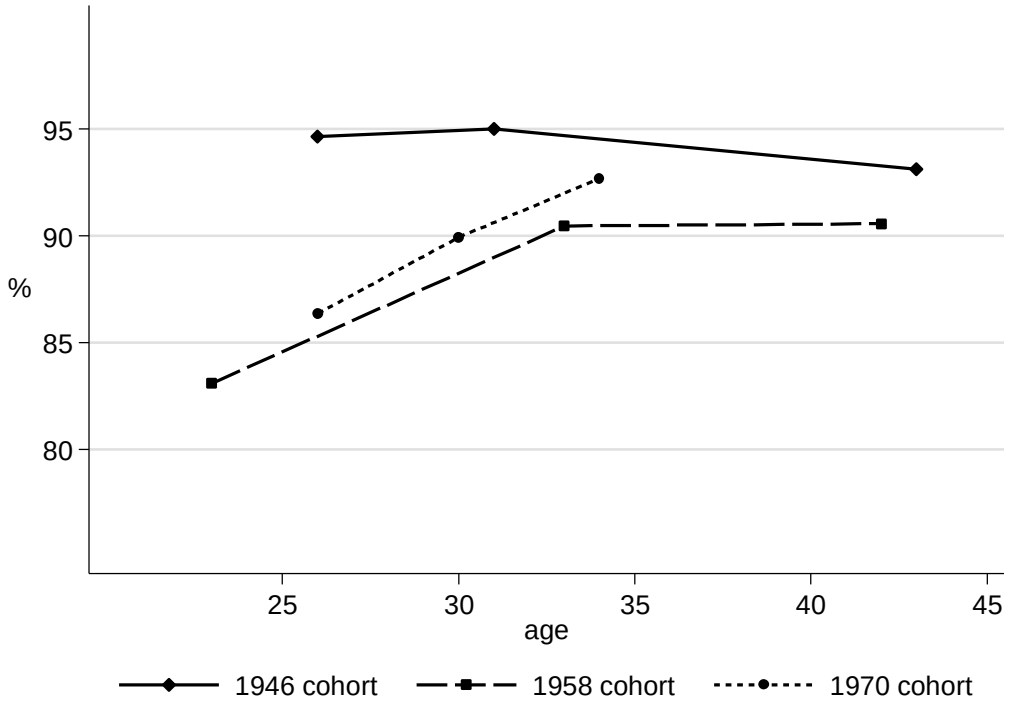


The length of mothers' employment breaks has also decreased across the three cohorts. For the 1946 cohort, only a fifth returned to work within a year of a first birth, compared to nearly two-fifths for the 1958 cohort and close to 60 per cent for the 1970 cohort. Again, the increase in women's earnings capacities has been identified as an important contributory factor (Joshi, Layard and Owen, 1985), along with the introduction of paid maternity leave in 1975. Using cross-sections of the Family Expenditure Survey for the period 1974-2000, Gregg, Gutierrez-Domenech and Waldfogel (2007) presented evidence that the introduction of maternity leave, and subsequent policy changes to increase its generosity and coverage, have helped raise employment amongst mothers of young children.

The 1980s and 1990s were decades of increasing differences in the employment rates of mothers with and without qualifications (Macran, Dex and Joshi 1996; Joshi, 2002). Gregg et al. (2007) found that mothers with the highest potential earnings responded most to the change in maternity legislation in the early 1980s, whilst those with middle earnings increased their rates of employment later in the 1980s and 1990s. Mothers in the lowest predicted earning third did not change their employment behaviour. The authors speculate that this lack of increase in employment for mothers with lower earnings prospects may have been due to increasing childcare costs, in line with growth in real earnings at the lower end of the wage distribution.

In contrast to women, men's rates of employment have decreased slightly across the three cohorts (Figure 5). Members of the 1958 and 1970 cohorts were affected by recessions at the start of their careers. There is considerable evidence on the long-term scarring effects of unemployment on future earnings and job prospects, including for members of the 1958 cohort (Arulampalam, 2001; Gregory and Jukes, 2001; Gregg, 2001). The decrease in relative wages for unskilled work, alongside the weak incentives to come off incapacity benefit once on it, contributed to the persistence of long-term economic inactivity over periods of economic growth and job creation (Faggio and Nickell, 2003; Disney and Webb, 1991).

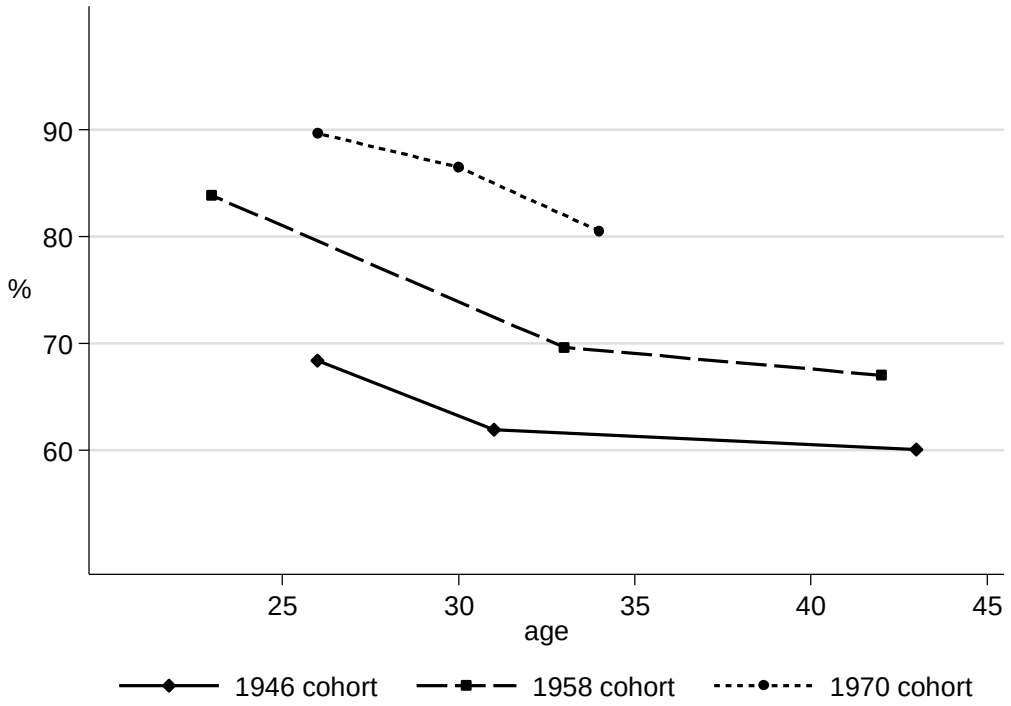
Figure 5: Percentage of men in work at each age, by cohort (including self-employees)



Women's median hourly earnings increased relative to men's across the three cohorts. Figure 6 illustrates this increase, alongside a decrease within cohorts over their twenties and early thirties.[3] There is strong evidence that the Equal Pay Act was directly responsible for the sharp increase women's relative pay in the mid 1970s (Neuburger, 1984; Zabalza and Tzannatos, 1985; Joshi et al., 1985; Manning, 1996).

[3] Women's and men's median real earnings both increase with age in each cohort, but men's increase more than women's.

Figure 6: Women's median hourly earnings as a % of men's at each age, by cohort (employees only)



The other major factor contributing to the increase in women's relative earnings and employment has been the increase in their levels of education. Educational opportunities increased for women and men born in the 1950s, with the introduction of comprehensive education under the 1964-1970 Labour Government, the raising of the school leaving age in 1974, and the expansion of university places in the 1960s and 1970s. Across the 1946 and 1970 birth cohorts, the proportion with no or very low qualifications when in their early thirties fell from around half in 1978 to a fifth in 2000 for both women and men. Even more striking, the proportion with degree-level or higher qualifications increased from just 10 per cent to 32 per cent for women, and from 21 per cent to 31 per cent for men (Makepeace, Dolton, Woods, Joshi and Galinda-Rueda., 2003, Table 2.1).

The complex and interdependent trends in women's and men's employment participation, wages, education and family formation mean that the relationship between trends in observed wages for the working population and underlying wage opportunities for the whole population are not simple to predict. The exercise of imputing potential wages for non-workers, though, requires certain assumptions, in addition to the rich evidence in the cohort studies.

3 Previous work on selection bias in wage analyses

Simple weighting approaches have been used for some time in applied labour economics to assess interdependent trends in wage opportunities and employment participation (e.g. Smith and Welch, 1986; Welch, 1990; Juhn, 1992). The

assumption in these studies is that non-working individuals are likely to face potential wages close to those of working individuals who have similar educational and employment characteristics. For example, to analyse the role of declining wage opportunities in men's decreasing labour market participation in the US, Juhn (1992) assumed that non-workers faced the same potential wages as workers with the same levels of education and employment experience, weighting wage trends to reflect the distribution of these across the male population at each period.

Concern about unobserved selectivity biases in wage analyses was raised in the econometric literature in the 1970s (Gronau, 1974; Heckman, 1977; Heckman, 1979). Both Gronau (1974) and Heckman (1977) had in mind scenarios in which working and non-working women with the same observed educational and employment characteristics faced different wage offers.

Gronau (1974) considered the problem of unobserved selectivity bias in the context of a job-search model of the labour market. Within this framework, he noted that workers and non-workers each faced a range of wage offers, not just one, and that lower offers would be more likely to be rejected. As a consequence, observed wage distributions for groups characterised by partial employment represented just the acceptable part of a distribution of wage offers facing the whole group. Gronau (1974) discussed a series of wage comparisons that would be affected by this form of selection bias. First, wage comparisons across women and men could give a misleading picture of gender differentials in wage offers. Second, the dynamic effect of an improvement in wage offers would be to increase rates of employment participation and average pay, but with the observed increase in average pay lagging behind the improvement in the average wage offer. Third, the effects of new mothers' selective withdrawal and re-entry into the workforce could mask the underlying decline in wage offers.

Heckman (1977) and Lewis (1974) pointed out that similar forms of bias could arise from self-selection within heterogeneous groups. For women, employment decisions may be affected by a complex and interdependent set of observed and unobserved factors, including: social norms and preferences around childcare and paid work; the cost of childcare relative to potential earnings; marital or cohabitation status; partners' earnings; and other sources of household income.

Heckman (1979) proposed the most widely-used method for dealing with unobserved selectivity bias. The method relies on finding suitable exclusion restrictions, i.e. variables which affect selection into a sample, but not the outcome of interest. The main problem with many applications of Heckman's method (and non-parametric versions of this) to the analysis of women's wages is that the exclusion restrictions are questionable. Commonly used variables include: partner's income; household wealth; non-wage household income; housing tenure; and number and ages of children (e.g. Zabalza and Arrufat, 1985; Dolton and Makepeace, 1987a; Ermisch and Wright, 1993; Joshi and Paci, 1998). These variables are in all cases strongly correlated with the probability of being in work. What is not testable and open to debate is whether these variables are also correlated with the unobserved wage offer, e.g. if motherhood directly affected potential wages; if couples' earnings were

directly correlated; or if household wealth was correlated with past earnings (and current prospects). More credible exclusion restrictions are those which draw on institutional arrangements affecting work incentives, such as out-of-work benefits entitlement (Blundell et al., 2007).[4] Angrist and Krueger (2001) have pointed out that the bias in estimates from models using a bad instrumental variable (which are correlated with the error term) can be greater than the original biases it is intended to correct. The same objections apply to mis-specified exclusion restrictions.

In practice, unobserved selectivity biases have frequently been estimated to be small and negligible, relative to differences in wage opportunities associated with differences in education and employment histories across workers and non-workers (e.g. Zabalza and Arrufat, 1985). An exception is an analysis by Mulligan and Rubinstein (2008), who found evidence of substantial unobserved selectivity biases in women's wages, using Heckman's method, treating the presence of young children in the household as excludable from the wage equation and controlling for education, potential work experience, marital status and region. The exclusion restriction in this study is questionable, since there are a number of ways in which motherhood might be directly correlated with the wage offer.

An alternative approach to accounting for selectivity bias was proposed by Manski (1989). He pointed out that it is possible to estimate upper and lower bounds on quantiles of the wage offer distribution without any assumptions about the wage offers of non-working individuals. The proportion of missing data arising from non-participation, and the maximum and minimum values within which the missing potential wage data are logically constrained to lie, determine the size of the bounds. Additional assumptions, including exclusion restrictions, can be introduced to tighten the bounds.

Blundell et al. (2007) looked at trends in gender differentials in potential pay for Britain, using cross-sectional data from the Family Expenditure Survey (FES) for 1978 and 1998. Dividing the population into four broad groups according to age and educational attainment, they estimated bounds on the wage distributions to account for selection, as proposed by Manski (1989). In the first instance, they found that the low proportion of women in employment in the 1978 sample made the unrestricted bound too wide to get an informative estimate of change for any of the groups. Imposing restrictions on the model of positive selection into employment and partial exclusion restrictions (related to changes in the out-of-work benefits system)[5], they found evidence that the gender differential in potential pay had decreased for younger, less-qualified women and men, although not for other groups. Their results

[4] Other less intuitive exclusion restrictions have also been used in the literature, including individual attitudes to work and family (e.g. Albrecht et al., 2004) and holding a loan or mortgage on a property (Dolton, Marcenaro-Gutierrez, and Skalli, 2009).

[5] They assumed that out-of-work benefits entitlement, holding fixed household composition, affects the employment decision (via the reservation wage) but is only weakly and positively correlated with an individual's potential wage. This is referred to as the monotonicity restriction in the paper.

also suggested that the gender differential in potential wages had decreased more than was observed in wage trends for this group.

More recent studies have exploited the longitudinal aspect of panel datasets to impute potential wages for non-working individuals based on their observed wage when in work on a previous or subsequent occasion. This implicitly accounts for fixed observed and unobserved influences on the wage offer, though not for dynamic influences. Blau and Kahn (2006) used an imputation approach to estimate potential wages for non-employed individuals, using the Panel Survey of Income Dynamics (PSID) for the years 1979, 1989 and 1998. They replaced missing potential wages for individuals out of work with the wage observed within a four-year window when that individual was last or next in work. For individuals who had not been in paid work over the four-year window, the missing potential wage was placed above or below the median based on their level of education and employment experience. They found that fixed selectivity biases were substantial and significant, and again, evidence of a greater convergence in women's and men's wage opportunities than is observed in wage trends for the working population.

Olivetti and Petrongolo (2008) also used imputation methods to estimate gender gaps in potential wages across fourteen OECD countries. Using data from the PSID and the European Community Household Panel Survey in the period 1994-2001, they placed the missing wages of non-employees either side of the median based on their most recently observed wage. For individuals who had not been employed over the whole period, their potential wage was estimated on the basis of their observed characteristics. They found that selectivity biases were positive and substantial in Southern European countries, with relatively low rates of female employment. They concluded that cross-country variation in the labour market opportunities of women and men was overstated in observed wage gaps. Comparing the imputation methods, they also concluded that fixed, unobserved selectivity biases were negligibly small in the study context.

4 Data and Methodology

4.1 Data

Our data come from three of the British Birth Cohort Studies, which are continuing national surveys of the same individuals born in single weeks of March 1946, March 1958 and April 1970. In total, more than 20,000 people are still taking part in the studies. For a full account of their histories and a review of findings, see the volume edited by Ferri, Bynner and Wadsworth (2003) and the studies' websites (http://www.nshd.mrc.ac.uk and http://www.cls.ioe.ac.uk).

The earliest birth cohort study – the MRC National Survey of Health and Development (NSHD) - started in March 1946 as a study of childbirth and maternity services (Wadsworth, 1991). All babies born in a single week across England, Scotland and Wales in March 1946 were included in the study. In 1948, a smaller,

stratified sample was followed up, including all children born to fathers in non-manual and agricultural occupations and one in four children born to fathers in manual occupations. The aim was to preserve roughly equal sample sizes from the two social class groups. Babies born to unmarried mothers (672) and multiple births (180) were excluded from the study. The resulting 5,362 two-year-olds (2,547 girls and 2,815 boys) constitutes the core target sample of the 1946 cohort. Weights are provided with the datasets to account for stratification. In childhood, information was collected regularly about family circumstances and schooling from mothers, teachers and school medical officers. Tests of general, mathematics and reading ability were administered at school. In adult life, cohort members have been asked in detail about their health, families and jobs. Face-to-face interviews were conducted at ages 26, 36, 43, 53 and 60 and postal surveys were carried out between the ages of 18 and 21 and at ages 25 and 31.

The 1958 cohort study - the National Child Development Study (NCDS) - included in its original sample all babies born in England, Scotland and Wales over a week in March 1958. 17,416 mothers were interviewed (99 per cent of those eligible). At ages 7, 11 and 16, information was obtained from interviews with mothers, medical records, questionnaires sent to school teachers, school medical examinations, ability tests and questionnaires completed by cohort members themselves. Children who had been born outside Britain in the sample week were recruited into the study over this period. In adulthood, face-to-face interviews have been carried out with cohort members at ages 23, 33, 42 and 50, and a telephone survey at age 46.

The 1970 British Cohort Study (BCS70) initially included all babies born in the United Kingdom in a week of April 1970, 17,200 were successfully interviewed. Those born in Northern Ireland were not followed up. Children born outside Britain in the study week were recruited at ages 5 and 10. Data were collected at ages 5, 10 and 16 from parents, school teachers, medical officers and cohort members themselves. A postal survey was sent out to cohort members at age 26 and face-to-face interviews have been conducted at ages 30 and 34, plus a telephone survey at 38.

A problem common to all long-running cohort studies is that some individuals take part intermittently and others do not remain in the study. By around age thirty, 60 per cent (62% for the 1946 cohort) of the original birth samples were still taking part in each of the three studies. A substantial fraction had died or emigrated. For the 1946 cohort, 11 per cent had emigrated by age 36 and 6 per cent had died, so that 79 per cent of the available sample were still taking part. Around 70 per cent of the target sample for the 1958 cohort study was still taking part at age 33 and just under 70 per cent of the 1970 cohort at age thirty. The figures for the 1970 cohort are more likely to be revised as more information comes to light about untraced individuals in future tracing exercises (Plewis, Calderwood, Hawkes and Nathan, 2004). For the 1958 and 1970 cohort studies, Nathan (1999) found that, although cumulative attrition was substantial, the evidence did not indicate serious biases in analyses for general population samples. Wadsworth et al. (2003) drew a similar conclusion for the 1946 cohort.

Information about earnings has been collected from each cohort on at least three occasions. Table 1 summarises information on the survey samples from which earnings data have been used covering the period 1972-2004. (all tables are shown at the end of the paper). Earnings data collected at several other surveys were not used owing to different methods used for capturing data in the 1946 cohort or unacceptably high levels of missing data in the 1958 cohort.[6] We found evidence that relatively low response rates to the postal survey of the 1970 cohort in 1996 (age 26) did affect the representativeness of the sample at this age. The later wages, at ages 30 and 34, of cohort members who did not respond to the postal survey were between 8 and 12 per cent lower than those of respondents. Consequently, we carried out our analysis both for the cross-sections who took part in any one of the adult surveys, which are presented, and also for the sub-sample who took part in all three of the adult surveys (not presented). Our conclusions were not altered.

Wages are measured as a before-tax hourly rate, calculated by dividing before-tax (gross) reported earnings by the number of weeks contained in the corresponding reported pay period and dividing again by reported weekly hours of work, including paid overtime. The earnings questions asked at each of the surveys are shown in tables 8-10. Hourly wages are adjusted to January 2000 prices using the long-term indicator of prices of goods and services produced by Office for National Statistics.[7]

4.2 Methodology

The focus of our analysis is trends in the relative potential wages of all women and men, including those of the non-working population. The expected potential wage, for a given group, can be written out as a weighted average of the expected potential wage for employees, which is observed, and the expected potential wage for non-employees, which is unobserved:

$$\mathrm{E}(w^0) = \underbrace{\mathrm{E}(w^0 \mid s=1)}_{Observed}.\Pr(s=1) + \underbrace{\mathrm{E}(w^0 \mid s=0)}_{Unobserved}.[1-\Pr(s=1)] \quad (1)$$

where w^0 is the individual potential wage and s represents employment status (1 = employed, 0 = not employed). From equation 1, it becomes clear the size of the difference between the expected wage offer for the whole population, $\mathrm{E}(w^0)$ and the expected wage for employees, $\mathrm{E}(w^0 \mid s=1)$, i.e. the selectivity bias, is affected by: firstly, the employment rate, $\Pr(s=1)$, which determines the fraction of missing data; and, secondly, the value of the expected potential wage for non-employees, $\mathrm{E}(w^0 \mid s=0)$. It is the latter term which is unobserved and is the object of estimation.

[6] Earnings data were collected at ages 36 and 53 from the 1946 cohort and at age 46 from the 1958 cohort.
[7] Available at http://www.statistics.gov.uk/StatBase/TSDdownload1.asp.

4.3 Selection on observables

The working assumption is that employed and non-employed individuals who have the same observed characteristics face the same wage offer.[8] The assumption of selection on observables is equivalent to the *conditional independence assumption.* This assumption implies:

$$\mathrm{E}(w^0 \mid \mathrm{X}, s = 0) = \mathrm{E}(w^0 \mid \mathrm{X}, s = 1) \tag{2}$$

where X is a vector of selected characteristics.

In selecting variables on which to match employed and non-employed individuals, the aim is to cover the main joint influences on employment decisions and potential earnings. A suitable wage donor is always of the same sex, age and cohort. Apart from this, non-working individuals are matched to working individuals who are similar across the following sets of selected characteristics.

Childhood characteristics include: mother's age at birth; the number and ages of any siblings; parents' school-leaving ages; and father's social class in childhood. There is a large literature indicating that coming from a larger family, having poorly educated parents, being born to a younger mother, being a younger sibling and having a father in a lower status job are all markers of childhood disadvantage and that these shape employment and earnings prospects in adulthood (e.g. see Kuh and Wadsworth, 1991; Kuh, Head, Hardy and Wadsworth, 1997; Blanden, Gregg and Macmillan, 2007; Plewis and Kallis, 2008; Flouri and Hawkes, 2008).

Scores from mathematics and reading tests taken at age ten or eleven are included as indicators of educational achievement at these ages. Similar to Joshi, Makepeace and Dolton (2007), we have standardised the scores to make them comparable, with a mean zero and standard deviation one.

Adult characteristics include: level of highest qualification obtained; years spent in a current job; years spent in full-time work; years spent in part-time work; region of residence; whether any children under 16 are living in the household; whether there is more than one child in the household; and whether there are children aged under five in the household. For the 1958 and 1970 cohorts, information on social class status of the first job after leaving full-time education was also included. These variables have all previously been shown to be strong predictors of wages and employment participation for cohort members (e.g. see Joshi and Paci, 1998; Blundell Dearden, Goodman and Reed, 2000; Joshi et al., 2007). Tables 4 & 5 summarises the means of the variables used.

[8] With the cohort data, we cannot reasonably impute wages for non-employees using their observed wage at a previous or subsequent survey, since the age-dynamics are themselves of interest and because the surveys were conducted too infrequently.

The assumption of selection on observables is arguably satisfactory in the present study owing to the wealth of data available on individual family and employment histories. Heckman-type methods to deal with unobserved selectivity bias were investigated but not pursued owing to the lack of credible instruments. Some further exploration of the data was also carried out to assess the likely presence and size of any unobserved biases, the results of which are presented alongside our main results.

4.4 Nearest-neighbour imputation using the propensity score

Imputation methods have come out of statistical work on methods to handle bias arising from missing data in surveys (see Rubin, 1976; Rubin, 1987; Little and Rubin, 2002; Sande, 1982; Skinner, Stuttard, Durrant and Jenkins, 2002; Durrant, 2006). For each missing item of data, an attempt is made to find a donor record without missing data, most similar to the survey record in question, and to use the observed value to fill in the missing item. Nearest-neighbour matching, predictive mean matching and hot deck imputation are some of the methods developed for identifying an acceptable match for a record with a missing item of data. An example of an application of predictive mean matching to impute low wages in the Labour Force Survey can be found in Skinner et al. (2002) and Dickens and Manning (2004).[9]

Separately, Rosenbaum and Rubin (1983) have shown that conditioning on the propensity score is sufficient to remove bias arising from selection on observables, so equation (2) can be re-written as:

$$\mathrm{E}(w^0 \mid p, s=0) = \mathrm{E}(w^0 \mid p, s=1) \tag{3}$$

where $p = \Pr(s=1 \mid \mathrm{X})$.

Combining these methods, the imputation method used for the present analysis involves two stages. First, the degree of similarity between employees and non-employees across the selected characteristics is reduced to a single propensity score. Second, each non-employed individual is matched to an employee with the most similar propensity score and their missing potential wage is replaced with the value observed for their matched donor.[10] Potential wages are imputed in the same way for self-employed individuals and for employees with incomplete wage data.

The propensity score is the individual's predicted probability of being in employment, calculated from probit models using the set of selected characteristics described above to discriminate between the employed and the non-employed, separately by

[9] Individuals with missing hourly wages have these replaced with values borrowed from individuals with observed wages who have the same predicted wage rate. Predicted rates are obtained from a linear regression of log wages on a set of observed characteristics.

[10] Two alternative methods were also used and gave similar results. The first involved re-weighting the observed wage data for employees based on propensity score matching. The second involved imputing predicted log wages from a least squares model.

age, gender and cohort. This score gives greatest weight to those characteristics that have the largest effect on employment participation. The estimated propensity score, $\hat{p}$, is given by:

$$\hat{p} = \Phi(\mathrm{X}'\hat{\beta}) = \Pr(s = 1 \mid \mathrm{X}) \tag{4}$$

where the probit link function, $\Phi(.)$ is the cumulative distribution function of the standard normal distribution and $\hat{\beta}$ is a vector of parameters estimated by maximum likelihood. Tables 6 and 7 summarises the estimated probit models.[11]

Nearest neighbour matching is carried out on a pair-by-pair basis. Each empty cell, representing the unobserved potential wage for a non-employee, is replaced with a value equal to the observed wage for the employee with the closest propensity score. Given a sample of r employees and $n-r$ non-employees, the estimated propensity scores for the sample are $\hat{p}_1 ..... \hat{p}_n$. The wages for employees are observed as $w_1 .... w_r$ whilst the potential wages for non-employees, $w_{r+1}^o .... w_n^o$, are unobserved. A potential wage is imputed for each non-working individual, i , using the wage of the most similar working individual, j , identified by the algorithm:

$$| \hat{p}_i - \hat{p}_j | = \min_{1 \le k \le r} | \hat{p}_i - \hat{p}_k | \tag{5}$$

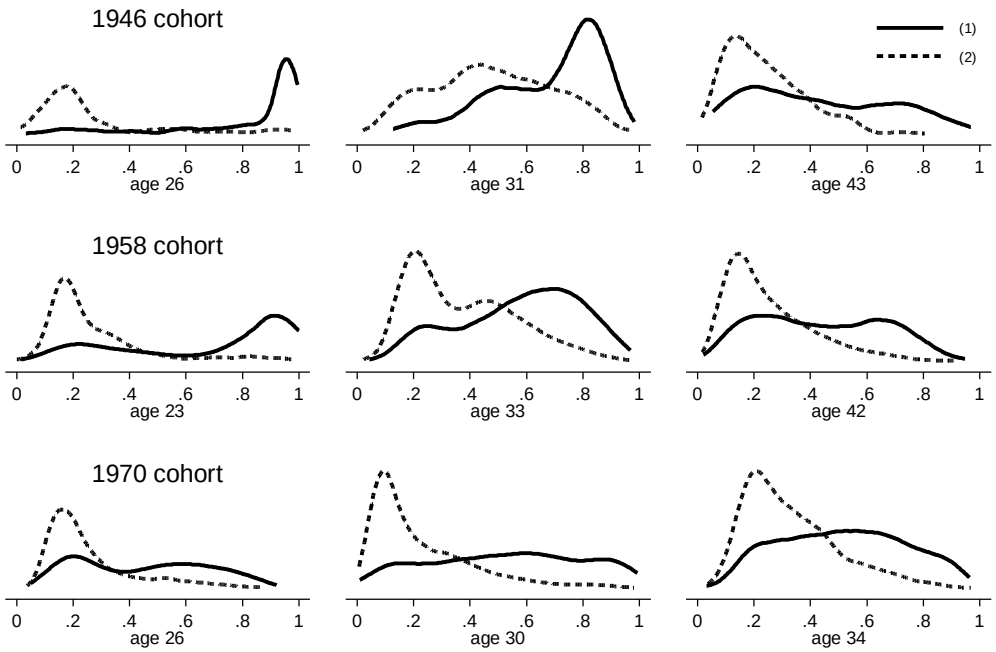
where k represents any other employed individual. In our analysis, a wage donor, i.e. an employed individual with a low predicted probability of being in work, could appear more than once.

A common support restriction was used, which meant not imputing potential wages for non-employed individuals who had a propensity score lower than the *lowest* score estimated for any employed individual in the same sample. Very few imputations failed because of this. Figures 7 and 8 show the distributions of propensity scores for the working and non-working groups in each survey sample. Table 2 show the numbers of observed and imputed cases for each sample.

Standard errors on imputed median wages, and on ratios of medians for women and men, were estimated using bootstrap methods.

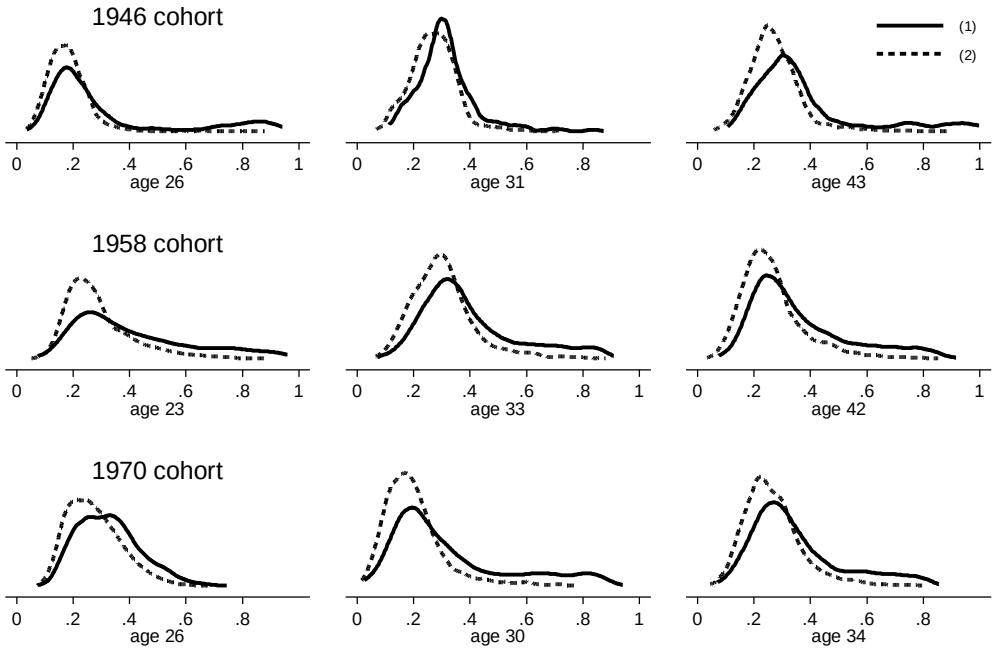
[11] We used pooled probit models, discriminating between employees with observed wages, on the one hand, and all other groups without observed wages, on the other. We also tested a set of separate models for employees vs. the different missing-wage groups and concluded that the pooled models were adequate.

Figure 7: Distributions of propensity scores for 1) employees and 2) imputed sample, women by age and cohort



(1) Employees
(2) Imputed sample (non-employed, self-employed & missing wage cases)

Figure 8: Distributions of propensity scores for 1) employees and 2) imputed sample, men by age and cohort



(1) Employees
(2) Imputed sample (non-employed, self-employed & missing wage cases)

5 Results

5.1 Multivariate analysis of selection into employment

For women and men in each cohort, higher qualifications, higher scores in maths tests taken at age ten or eleven, higher status upon entering the job market and continuity of employment are generally associated higher employment probabilities. Tables 6 and 7 show roughly this pattern, although the signs on the coefficients are reversed - it is non-employment that is modelled - and are smaller owing to the inclusion of self-employees and missing wage cases in the 'non-employment' group for the purposes of imputation.

This pattern of positive selection into employment is less uniform for women, owing to the complexity introduced by differences in the timing of childbirth by social and educational status. More educated women are likely to have children later in each of the three cohorts. As a consequence, non-working women become a more diverse group with age. For the 1946 cohort, childbearing years are spread over a relatively short period. By age 31, most women in the cohort had become mothers (Figure 4), and there were few systematic differences in the educational and family background characteristics of women in and out of work. For the 1958 and 1970 cohorts, social and educational differences between employed and non-employed women persisted into their thirties, but also weakened with age.

5.2 Tests for unobserved selectivity bias

Three data exercises were used to investigate possible unobserved selectivity biases. First, the estimated coefficients on variables in the probit models (Table 6) were compared to a set of coefficients on the same variables in log-linear wage models estimated by least squares (not shown). For the 1958 and 1970 cohorts, gthere were signs of positive selection for women in their thirties and forties. Having young children was a positive predictor both of being out of work at these ages and of having higher wages when in work. This result is suggestive of either: positive selection into motherhood at older ages, which would not affect our wage imputations; or, positive selection into work amongst older mothers, which would result in an upward bias in imputed potential wages for these samples.

A second exercise was to compare the means of each characteristic used in the probit models across the matched samples. We found that the means were generally not significantly different across the matched samples, with some important exceptions. For the 1946 cohort at age 26 and for the 1958 cohort at age 23, working women as a whole group had higher levels of education and higher average maths scores than non-working women. In contrast, the matched working samples had fewer qualifications and significantly lower scores than the matched non-working samples. This pattern suggests negative selection into work amongst mothers at these ages, not adequately accounted for by the matching exercise. For the 1946 cohort, within the group of mothers with young children, it was less educated mothers

who had given birth slightly earlier and were more likely to have returned to work by the age of 26. Imputed wage offers for these non-working groups may consequently be biased downwards. The size of this downward bias is likely to be small and our estimated trends appear to be robust to alternative specifications, e.g. imputing the predicted wage from a log-linear wage model.

A third investigation was carried out for women in the 1958 cohort only, using additional information collected about their weekly wage in their first job after leaving full-time education, requested at the age 23 survey. For the 281 women who provided this information, there was a difference in mean log first wage across the unmatched working and non-working samples, which was marginally significant (at the 10% level), supporting the view that observed selection biases show up slightly in the first wage. In contrast, there was no difference in mean log first wage across the matched working and non-working samples, suggesting that the matching successfully removed any fixed, unobserved biases that appear in the first wage.

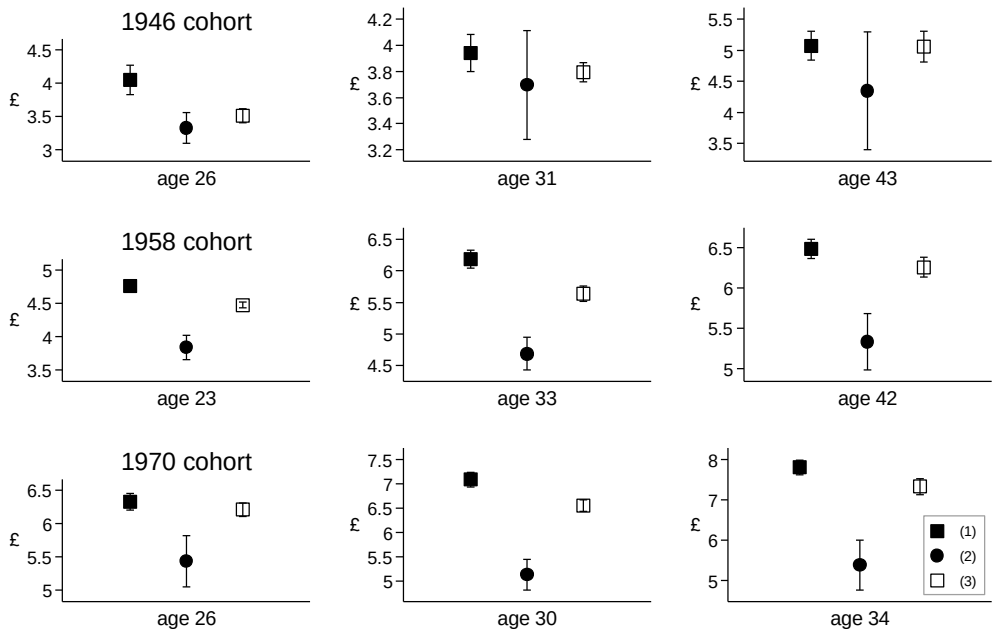
Positive selectivity biases associated with job search intensity or with having a better potential employer would not show up in the tests carried out. Gronau's (1974) observation that similar individuals face a range of wage offers, and that the better wage offers are more likely to be accepted, would further imply that the imputed wage offers in our analysis overstate the actual wage offers faced by non-working individuals. Without exclusion restrictions, it is not possible to quantify the likely size of such biases, although the evidence in the literature (discussed above) suggests that these are likely to be small relative to those captured by differences in education and work experience.

5.3 Estimates of women's and men's relative wage opportunities

Potential wages imputed for non-working women and men tend to feature in the lower parts of the wage distribution, indicating positive selection into work based on observed characteristics. In contrast, self-employees and employees with missing wages do not have systematically different imputed wages to observed wages for employees in the majority of surveys (Table 3).

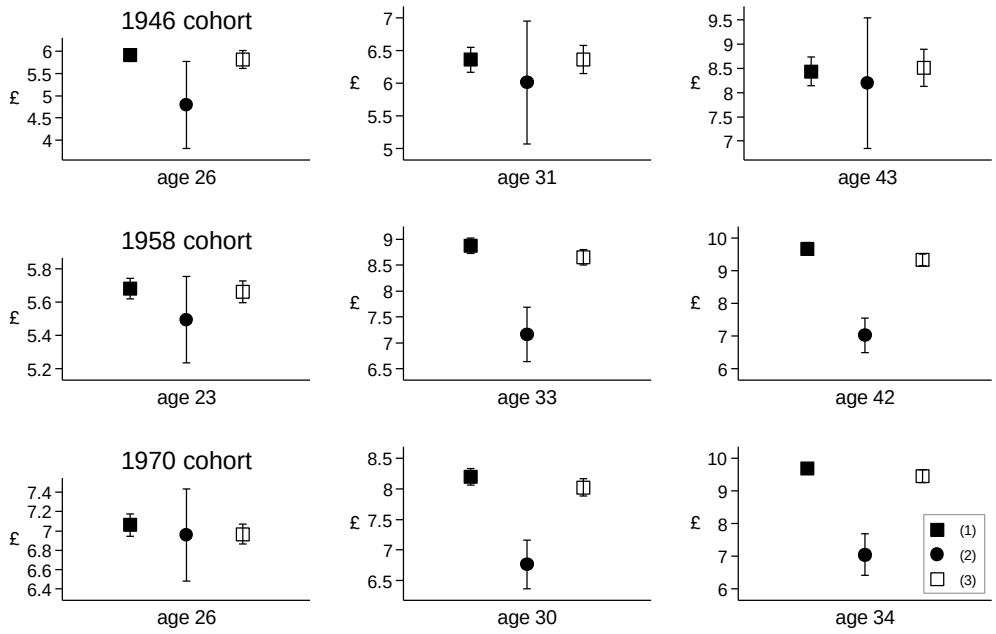
For women in their twenties and thirties, the size of the difference between the median imputed potential wage for non-employees and median observed wage for employees has increased across the three cohorts (Figure 9). The increase in the strength of the selection effect may be owing to the widening of the wage distribution over this period, as well as to increasing differences in the social and educational characteristics of employed and non-employed women. However, the aggregate effect of selectivity bias has reduced across the cohorts, as the fraction of women out of work has decreased. This effect shows up as a decreasing gap between the median observed wage for the employed population and the median potential wage for the whole population (Figure 9).

Figure 9: Medians of 1) observed wages for employees 2) imputed potential wages for non-employed and 3) potential wages for population (observed plus imputed), women by age and cohort



(1) Employees (observed)
(2) Non-employed (imputed)
(3) Population (observed plus imputed)

Figure 10: Medians of 1) observed wages for employees 2) imputed potential wages for non-employed and 3) potential wages for population (observed plus imputed), men by age and cohort

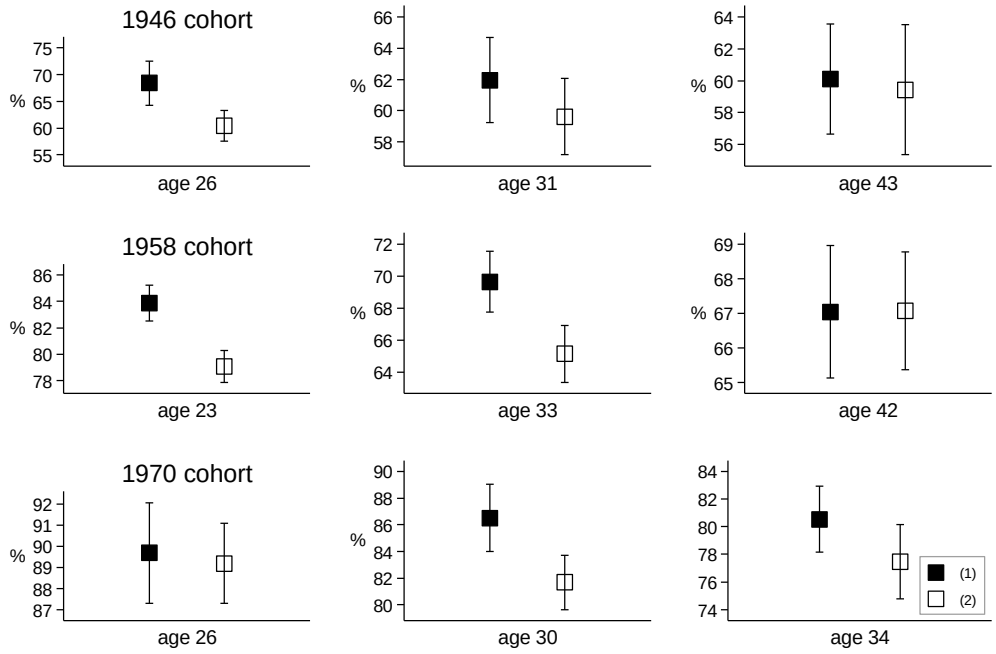


(1) Employees (observed)
(2) Non-employed (imputed)
(3) Population (observed plus imputed)

Only a small fraction of men were not employed for some period, although the potential wages of the non-employed minority are significantly lower than those of employees for men in the 1958 and 1970 cohorts in their thirties and forties (Figure 10). Owing to the small fraction out of work, the median potential wage for the whole male population is not significantly different from the median observed wage for employees at any of the surveys (Figure 10).

Taking the trends for women and men together, women's relative wage opportunities have improved across the three cohorts, moving in the same direction as observed wage trends (Figure 11). However, the increase in younger women's wage opportunities is greater than the observed increase in wages for this group. In particular, the ratio of women's to men's median potential wages is substantially and significantly lower than the observed wage ratio for the 1946 cohort at age 26, when only half of women were in work. This finding is in line with results obtained by Blundell et al. (2007).

Figure 11: Female-male ratio of medians for: 1) employees (observed wages); and 2) all women and men (imputed potential ratios), by age and cohort

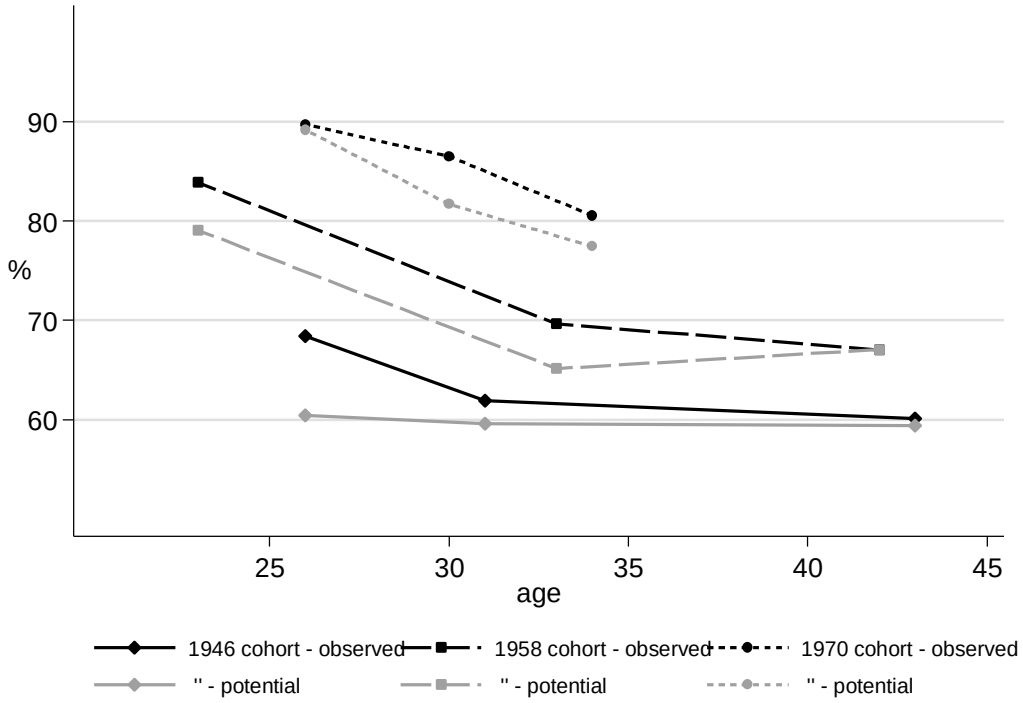


(1) Employees (observed)
(2) Population (observed plus imputed)

There is also weak evidence in the 1970 cohort that women's selective withdrawal and re-entry into the workforce around childbearing years conceals some of the decline in women's labour market position after childbirth. Between the ages of 26 and 30, around a fifth of women in the cohort had a first baby (Figure 4). The proportion of women in work fell from around 78% to around 74% (Figure 3) and the estimated potential wage gap between employed and non-employed women increased (Figure 9). Median hourly wages for working women decreased from

89.7% of men's to 86.5% of men's between these ages. Median potential wages for all women decreased more though, from 89.2% to 81.7% (Figure 12).

Figure 12: Women's median actual and potential hourly earnings as a % of men's at each age, by cohort



6 Conclusions

Women's position in the labour market has changed dramatically across three British generations since the 1970s, with a simultaneous increase in rates of employment and in relative rates of pay.

The analysis presented in this paper tackled the problem of how to assess changes in women's and men's wage opportunities, in light of simultaneous and interdependent changes in wage opportunities and employment participation. The approach taken was to impute potential wages for individuals not in work. These were imputed using the assumption that the potential wage of each non-employed individual would be equal to the observed wage of an employed individual of the same gender, cohort and age, with a similar educational, job and family history. The results indicate that the cross-cohort increase in younger women's wage opportunities, relative to men's, exceeds the observed increase in wages. Some evidence is also found to suggest that the loss of earning power for women after childbirth is understated in observed wage dynamics.

References

Angrist, J.D. and Krueger, A.B. Instrumental variables and the search for identification: from supply and demand to natural experiments. *Journal of Economic Perspectives*, 15(4):69-85, Fall 2001.

Arulampalam, W. Is unemployment really scarring? Effects of unemployment experiences on wages. *The Economic Journal*, 111(475):F585-F606, November 2001.

Blau, F.D. and Kahn, L.M. The U.S. Gender Pay Gap in the 1990s: slowing convergence. *Industrial and Labor Relations Review*, 60(1):45-66, October 2006.

Blanden, J., Gregg, P. and Macmillan, L. Accounting for intergenerational income persistence: noncognitive skills, ability and education. *The Economic Journal*, 117:C43-C60, March 2007.

Blundell, R., Dearden, L., Goodman, A. and Reed, H. The returns to higher education in Britain: evidence from a British cohort. *The Economic Journal*, 110:F82-F99, February 2000.

Blundell, R., Gosling, A., Ichimura, H. and Meghir, C. Changes in the distribution of male and female wages accounting for employment composition using bounds. *Econometrica*, 75(2):323-363, March 2007.

De Cooman, E, Ermisch, J. F. and Joshi, H. E. The Next Birth and the Labour Market: a dynamic econometric model of births in England and Wales. *Population Studies*, 41(2):237-268, 1987

Dickens, R. and Manning, A. Has the national minimum wage reduced UK wage inequality? *Journal of the Royal Statistical Society Series A*, 167(4):613-626, Statistics in Society, 2004.

Disney, R. and Webb, S. Why are there so many long term sick in Britain? *The Economic Journal*, 3(1):252-262, March 1991.

Dolton, P. J. and Makepeace, G. H. Marital status, child rearing and earnings differentials in the graduate labour market. *The Economic Journal*, 97(388):897-922, December 1987a.

Dolton, P. Marcenaro-Gutierrez, O. and Skalli, A. Gender wage differentials in Europe. In Peter Dolton, Rita Asplund, and Erling Barth, editors, *Education and inequality across Europe*, chapter 9. Edward Elgar Pub, 2009.

Durrant, G.B. Missing data methods in official statistics in the United Kingdom: some recent developments. *Allgemeines Statistiches Archiv*, 90:577-593, November 2006.

Ermisch, J.F. and Wright, R.E. Wage offers and full-time and part-time employment by British women. *The Journal of Human Resources*, 28(1):111-133, Winter 1993.

Faggio, G. and Nickell, S. The rise in inactivity among adult men. In Richard Dickens, Paul Gregg, and Jonathan Wadsworth, editors, *The labour market under New Labour: The state of working Britain*, chapter 3, pages 40-52. Palgrave Macmillan, 2003.

Ferri, E., Bynner, J. and Wadsworth, M. (editors) *Changing Britain, changing lives. Three generations at the turn of the century*, chapter 2, pages 29-70. Institute of Education, University of London, London, 2003.

Flouri, E. and Hawkes, D. Ambitious mothers - successful daughters: Mothers' early expectations for children's education and children's earnings and sense of control in adult life. *British Journal of Educational Psychology*, 78:411-433, 2008.

Gregory, M. and Jukes, R. Unemployment and subsequent earnings: Estimating scarring among British men 1984-94. *The Economic Journal*, 111(475):F607-F625, November 2001.

Gregg, P. The impact of youth unemployment on adult unemployment in the NCDS. *The Economic Journal*, 111(475):F626-F653, November 2001.

Gregg, P. Gutierrez-Domenech, M. and Waldfogel, J. The employment of married mothers in Great Britain, 1974-2000. *Economica*, 75(296):842-864, November 2007.

Gronau, R. Wage comparisons - a selectivity bias. *The Journal of Political Economy*, 82(6):1119-1143, December 1974.

Heckman, J.J. Sample selection bias as a specification error with an application to the estimation of labor supply functions. NBER Working Paper 172, National Bureau of Economic Research, March 1977.

Heckman, J.J. Sample selection bias as a specification error. *Econometrica*, 47(1): 153-162, January 1979.

Hobcraft, J. Fertility in England and Wales: A Fifty-Year Perspective *Population Studies*, Vol. 50, No. 3 pp. 485-524, 1996.

Jenkins, A., Joshi, H. and Killingsworth, M. Educational attainment, labour market conditions and unobserved heterogeneity: The timing of first and higher-order births in Britain. Working Paper 2008/12, Centre for Longitudinal Studies, Institute of Education, November 2008.

Joshi, H. Production, reproduction and education: women, children and work in a British perspective. *Population and Development Review*, 28(3):445-474, September 2002.

Joshi, H. and Paci, P. *Unequal pay for women and men: Evidence from the British Birth Cohort Studies.* The MIT Press, Cambridge, Massachusetts, with Gerald Makepeace and Jane Waldfogel, 1998.

Joshi, H.E., Layard, R. and Owen, S.J. Why are more women working in Britain? *Journal of Labor Economics*, 3(1):S147-S176, January 1985.

Joshi, H. Makepeace, G. and Dolton, P. More or less unequal? Evidence on the pay of men and women from the British Birth Cohort Studies. *Gender, Work and Organization*, 14(1):37-55, January 2007.

Juhn, C. Decline of male labor participation: the role of declining market opportunities. *The Quarterly Journal of Economics*, 107(1):79-121, February 1992.

Kneale, D. and Joshi, H. Postponement and childlessness: Evidence from two British cohorts. Demographic Research, 19(58):1936-1968, November 2008.

Kuh, D. and Wadsworth, M. Childhood influences on adult male earnings in a longitudinal study. *The British Journal of Sociology*, 42(4):537-555, December 1991.

Kuh, D., Head, J., Hardy, R. and Wadsworth, M. The influence of education and family background on women's earnings in midlife: evidence from a British national birth cohort study. British Journal of Sociology of Education, 18(3): 385-405, 1997.

Lewis, H.G. Comments on selectivity biases in wage comparisons. *The Journal of Political Economy*, 82(6):1145-1155, December 1974.

Little, R.J.A. and Rubin, D.B. *Statistical analysis with missing data*. John Wiley & Sons, Inc., New York, 2nd edition, 2002.

Macran, S., Dex, S. and Joshi, H. Employment after childbearing: A survival analysis. *Work, Employment and Society*, 10(2):273-296, June 1996.

Makepeace, G., Dolton, P., Woods, L., Joshi, H. and Galinda- Rueda, F. From school to the labour market. In Elsa Ferri, John Bynner, and Michael Wadsworth, editors, *Changing Britain, changing lives. Three generations at the turn of the century*, chapter 2, pages 29-70. Institute of Education, University of London, London, 2003.

Manning, A. The Equal Pay Act as an experiment to test theories of the labour market. *Economica*, 63(250):191-212, May 1996.

Manski, C.F. Anatomy of the selection problem. *The Journal of Human Resources*, 24(3):343-360, Summer 1989.

Mulligan, C.B, and Rubinstein, Y. Selection, investment and women's relative wages over time. *The Quarterly Journal of Economics*, 123(3):1061-1110, August 2008.

Nathan, G. *A review of sample attrition and representativeness in three longitudinal surveys.* Methodology Series No. 13, Government Statistical Service, 1999.

Neuburger, H. Unemployment: are wages to blame? a critique of the `priced-out' hypothesis. Discussion series, Low Pay Unit, 1984.

Olivetti, C. and Petrongolo, B. Unequal pay or unequal employment? a cross-country analysis of gender gaps*. Journal of Labor Economics*, 26(4):621-654, October 2008.

Plewis, I., Calderwood, L., Hawkes, D. and Nathan, G. National Child Development Survey and 1970 British Cohort Study Technical Report: Changes in the NCDS and BCS70 populations and samples over time. Centre for Longitudinal Studies, Institute of Education. October 2004.

Plewis, I. and Kallis, C. Changing economic circumstances in childhood and their effects on subsequent educational and other outcomes. Working Paper No 49, Department for Work and Pensions, February 2008.

Rosenbaum, P.R. and Rubin, D.B. The central role of the propensity score in observational studies for causal effects. *Biometrika,* 70(1):41-55, April 1983.

Rubin, D.B. *Multiple imputation for nonresponse in surveys*. John Wiley & Sons, Inc., Hoboken, New Jersey. Wiley series in probability and mathematic statistics, 1987.

Rubin, D.B. Inference and missing data. *Biometrika*, 63(3):581-592, December 1976.

Sande, I.G. Imputation in surveys: coping with reality. *The American Statistician*, 36(3):145-152, August 1982.

Skinner, C., Stuttard, N. Durrant, G.B. and Jenkins, J. The measurement of low pay in the UK Labour Force Survey. *Oxford Bulletin of Economics and Statistics*, 64:653-676, December 2002. Special Issue: Issue supplement.

Wadsworth, M.E.J. *The imprint of time: childhood history and adult life*. Oxford University Press, Oxford, UK, 1991.

Wadsworth, M.E.J., Butterworth, S.L., Kuh, D.J., Richards, M., Langenberg, C., Hilder, W.S. and Connor, M. The life course prospective design: an example of benefits and problems associated with study longevity. *Social Science & Medicine*, 57:2193-2205, 2003.

Zabalza, A. and Arrufat, J.L. The extent of sex discrimination in Great Britain. In A. Zabalza and Z. Tzannatos, editors, *Women and equal pay: The effects of legislation on female employment and wages in Britain*, chapter 5, pages 70-94. Cambridge University Press, 1985.

Zabalza,A. and Tzannatos, Z. *Women and equal pay: The effects of legislation on female employment and wages in Britain*. Cambridge University Press, 1985.

Tables

Table 1: Description of samples used for analysis of employment and earnings

Birth year	Survey year	Age	Survey method	Men (N)	Women (N)
1946	1972	26	Interview	1,768	1,712
1946	1977	31	Postal questionnaire	1,467	1,433
1946	1989	43	Interview	1,283	1,169
1958	1981	23	Interview	5,535	5,721
1958	1991	33	Interview	5,441	5,682
1958	2000	42	Interview	5,585	5,764
1970	1996	26	Postal questionnaire	3,367	3,836
1970	2000	30	Interview	5,360	5,735
1970	2004	34	Interview	4,588	5,014

Notes to table: The samples exclude cohort members who did not respond to the survey in question. They also exclude cohort members for whom data are missing on their qualifications, employment histories or, for women only, their childbearing histories.

Table 2: Numbers of observed and imputed potential wages

Birth cohort	Age at survey	Employed with observed wage	Imputed cases		
			Non-employed (1)	Self-employed	Employed, missing wage
Women					
1946	26	739	845	50	74
	31	630	684	70	124
	43	1,129	183	88	25
1958	23	3,400	1,947	87	449
	33	3,078	1,816	385	439
	42	3,999	1,207	426	135
1970	26	3,237	802	152	248
	30	3,894	1,459	261	83
	34	2,986	1,280	300	451
Men					
1946	26	1,463	74	163	119
	31	1,184	60	189	159
	43	1,105	51	285	27
1958	23	4,084	788	304	619
	33	3,697	503	865	436
	42	3,959	507	992	121
1970	26	2,779	385	333	254
	30	4,167	451	605	89
	34	3,270	319	658	342

(1) This group includes full-time carers and home workers, unemployed individuals, students and those not working due to long-term health conditions or for other reasons.

Table 3: Medians of observed and imputed potential wages

Birth cohort	Age at survey	Employed with observed wage	Imputed cases		
			Non-employed	Self-employed	Employed, missing wage
Women					
1946	26	4.06	3.33	4.23	3.84
	31	3.95	3.71	3.63	3.94
	43	5.07	4.35	5.14	5.87
1958	23	4.76	3.84	4.37	4.85
	33	6.26	4.73	5.74	6.14
	42	6.51	5.34	6.38	5.54
1970	26	6.33	5.43	6.46	6.44
	30	7.10	5.13	6.66	6.39
	34	7.83	5.38	7.47	7.66
Men					
1946	26	5.94	4.79	5.43	5.42
	31	6.29	6.01	6.43	6.82
	43	8.40	8.19	9.30	7.73
1958	23	5.69	5.50	5.77	5.66
	33	8.96	7.10	8.31	8.52
	42	9.70	7.02	9.10	8.94
1970	26	7.06	6.96	6.61	7.02
	30	8.21	6.76	7.72	7.17
	34	9.72	7.05	9.57	9.04

Table 4: Means of variables, women

	1946 cohort			1958 cohort			1970 cohort		
	26	31	43	23	33	42	26	30	34
Mother's age at birth at CM birth									
Youngest quartile	0.28	0.29	0.29	0.26	0.26	0.26	0.22	0.23	0.23
Second quartile	0.20	0.20	0.19	0.26	0.26	0.26	0.21	0.20	0.20
Third quartile	0.22	0.22	0.22	0.20	0.20	0.20	0.21	0.20	0.20
Oldest quartile	0.20	0.20	0.20	0.23	0.23	0.23	0.20	0.19	0.18
Father's social class									
I	0.05	0.05	0.06	0.05	0.05	0.05	0.06	0.05	0.06
II	0.15	0.16	0.15	0.13	0.13	0.13	0.17	0.16	0.17
III	0.17	0.18	0.19	0.09	0.09	0.09	0.08	0.07	0.07
IV	0.30	0.29	0.28	0.39	0.39	0.38	0.36	0.35	0.35
V & VI	0.30	0.20	0.20	0.15	0.15	0.15	0.10	0.10	0.10
Mother's schooling									
Left at 16 or younger	0.84	0.84	0.83	0.65	0.65	0.65	0.71	0.70	0.70
Left at 17	0.04	0.05	0.04	0.03	0.04	0.03	0.06	0.06	0.06
Left at 18	0.03	0.03	0.03	0.04	0.05	0.05	0.06	0.05	0.06
(variable missing)	0.09	0.08	0.09	0.27	0.26	0.27	0.17	0.19	0.18
Father's schooling									
Left at 16 or younger	0.80	0.81	0.79	0.63	0.62	0.62	0.66	0.65	0.65
Left at 17	0.05	0.06	0.05	0.03	0.03	0.03	0.05	0.05	0.05
Left at 18	0.05	0.05	0.05	0.06	0.06	0.06	0.08	0.07	0.07
(variable missing)	0.10	0.09	0.10	0.29	0.29	0.29	0.21	0.23	0.22
Siblings at age 16									
Only child	0.12	0.12	0.12	0.05	0.05	0.05	0.09	0.09	0.09
One sibling	0.30	0.30	0.31	0.22	0.22	0.22	0.44	0.41	0.42
Two or three siblings	0.34	0.34	0.33	0.31	0.32	0.32	0.32	0.33	0.33
Four or more siblings	0.16	0.14	0.14	0.16	0.15	0.15	0.04	0.05	0.04
Older siblings at age 16									
No older sibling	0.41	0.42	0.42	0.30	0.30	0.30	0.40	0.37	0.38
One older sibling	0.33	0.32	0.33	0.23	0.23	0.22	0.31	0.30	0.31

Table 4: Means of variables, women (continued)

	1946 cohort			1958 cohort			1970 cohort		
	26	31	43	23	33	42	26	30	34
Two or more	0.27	0.26	0.25	0.21	0.22	0.22	0.18	0.19	0.19
(siblings variables missing)	0.09	0.09	0.09	0.26	0.25	0.26	0.12	0.13	0.12
Maths score at age 11	0.04	0.08	0.07	0.03	0.05	0.05	0.10	0.01	0.05
Reading score at age 11	0.05	0.08	0.08	0.06	0.06	0.06	0.19	0.10	0.13
Missing score	0.12	0.12	0.11	0.14	0.14	0.14	0.24	0.25	0.24
Highest qualification (at current survey)									
None or below O-level	0.48	0.46	0.40	0.32	0.27	0.25	0.19	0.22	0.20
O-level or equivalent	0.25	0.27	0.26	0.37	0.38	0.31	0.42	0.30	0.26
A-level or equivalent	0.13	0.13	0.13	0.12	0.10	0.13	0.11	0.15	0.16
Diploma	0.09	0.09	0.10	0.10	0.14	0.16	0.10	0.15	0.18
Degree or higher	0.05	0.05	0.10	0.10	0.11	0.14	0.18	0.17	0.19
Social class of first job									
I	m	m	m	0.01	0.01	0.01	0.03	0.02	0.03
II	m	m	m	0.12	0.12	0.12	0.18	0.16	0.17
III	m	m	m	0.56	0.55	0.53	0.47	0.46	0.44
IV	m	m	m	0.09	0.09	0.09	0.11	0.11	0.10
V	m	m	m	0.18	0.18	0.17	0.16	0.19	0.18
VI	m	m	m	0.01	0.01	0.01	0.02	0.02	0.02
(variable missing)	m	m	m	0.03	0.04	0.07	0.05	0.03	0.06
Years in full-time work	5.0*	2.9	7.5	4.5	8.6	12.1	5.2	7.9	10.1
Years in part-time work	1.0*	0.6	4.6	0.1	1.7	4.5	0.5	1.3	2.5
Children in household	0.61	0.83	0.79	0.33	0.75	0.80	0.31	0.54	0.70
Children under five in hhld	0.55	m	0.04	0.30	0.40	0.10	0.26	0.40	0.38
More than one child in hhld	0.37	m	0.20	0.13	0.56	0.60	0.13	0.31	0.46
Living in London/SE	0.33	0.32	0.32	0.10	0.31	0.30	m	0.30	0.30
Sample size	1,710	1,431	1,167	5,721	5,682	5,764	3,836	5,735	5,014

Notes to table: Means for 1946 cohort are unweighted. Maths and reading scores are standardised for the sample of boys and girls who took the tests with mean 0 and standard deviation 1. A dummy variable is included if the test was not taken and scores are missing. The qualification variables represent the level of the highest qualification attained at the given survey. Academic or vocational qualifications are counted. The variables taking the value 1 if the qualification level is the highest attained and 0 otherwise. The variable `London/SE' is also a dummy variable taking the value 1 for individuals living in London or the South East and zero for those living in other parts of Britain. This was missing for the postal survey of the 1970 cohort carried out at age 26.

Table 5: Means of variables, men

	1946 cohort			1958 cohort			1970 cohort		
	26	31	43	23	33	42	26	30	34
Mother's age at birth at CM birth									
Youngest quartile	0.28	0.29	0.27	0.25	0.25	0.25	0.22	0.23	0.22
Second quartile	0.21	0.21	0.21	0.26	0.26	0.27	0.20	0.20	0.20
Third quartile	0.20	0.21	0.21	0.21	0.21	0.21	0.21	0.19	0.21
Oldest quartile	0.19	0.20	0.20	0.23	0.23	0.22	0.20	0.18	0.18
Father's social class									
I	0.05	0.05	0.06	0.05	0.05	0.05	0.07	0.06	0.06
II	0.15	0.16	0.16	0.13	0.14	0.13	0.18	0.16	0.17
III	0.17	0.18	0.18	0.09	0.09	0.09	0.08	0.07	0.08
IV	0.28	0.28	0.29	0.38	0.38	0.38	0.34	0.36	0.34
V & VI	0.20	0.20	0.20	0.15	0.15	0.15	0.08	0.22	0.22
Mother's schooling									
Left at 16 or younger	0.83	0.83	0.84	0.66	0.65	0.65	0.69	0.70	0.70
Left at 17	0.05	0.05	0.05	0.03	0.03	0.03	0.07	0.06	0.06
Left at 18	0.03	0.03	0.03	0.04	0.04	0.04	0.07	0.05	0.05
(variable missing)	0.09	0.09	0.09	0.26	0.28	0.28	0.18	0.19	0.19
Father's schooling									
Left at 16 or younger	0.78	0.78	0.79	0.63	0.62	0.62	0.65	0.65	0.65
Left at 17	0.05	0.06	0.06	0.03	0.03	0.03	0.06	0.05	0.06
Left at 18	0.06	0.05	0.05	0.06	0.06	0.06	0.08	0.07	0.07
(variable missing)	0.11	0.09	0.10	0.28	0.29	0.29	0.21	0.22	0.22
Siblings at age 16									
Only child	0.13	0.13	0.12	0.05	0.05	0.05	0.08	0.08	0.08
One sibling	0.31	0.31	0.32	0.23	0.23	0.23	0.43	0.40	0.41
Two or three siblings	0.33	0.33	0.33	0.31	0.31	0.31	0.34	0.34	0.34
Four or more siblings	0.15	0.14	0.14	0.15	0.14	0.15	0.03	0.04	0.04

Table 5: Means of variables, men (continued)

	1946 cohort			1958 cohort			1970 cohort		
	26	31	43	23	33	42	26	30	34
Older siblings at age 16									
No older sibling	0.43	0.43	0.42	0.30	0.29	0.29	0.38	0.36	0.36
One older sibling	0.32	0.32	0.32	0.24	0.24	0.24	0.32	0.32	0.32
Two or more	0.26	0.25	0.26	0.21	0.20	0.20	0.18	0.19	0.19
(Siblings variables missing)	0.09	0.09	0.08	0.15	0.27	0.27	0.12	0.13	0.13
Maths score at age 11	0.01	0.03	0.07	0.10	0.10	0.10	0.22	0.10	0.14
Reading score at age 11	-0.02	-0.01	0.02	0.07	0.08	0.07	0.14	0.02	0.05
Missing score	0.12	0.12	0.11	0.14	0.14	0.14	0.25	0.25	0.25
Highest qualification (at current survey)									
None or below O-level	0.45	0.44	0.38	0.27	0.22	0.21	0.22	0.21	0.19
O-level or equivalent	0.14	0.15	0.14	0.30	0.31	0.28	0.33	0.24	0.21
A-level or equivalent	0.16	0.15	0.17	0.23	0.18	0.18	0.14	0.23	0.24
Diploma	0.12	0.12	0.12	0.09	0.14	0.17	0.09	0.13	0.15
Degree or higher	0.14	0.14	0.19	0.10	0.14	0.16	0.22	0.19	0.21
Social class of first job									
I	m	m	m	0.04	0.04	0.04	0.06	0.05	0.05
II	m	m	m	0.10	0.10	0.10	0.19	0.17	0.18
III	m	m	m	0.18	0.18	0.17	0.19	0.18	0.17
IV	m	m	m	0.39	0.38	0.37	0.27	0.32	0.29
V	m	m	m	0.16	0.16	0.15	0.17	0.18	0.18
VI	m	m	m	0.08	0.07	0.07	0.04	0.06	0.05
(variable missing)	m	m	m	0.05	0.06	0.11	0.07	0.05	0.08
Years in full-time work	11.3*	5.8	17.2	5.4	13.4	21.0	5.0	10.4	14.3
Years in part-time work	0.0*	0.0	0.1	0.0	0.1	0.7	0.1	0.2	0.3
Children in hhld	0.40	m	0.75	0.16	0.62	0.70	0.17	0.36	0.53
Children under five	0.37	m	0.09	0.15	0.41	0.15	0.14	0.29	0.36
More than one child	0.19	m	0.17	0.05	0.43	0.54	0.06	0.20	0.33
Living in London/SE	0.32	0.33	0.31	0.10	0.30	0.29	m	0.30	0.30
Sample size	1,768	1,467	1,283	5,535	5,441	5,585	3,367	5,360	4,588

Table 6: Summary of probit models used to estimate propensity scores, women

	1946 cohort			1958 cohort			1970 cohort		
	26	31	43	23	33	42	26	30	34
Father in non-manual work at birth (weighting factor)	-0.01 (0.03)	-0.04 (0.03)	0.01 (0.04)	n/a	n/a	n/a	n/a	n/a	n/a
Mother's age at birth in quartiles (youngest = ref)									
Second quartile	-0.17 (0.12)	0.15 (0.11)	0.02 (0.13)	-0.01 (0.06)	0.04 (0.05)	-0.06 (0.05)	0.07 (0.07)	0.09 (0.06)	0.11 (0.06)
Third quartile	-0.14 (0.13)	0.04 (0.11)	-0.07 (0.13)	-0.07 (0.06)	0.03 (0.06)	-0.07 (0.06)	0.04 (0.07)	0.09 (0.06)	0.11 (0.06)
Oldest quartile	-0.21 (0.14)	0.09 (0.12)	-0.01 (0.14)	-0.04 (0.07)	0.03 (0.06)	-0.01 (0.06)	0.07 (0.08)	0.03 (0.07)	0.12 (0.07)
Father's social class (V & VI = ref)									
I	0.08 (0.23)	-0.10 (0.21)	-0.01 (0.23)	0.24 (0.10)	-0.05 (0.10)	0.17 (0.10)	0.19 (0.13)	0.12 (0.11)	0.13 (0.11)
II	0.03 (0.14)	0.12 (0.13)	0.11 (0.15)	0.06 (0.08)	0.04 (0.07)	0.11 (0.07)	0.13 (0.09)	0.07 (0.08)	0.01 (0.08)
III	0.15 (0.14)	-0.23 (0.13)	0.07 (0.15)	-0.03 (0.08)	0.03 (0.07)	-0.02 (0.08)	0.00 (0.11)	0.12 (0.10)	0.07 (0.09)
IV	-0.05 (0.12)	-0.12 (0.10)	0.14 (0.12)	0.02 (0.06)	-0.09 (0.05)	-0.03 (0.05)	0.03 (0.08)	0.11 (0.07)	0.00 (0.06)
Mother's schooling (minimum = ref)									
Left at 17	-0.12 (0.21)	-0.12 (0.19)	-0.01 (0.23)	-0.05 (0.11)	0.05 (0.10)	-0.03 (0.11)	-0.04 (0.10)	-0.17 (0.09)	0.05 (0.08)
Left at 18	0.09 (0.26)	0.23 (0.25)	0.01 (0.26)	0.10 (0.10)	-0.01 (0.10)	0.08 (0.10)	0.00 (0.11)	-0.02 (0.10)	-0.10 (0.09)
(variable missing)	0.82 (0.31)	0.74 (0.31)	0.02 (0.33)	0.06 (0.16)	0.00 (0.15)	-0.23 (0.15)	-0.53 (0.62)	0.54 (0.37)	0.16 (0.42)
Father's schooling (minimum = ref)									
Left at 17	0.22 (0.20)	0.00 (0.19)	0.26 (0.32)	-0.15 (0.12)	-0.05 (0.11)	0.05 (0.12)	0.03 (0.11)	0.26 (0.17)	0.04 (0.09)
Left at 18	-0.16 (0.22)	-0.08 (0.20)	0.14 (0.21)	-0.05 (0.09)	0.04 (0.09)	0.04 (0.09)	0.05 (0.10)	-0.11 (0.10)	0.04 (0.09)
(variable missing)	-0.64 (0.30)	-0.70 (0.31)	0.26 (0.21)	-0.09 (0.12)	-0.16 (0.10)	0.11 (0.11)	0.20 (0.19)	0.10 (0.09)	-0.03 (0.17)
Siblings (4+ = ref)									
Only child	0.03 (0.19)	0.01 (0.17)	0.07 (0.20)	0.00 (0.11)	0.19 (0.11)	-0.04 (0.11)	-0.25 (0.16)	-0.12 (0.11)	0.12 (0.13)
One sibling	0.30 (0.16)	0.07 (0.14)	0.12 (0.16)	-0.03 (0.08)	0.12 (0.07)	0.09 (0.08)	-0.22 (0.14)	-0.14 (0.13)	0.15 (0.11)
Two or three siblings	0.26 (0.13)	0.05 (0.12)	0.09 (0.14)	0.04 (0.07)	0.14 (0.06)	0.07 (0.06)	-0.21 (0.13)	-0.14 (0.11)	0.10 (0.10)
(variable missing)	-0.10 (0.25)	0.29 (0.23)	-0.31 (0.26)	0.00 (0.16)	0.29 (0.15)	d	-0.27 (0.14)	-0.12 (0.11)	0.12 (0.11)
Older siblings (ref = 2 +)									
No older sibling	-0.11 (0.14)	-0.07 (0.13)	-0.05 (0.15)	-0.03 (0.08)	-0.05 (0.07)	0.13 (0.15)	0.01 (0.09)	-0.02 (0.08)	-0.03 (0.07)
One older sibling	-0.04 (0.13)	0.07 (0.11)	0.10 (0.13)	-0.06 (0.07)	0.00 (0.06)	-0.11 (0.07)	0.07 (0.09)	0.03 (0.07)	-0.06 (0.07)

	1946 cohort			1958 cohort			1970 cohort		
	26	31	43	23	33	42	26	30	34
Maths score at age 11	-0.11 (0.07)	-0.08 (0.06)	-0.02 (0.07)	-0.08 (0.03)	-0.07 (0.03)	-0.02 (0.03)	-0.10 (0.04)	-0.06 (0.03)	-0.03 (0.03)
Reading score at age 11	0.10 (0.07)	0.12 (0.06)	0.02 (0.07)	0.10 (0.04)	0.03 (0.03)	0.06 (0.03)	0.01 (0.04)	0.03 (0.03)	-0.02 (0.03)
Missing score	0.19 (0.17)	0.06 (0.16)	0.28 (0.18)	-0.06 (0.06)	-0.07 (0.05)	0.00 (0.05)	0.03 (0.06)	0.02 (0.05)	0.04 (0.05)
Highest qualification (no quals = ref)									
O-level or equivalent	0.06 (0.12)	-0.00 (0.10)	-0.29 (0.12)	-0.17 (0.06)	-0.03 (0.05)	0.04 (0.05)	-0.18 (0.06)	-0.08 (0.06)	-0.11 (0.06)
A-level or equivalent	-0.16 (0.15)	0.20 (0.13)	-0.17 (0.15)	-0.41 (0.08)	-0.05 (0.07)	0.02 (0.07)	-0.12 (0.09)	0.02 (0.07)	-0.12 (0.07)
Diploma	0.07 (0.17)	0.35 (0.16)	-0.52 (0.18)	-0.47 (0.09)	-0.21 (0.07)	-0.11 (0.07)	-0.41 (0.10)	-0.27 (0.07)	-0.38 (0.07)
Degree or higher	0.18 (0.22)	0.11 (0.21)	-0.06 (0.17)	-0.72 (0.11)	-0.56 (0.09)	-0.29 (0.08)	-0.15 (0.10)	-0.75 (0.08)	-0.68 (0.07)
Social class of first job (V = ref)									
I	m	m	m	0.01 (0.16)	0.57 (0.17)	0.40 (0.16)	-0.59 (0.18)	-0.44 (0.15)	-0.12 (0.13)
II	m	m	m	-0.21 (0.08)	-0.04 (0.07)	-0.08 (0.08)	-0.34 (0.08)	-0.26 (0.07)	-0.15 (0.07)
III	m	m	m	0.03 (0.06)	0.01 (0.05)	0.01 (0.05)	-0.20 (0.07)	-0.22 (0.05)	-0.12 (0.05)
IV	m	m	m	0.23 (0.08)	0.11 (0.07)	-0.05 (0.08)	0.10 (0.09)	0.11 (0.07)	0.18 (0.07)
VI	m	m	m	0.19 (0.23)	-0.12 (0.20)	-0.15 (0.21)	-0.10 (0.18)	-0.39 (0.15)	-0.26 (0.15)
(variable missing)	m	m	m	1.28 (0.21)	0.05 (0.11)	-0.33 (0.08)	0.09 (0.12)	0.46 (0.14)	-0.01 (0.09)
Years in full-time emp.	-0.14 (0.01)	-0.18 (0.02)	-0.10 (0.01)	-0.23 (0.01)	-0.10 (0.04)	-0.09 (0.00)	-0.05 (0.01)	-0.15 (0.01)	-0.10 (0.01)
Years in part-time emp.	-0.20 (0.01)	-0.16 (0.03)	-0.14 (0.01)	-0.57 (0.04)	-0.16 (0.01)	-0.11 (0.00)	-0.09 (0.02)	-0.21 (0.01)	-0.12 (0.01)
Children in hhld	0.26 (0.17)	0.73 (0.11)	-0.09 (0.12)	0.42 (0.15)	0.18 (0.06)	-0.09 (0.06)	0.74 (0.14)	0.26 (0.07)	0.03 (0.06)
Children under five	0.92 (0.16)	d	1.10 (0.21)	1.16 (0.15)	0.64 (0.04)	0.67 (0.06)	0.14 (0.11)	0.62 (0.06)	0.50 (0.05)
More than one child	-0.20 (0.12)	d	-0.09 (0.11)	-0.02 (0.09)	-0.01 (0.05)	-0.07 (0.05)	0.33 (0.08)	0.06 (0.05)	0.12 (0.05)
Living in London/SE	0.02 (0.09)	0.30 (0.08)	-0.06 (0.09)	-0.05 (0.07)	0.10 (0.04)	0.13 (0.04)	m	0.15 (0.04)	0.14 (0.04)
Constant term	0.70 (0.23)	0.32 (0.21)	0.77 (0.25)	0.81 (0.13)	0.64 (0.11)	1.07 (0.11)	-0.15 (0.16)	0.83 (0.14)	0.88 (0.14)
Pseudo R-squared	0.46	0.16	0.16	0.34	0.16	0.15	0.15	0.26	0.15
Sample size	1,710	1,431	1,167	5,721	5,682	5,764	3,836	5,735	5,014

Notes to table: Estimated coefficients from nine separate probits are shown, with standard errors shown in the brackets. These model the probability of being non-employed, self-employed or having a missing wage, on the one hand, versus being employed with an observed wage. The coefficients are close to those estimated from probits modelling just the probability of non-employment vs. employment m indicates that variables are missing. d indicates that a variable was dropped from the model owing to collinearity

Table 7: Summary of probit models used to estimate propensity scores, men

	1946 cohort			1958 cohort			1970 cohort		
	26	31	43	23	33	42	26	30	34
Father in non-manual work at birth (weighting factor)	-0.04 (0.03)	0.00 (0.03)	0.02 (0.03)	n/a	n/a	n/a	n/a	n/a	n/a
Mother's age at birth by quartile (youngest = ref)									
Second quartile	0.01 (0.11)	0.00 (0.11)	-0.08 (0.12)	-0.06 (0.05)	-0.10 (0.05)	-0.04 (0.05)	0.07 (0.07)	-0.03 (0.06)	0.06 (0.06)
Third quartile	-0.14 (0.12)	0.02 (0.11)	-0.09 (0.12)	0.00 (0.06)	-0.09 (0.06)	-0.16 (0.06)	0.01 (0.08)	-0.01 (0.07)	0.03 (0.07)
Oldest quartile	0.03 (0.12)	-0.06 (0.12)	-0.05 (0.13)	0.05 (0.06)	-0.05 (0.06)	-0.05 (0.06)	0.13 (0.08)	-0.02 (0.07)	-0.06 (0.07)
Father's social class (V & VI = ref)									
I	-0.07 (0.20)	-0.06 (0.19)	0.00 (0.20)	0.04 (0.10)	0.14 (0.10)	0.20 (0.10)	-0.24 (0.13)	0.03 (0.11)	0.22 (0.11)
II	0.13 (0.12)	0.08 (0.12)	0.05 (0.13)	0.16 (0.07)	0.26 (0.07)	0.23 (0.07)	-0.08 (0.09)	0.11 (0.08)	0.22 (0.08)
III	0.04 (0.12)	-0.06 (0.12)	-0.13 (0.14)	0.04 (0.08)	-0.03 (0.08)	-0.05 (0.08)	-0.22 (0.11)	-0.20 (0.10)	-0.04 (0.10)
IV	0.23 (0.10)	-0.04 (0.10)	-0.05 (0.11)	0.05 (0.05)	0.06 (0.05)	0.09 (0.05)	-0.13 (0.08)	-0.12 (0.06)	0.02 (0.07)
Mother's schooling (minimum = ref)									
Left at 17	-0.13 (0.19)	0.25 (0.18)	-0.38 (0.22)	0.00 (0.10)	-0.12 (0.11)	0.14 (0.11)	0.04 (0.10)	0.02 (0.09)	0.00 (0.09)
Left at 18	-0.20 (0.25)	0.12 (0.24)	0.33 (0.25)	-0.02 (0.10)	0.15 (0.10)	0.13 (0.11)	0.05 (0.11)	0.14 (0.10)	-0.08 (0.10)
(variable missing)	0.05 (0.25)	-0.09 (0.27)	-0.06 (0.30)	0.11 (0.15)	0.15 (0.14)	0.12 (0.15)	0.08 (0.37)	0.09 (0.30)	-0.21 (0.32)
Father's schooling (minimum = ref)									
Left at 17	-0.09 (0.17)	-0.26 (0.17)	0.28 (0.18)	0.09 (0.11)	0.15 (0.11)	0.14 (0.12)	0.13 (0.11)	-0.13 (0.10)	-0.08 (0.10)
Left at 18	0.10 (0.19)	0.11 (0.18)	0.11 (0.20)	0.07 (0.09)	-0.09 (0.09)	-0.05 (0.09)	0.13 (0.10)	-0.25 (0.10)	-0.06 (0.09)
(variable missing)	0.22 (0.25)	-0.09 (0.25)	-0.35 (0.28)	0.08 (0.12)	-0.06 (0.12)	0.09 (0.12)	0.20 (0.20)	0.27 (0.16)	-0.02 (0.17)
Siblings, age 16 (4+ = ref)									
Only child	-0.19 (0.17)	-0.17 (0.17)	0.11 (0.20)	0.07 (0.11)	-0.02 (0.11)	0.09 (0.11)	-0.33 (0.17)	0.12 (0.14)	0.22 (0.14)
One sibling	-0.07 (0.14)	-0.16 (0.14)	0.26 (0.15)	0.12 (0.08)	0.00 (0.08)	0.10 (0.08)	-0.32 (0.15)	0.00 (0.12)	0.10 (0.12)

	1946 cohort			1958 cohort			1970 cohort		
	26	31	43	23	33	42	26	30	34
Two or three siblings	-0.13 (0.12)	0.04 (0.12)	0.13 (0.13)	0.12 (0.06)	0.04 (0.06)	0.03 (0.06)	-0.27 (0.14)	0.04 (0.12)	0.16 (0.11)
(variable missing)	0.25 (0.22)	0.30 (0.22)	0.36 (0.24)	-0.05 (0.14)	-0.05 (0.13)	-0.11 (0.14)	-0.24 (0.15)	0.06 (0.11)	0.09 (0.12)
Older siblings (ref = 2 +)									
No older sibling	0.00 (0.13)	0.08 (0.13)	-0.06 (0.14)	-0.02 (0.07)	-0.06 (0.07)	-0.11 (0.07)	0.09 (0.09)	-0.10 (0.08)	-0.17 (0.07)
One older sibling	0.07 (0.11)	0.12 (0.11)	0.14 (0.12)	0.04 (0.07)	0.02 (0.07)	0.05 (0.07)	0.14 (0.09)	-0.05 (0.07)	-0.05 (0.07)
Maths score at age 11	-0.02 (0.06)	-0.14 (0.05)	0.01 (0.06)	0.07 (0.03)	-0.04 (0.03)	-0.03 (0.03)	-0.10 (0.04)	-0.02 (0.03)	-0.02 (0.03)
Reading score at age 11	0.02 (0.05)	0.05 (0.05)	0.01 (0.06)	-0.09 (0.03)	-0.05 (0.03)	-0.04 (0.03)	0.02 (0.04)	-0.01 (0.03)	-0.01 (0.03)
Missing score	-0.31 (0.17)	-0.35 (0.17)	-0.04 (0.17)	-0.07 (0.05)	-0.02 (0.05)	0.02 (0.05)	0.02 (0.06)	0.11 (0.05)	0.05 (0.05)
Highest qualification (no quals = ref)									
O-level or equivalent	0.01 (0.12)	-0.04 (0.12)	-0.05 (0.13)	-0.15 (0.05)	0.01 (0.05)	0.10 (0.06)	-0.14 (0.07)	-0.07 (0.06)	0.08 (0.06)
A-level or equivalent	-0.06 (0.12)	0.00 (0.12)	-0.05 (0.12)	-0.29 (0.06)	-0.12 (0.06)	0.09 (0.06)	-0.29 (0.09)	-0.18 (0.06)	-0.01 (0.06)
Diploma	-0.11 (0.13)	-0.32 (0.13)	-0.27 (0.14)	-0.34 (0.08)	-0.31 (0.07)	-0.14 (0.07)	-0.33 (0.10)	-0.44 (0.08)	-0.18 (0.07)
Degree or higher	-0.29 (0.15)	0.01 (0.14)	-0.26 (0.14)	-0.81 (0.10)	-0.64 (0.08)	-0.48 (0.08)	-0.26 (0.09)	-0.84 (0.08)	-0.48 (0.08)
Social class of first job (V = ref)									
I	m	m	m	-0.29 (0.10)	0.05 (0.11)	0.15 (0.11)	-0.28 (0.12)	-0.20 (0.12)	-0.16 (0.11)
II	m	m	m	-0.14 (0.08)	-0.26 (0.08)	-0.07 (0.08)	-0.30 (0.08)	-0.20 (0.07)	-0.10 (0.07)
III	m	m	m	-0.17 (0.06)	-0.24 (0.06)	-0.17 (0.08)	-0.17 (0.08)	-0.22 (0.07)	-0.15 (0.07)
IV	m	m	m	0.09 (0.06)	-0.02 (0.05)	0.00 (0.06)	0.08 (0.07)	0.11 (0.06)	0.22 (0.06)
VI	m	m	m	0.02 (0.08)	-0.06 (0.08)	-0.11 (0.08)	0.48 (0.12)	0.13 (0.09)	0.26 (0.10)
(variable missing)	m	m	m	0.46 (0.10)	-0.15 (0.09)	-0.33 (0.08)	0.15 (0.10)	0.17 (0.10)	0.16 (0.08)
Years in full-time emp.	-0.18 (0.02)	-0.18 (0.05)	-0.18 (0.02)	-0.25 (0.01)	-0.10 (0.01)	-0.07 (0.00)	-0.02 (0.01)	-0.12 (0.01)	-0.08 (0.01)

	1946 cohort			1958 cohort			1970 cohort		
	26	31	43	23	33	42	26	30	34
Years in part-time emp.	-0.03 (0.05)	0.08 (0.13)	-0.18 (0.05)	-0.17 (0.06)	-0.06 (0.03)	-0.03 (0.01)	-0.01 (0.04)	-0.07 (0.02)	0.00 (0.02)
Children in hhld	-0.19 (0.26)	d	0.08 (0.10)	0.16 (0.20)	0.01 (0.06)	-0.18 (0.06)	0.14 (0.15)	-0.02 (0.09)	-0.03 (0.07)
Children under five	0.30 (0.25)	d	0.04 (0.14)	-0.09 (0.20)	-0.03 (0.05)	0.08 (0.05)	-0.31 (0.15)	-0.01 (0.08)	-0.08 (0.06)
More than one child	0.13 (0.11)	d	-0.05 (0.11)	-0.07 (0.10)	0.04 (0.05)	0.08 (0.05)	0.16 (0.12)	0.13 (0.07)	0.12 (0.06)
Living in London/SE	0.07 (0.08)	0.09 (0.08)	-0.03 (0.09)	-0.01 (0.06)	0.07 (0.04)	0.04 (0.04)	m	0.08 (0.04)	0.07 (0.04)
Constant term	1.24 (0.24)	0.45 (0.31)	2.43 (0.44)	0.93 (0.12)	1.04 (0.12)	1.02 (0.12)	0.02 (0.17)	0.86 (0.15)	0.45 (0.16)
Pseudo R-squared	0.11	0.04	0.08	0.10	0.07	0.08	0.04	0.10	0.07
Sample size	1,768	1,467	1,283	5,535	5,441	5,585	3,367	5,360	4,588

Notes to table: Estimated coefficients from nine separate probits are shown, with standard errors shown in the brackets. These model the probability of being non-employed, self-employed or having a missing wage, on the one hand, versus being employed with an observed wage. The coefficients are close to those estimated from probits modelling just the probability of non-employment vs. employment, owing to the heterogeneity amongst self-employees. m indicates that variables are missing. d indicates that a variable was dropped from the model owing to collinearity

Table 8: Wage questions at each survey for the 1946 cohort (referring to main job only)

Year	Age	Pay questions	Pay period	Hours question
1972	26	`Including all regular payments such as overtime, bonuses etc..., how much do you earn in a typical week or month before deductions for tax, national insurance etc..?'  `In your job is there a basic or standard rate of pay? If yes, how much is this, before deductions for tax, national insurance etc...'	(per week, per calendar month, per four weeks, other)	`how many hours would you say you work in a typical week, excluding over-time?'  `Do you ever work paid overtime hours?' `how much overtime do you work regularly and how much occasionally?'
1977	31	`On average, how much do you earn a week? (including overtime and other payment) before deductions'		`How many hours a week do you usually work including overtime?'
1989	43	`Would you mind telling me which of the letters on this card represents your own average gross earnings, before deduction of income tax and national insurance?'	26 wage bands shown in annual, monthly and weekly amounts	`How many hours a week on average do you have to work to earn this amount?'  `How many months a year on aver-age do you have to work to earn this amount? (if in part-time or seasonal work)' `Last week (or last full working week) how many hours did you actively spend working including overtime and working at home?'

Table 9: Wage questions at each survey for the 1958 cohort (referring to main job only)

Year	Age	Pay questions	Pay period	Hours question
1981	23	'On the last occasion what was your pay before deductions for tax and National Insurance: including any overtime, bonus, commission, tips? (if last occasion was usual amount)' 'And what is your usual pay before any deductions for tax and National Insurance: including any overtime, bonus, commission, tips, etc., that you usually receive? (if last pay was unusual)	per day, per week, per 2 weeks, per month, per 3 months, per 6 months, per year, other	'How many hours of paid work do you actually do in an average week – including any paid overtime you usually do, but excluding meal breaks?'
1991	33	'What is your usual gross pay before deductions?' 'Last time you were paid, what was your gross pay before deductions?' (including overtime, bonuses, commission and tips)'	'How long a period does that pay cover?' (1 week, fortnight, four weeks, calendar month, year, other)	'How many hours a week do you usually work for that pay, excluding meal breaks but including paid overtime?'
2000	42	'Last time you were paid, what was your gross pay before deductions?' (including overtime, bonuses, commission and tips)	as above	'(still thinking of your main job) Do you ever do any work which you would regard as paid or unpaid overtime?' 'How many hours per week do you usually work in your (main) job/business not including meal breaks? (if no overtime)' 'How many hours a week do you usually work not including meal breaks and overtime (if overtime)' 'How many hours paid overtime do you usually work per week?'

Table 10: Wage questions at each survey for the 1970 cohort (referring to main job only)

Year	Age	Pay questions	Pay period	Hours question
1996	26	'What is your usual take home pay (after deductions, but including any bonuses or overtime)? Please write in amount'	'tick one box for period covered' (hour, day, week, month, year, other period)	'How many hours do you usually work each week? Please include any paid overtime you usually do, but exclude meal breaks'
1991	33	'Last time you were paid, what was your gross pay before deductions?' (including overtime, bonuses, commission and tips)'	'How long a period does that pay cover?' (1 week, fortnight, four weeks, calendar month, year, other)	'(still thinking of your main job) Do you ever do any work which you would regard as paid or unpaid overtime?' 'How many hours per week do you usually work in your (main) job/business not including meal breaks? (if no overtime)' 'How many hours a week do you usually work not including meal breaks and overtime (if overtime)' 'How many hours paid overtime do you usually work per week?'
2000	42	as above	as above	as above

Centre for Longitudinal Studies
Institute of Education
20 Bedford Way
London WC1H 0AL
Tel: 020 7612 6860
Fax: 020 7612 6880
Email cls@ioe.ac.uk
Web http://www.cls.ioe.ac.uk