Labour Economics 91 (2024) 102614

Contents lists available at ScienceDirect

Labour Economics

journal homepage: www.elsevier.com/locate/labeco

Gender wage gap among young adults: A comparison across British cohorts ☆

Francesca Foliano *, Alex Bryson, Heather Joshi, Bożena Wielgoszewska, David Wilkinson

Social Research Institute, UCL, United Kingdom

ARTICLE INFO

JEL classification:
J16
J2
J3

Keywords:
Gender wage gap
Birth cohorts
Employment selection
Graduates
Occupational segregation

ABSTRACT

We study the evolution of the gender wage gap among young adults in Britain between 1972 and 2015 using data from four British cohorts born in 1946, 1958, 1970 and 1989/90 on early life factors, human capital, family formation and job characteristics. We account for non-random selection of men and women into the labour market and compare the gender wage gap among graduates and non-graduates. The raw and covariate-adjusted gender wage gaps at the mean decline over the period among non-graduates, but they rise among young graduates. The gender wage gap across the wage distribution narrows over time for lower wages. Allowing for positive selection into employment increases the size of the gender wage gap in earlier cohorts, but selection is not apparent in the two most recent cohorts. Thus the rate of convergence in the wages of young men and women is understated when estimates do not account for positive selection in earlier cohorts. Differences in traditional human capital variables explain only a very small component of the gender wage gaps among young people in all four cohorts, but occupational gender segregation plays an important role in the later cohorts.

1. Introduction

In the UK gender inequalities in wages appear early in working life and gradually widen over the life course (Manning and Swaffield, 2008; Bryson et al., 2020; Benny et al., 2021), especially with the arrival of children (Costa Dias et al., 2020; Kleven et al., 2019). For most workers employment in early adulthood precedes the acquisition of family responsibilities, therefore a study of the gender wage gap at the beginning of the working life can help understand some of the mechanisms through which the gender wage gap first appears and creates path dependence over the subsequent ages.

In recent cohorts, women have surpassed men in terms of their academic attainment while they delayed the birth of the first child and increased their participation in the labour market (Albanesi et al., 2022). At the same time a series of policies, the Equal Pay Act of 1970 implemented in 1976 with other Equal Opportunities legislation, the Equal Pay Amendment Act of 1980, the National Minimum Wage Act 1998 (NMW) and the Equality Act 2010, aimed to reduce discrimination and should have created a more favourable environment for gender pay equity for later cohorts of workers. As a consequence, studying the evolution of the gender wage gap and its main sources across successive cohorts of young adults can shed light on how secular demographic, social and economic shifts, including changing non-random selection into the labour market, may have affected the gender disparity in earnings over time while legislation likely contributed to reduce discrimination.

The gender wage gap for all workers in the UK, as for most developed countries, has substantially decreased over time (Kunze, 2018). While we might anticipate a similar decline among young adults, convergence in qualifications and experience is unlikely to have affected the gender wage gap among young people because previous studies find that human capital variables explain only a small proportion of the earnings gap among this demographic (Fortin, 2008; Combet and Oesch, 2019). Indeed, some research suggests gender convergence in earnings may be heterogeneous across the skill and wage distribution. Goldin (2021) shows that in the United States graduate women's earnings stopped catching up with graduate men's earnings after 1990. In Europe, research shows that the gender wage gap is larger at the top of the earnings distribution than at the bottom, but that, in some cases, it has widened among low earners, where women are concentrated (Albrecht et al., 2003; Arulampalam et al., 2007).

☆ The project has been funded by the ESRC, United Kingdom (grant number ES/S012583/1). We thank the participants at the Royal Economic Society 2021 conference, the QSS Seminar at UCL, the WPEG 2022 conference in Sheffield, the on-line ISER policy event on Gender Wage Gap, the EALE 2022 conference in Padova, the AIEL 2022 conference in Salerno, the GRAPE 2023 conference in Warsaw and Emma Duchini and Ian Walker for useful comments. We also thank the MRC Lifelong Health and Ageing Unit UCL for access to NSHD data. We are grateful to the editor and four anonymous referees for their careful and insightful comments.

* Corresponding author.
E-mail address: f.foliano@ucl.ac.uk (F. Foliano).

https://doi.org/10.1016/j.labeco.2024.102614
Received 7 March 2023; Received in revised form 22 July 2024; Accepted 26 July 2024
Available online 16 August 2024
0927-5371/© 2024 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

A complicating factor in the study of the gender convergence in earnings is that the increasing labour market participation of young women may have changed the profile of their selection into employment. It is important to consider any changing sample selection when comparing the observed wages of men and women over time. Non-random selection into employment may differ by stage of life and across cohorts. If women are more positively selected into employment than men, as earlier research for Britain suggested (Neuburger et al., 2011), failure to account for this bias leads to an underestimation of the gender wage gap. If the positive selection into employment has diminished over time, as Bryson et al. (2020) show for workers born in 1958 and 1970, failure to account for such a change will underestimate convergence in the pay of young men and women.

In this paper we examine the gender wage gap among young adults across four British cohorts, over five decades. We use four nationally representative British birth cohort studies whose members were born in 1946, 1958, 1970 and 1989–90, and who were interviewed at age 26 (1972), 23 (1981), 26 (1996) and 25/26 (2015) respectively.[1] We ask four questions: first, how has the gender wage gap evolved among young adults in Britain, over a period when policies were being introduced to promote equal pay? Secondly, were graduates more or less likely than non-graduates to have benefited from improved opportunities? Third, how does selection into the labour market affect these trends? Fourth, which human capital, family and job characteristics help explain the size of the gender wage gap?

We study the evolution of the raw gender wage gap at the (logarithmic) mean – for the whole sample and, separately, for graduates and non-graduates – and across the wage distribution. In doing so we compare raw (observed) wage gaps with those adjusted for covariates on personal characteristics. We then present an additional set of estimates of the gender wage gap that accounts for selection into employment, where selection is assumed to be driven by observable characteristics. We use standard decomposition methods to explore the role of early life factors, human capital, marital and parental status in explaining the gender wage gap. Similarly we estimate the gender wage gap across the wage distribution to investigate whether wage inequalities among young adults vary for low and high earners over time and after accounting for selection into employment; we then decompose this gap into a component explained by the distribution of characteristics and an unexplained component. For the two most recent cohorts we have detailed information about the occupations held by respondents: we therefore include observed job characteristics in the covariates to assess how they contribute to the observed gender wage gap at the mean and across the wage distribution.

The paper contributes to the literature on the gender wage gap in four ways. First, it extends earlier studies examining the gender wage gap among young adults using nationally representative samples (see, for example, Neuburger, 2010; Neuburger et al., 2011; Fortin, 2008; Manning and Swaffield, 2008; Combet and Oesch, 2019), by incorporating a cohort of individuals born as early as 1946 and a much more recent birth cohort – born in 1989/90 – to examine trends over 43 years. The 1946 cohort study, although it has limited sample size, includes unique microdata evidence from the pre-equal pay and opportunities legislation era. We find suggestive evidence that the equal pay policies had an impact in reducing the gap between the 1970s and 1980s, and that the NMW policy reduced the gap for the lowest wages in the later cohort.

The second contribution is to assess and compare the trends in the gender wage gap among graduates and non-graduates in the population at large. Other studies on gender inequalities in wages and earnings among young people focus on homogeneous samples of high-skilled young adults (Dolton et al., 1996; Chevalier, 2007; Goldin and Katz, 2008; Bertrand et al., 2010; Azmat and Ferrer, 2017; Bütikofer et al., 2018). In particular Francesconi and Parey (2018) use representative survey data from Germany to study gender gaps in earnings among young graduates across multiple cohorts. We follow a similar approach while studying the gender wage gap in Britain and we complement it with evidence on the evolution of gender disparities in wages among non-graduates. We find that the gender wage gap among non-graduates declines over time while wages stop converging after the mid-1990s among young graduates.

Our third contribution is to study gender inequalities in earnings across cohorts by accounting for differential selection into employment over time, given the changing participation of young men and women in the labour market and their shifts in educational attainment and fertility. We find that adjusting for positive selection into employment increases the size of the gender wage gap in earlier cohorts, but selection is not apparent in the two most recent cohorts. Thus, the rate of convergence in the wages of young men and women is understated when estimates do not adjust for positive selection in earlier cohorts.

Our fourth contribution is to investigate the role of specific job characteristics in explaining the gender inequality in this particular age group as well as human capital and family. As women started to look more like men by catching up with them in terms of their main human capital traits, researchers turned to differences in job characteristics to explain the gender wage gap (Cortes and Pan, 2018). Groshen (1991) reports that occupational segregation explains a sizeable part of the gap in five US industries. Bayard et al. (2003) use administrative data for all sectors and occupations in the US to study the contribution of occupational gender segregation to the gender wage gap. More recently Blau and Kahn (2017) show that differences over time in occupations and sectors explain a rising proportion of the gender wage gap for all women in the US. We find that occupational gender segregation plays an increasingly important role, accounting for much of the observed gap in 2015.

The rest of the paper is structured as follow: Section 2 describes the data; Section 3 details the empirical strategy to estimate the gap; Section 4 presents the results of our decomposition analysis; Section 5 concludes.

2. Data

We examine gender wage gaps among young people in their twenties from four nationally representative British cohort studies (see Medical Research Council, 2016; University College London, 0000a,b,c, various years, for documentation). The first is the National Study of Health and Development (NSHD), a cohort born in 1946 and interviewed as young adults, at 26 in 1972. The second is the National Child Development Study (NCDS) whose members were born in 1958 with an interview in 1981 at age 23. The third is the British Cohort Study (BCS), born in 1970, whose members responded to a postal survey in 1996 at age 26. The fourth is Next Steps,[2] born in 1989–90, in England only, whose members were interviewed in 2015 at age 25/26.

All four cohort studies include a rich set of demographic and longitudinal socio-economic characteristics of the cohort members. NSHD, NCDS and BCS follow the cohort members since birth. Next Steps started following cohort members from the age of 14 which means that early individual and family characteristics are not available for this cohort. All four studies include cognitive measures recorded at pre-employment ages which would not be available in cross-sectional

[1] 1958 cohort members interviewed in 1981 were three years younger than the respondents in other cohorts. Previous studies based on British data have shown that gender differences in wages are smaller when workers enter the labour market and then widen over time with the arrival of family responsibilities (Costa Dias et al., 2020). The implication is that the gender wage gap observed throughout our analysis in 1981 is likely a lower bound for the gap we could have observed among individuals aged 25 years old in the same year. This is further discussed in the interpretation of our results.

[2] Originally known as Longitudinal Survey of Young People in England.

sources.[3] Detailed information about the highest qualifications achieved by the cohort members at the time of data collection allows us to map these qualifications into levels of education across cohorts, by taking into account the significant changes in the structure of education that happened over time. This allows us to study the evolution of the gender wage gap separately among graduates and non-graduates. Graduates are defined consistently across the four cohort studies as those respondents who achieved post-secondary professional qualifications and higher education diplomas, university degrees or higher degrees.[4] Gross hourly wages are derived from information collected in every cohort about gross weekly wages and hours worked per week; these are then deflated by the RPI to January 2000 prices. Survey questionnaires' items used to derive gross hourly wages in all cohorts are detailed in Appendix A (available online).

The studies were affected by sample attrition over time, an issue we tackle with attrition weights as described in Section 3. The initial samples used in our study include only respondents who participated in the survey in the year of interest (1972 for NSHD, 1981 for NCDS, 1996 for BCS and 2015 for Next Steps). The full number of observations for men and women in each cohort (with or without wages) is the following: 1901 and 1851 in the NSHD, 6043 and 6033 in the NCDS, 3959 and 4756 in the BCS and 3371 and 4198 in Next Steps.

The evolution of female and male employment rates is crucial to understanding any change in selection affecting the comparison of cohorts. The proportion of young men and women in employment across the four cohorts is reported in the left plot of Fig. 1. For young men this dropped sharply between 1972 and 1981 (from 94% to 76%) and remained constant after that. The employment rate of young women increased between 1972 and 1981 by 14 percentage points (from 48% to 62%), coinciding with the Equal Pay and opportunities legislation that came into force in 1976, and a sharp drop in childbearing by women in their early twenties. It continued to increase between 1981 and 1996 by 9 percentage points. The increase was only marginal in the last cohort (from 73% to 77%). Similar trends in employment are found among the full sample of working age respondents, as well as young men and women, in the Labour Force Survey that is a nationally representative survey for the UK population (Figure A1, in Appendix A). The sharp relative change in female and male employment between the late 1970s and the 1990s was caused by the interaction between a large labour demand shock, the shift of labour demand away from the manufacturing sector towards the service sector, the welfare system and the role of unions (Blundell et al., 2007; Harkness, 1996; Machin, 1994). The steep convergence in employment rates between 1972 and 1996 suggests the types of men and women entering the labour market possibly changed too, reducing the positive selection of women and increasing that of men.

The comparison of gender inequalities in wages by education levels over time requires the consideration of possible differential selection into higher education. If young women become less positively selected into higher education over time compared to men, the convergence of wages will be underestimated for both graduates and non-graduates. Young women were less likely than young men to have a degree in 1972, as shown by the right plot in Fig. 1. After that, the proportion of young men and women with higher education was similar and it increased over time until 2015, when the young women were more likely than men to be graduates.[5] The low proportion of graduate women compared to men in the earliest cohort suggests more positive selection of women compared to men into higher education in that cohort.[6] This is confirmed by descriptive statistics of cognitive test scores recorded pre-labour market entry presented in Tables B1 to B4 in Appendix B. A comparison between graduates and non-graduates shows that across all cohorts there is positive selection into higher education, i.e. those with a degree have higher pre-market cognitive scores in school. However, the advantage has declined over time particularly for women. Graduate women had higher scores than men in reading and maths cognitive tests in 1972. This difference changed in favour of men in 1981, and then it appeared to settle in later cohorts where women performed better in reading and men performed better in maths. We do not observe gender differences in test scores among non-graduates. If young women have been less positively selected into higher education over time, compared to men, this might have contributed to a growing divergence in the earnings of graduate men and women over time, and increased the rate of convergence in earnings between non-graduate men and women.

The variables used in the covariate adjustment of the gender wage gap can be grouped into four main sets.

- Early years factors: mother's school leaving age, father's school living age and father's social class.
- Human capital variables: maths and reading standardized test scores, indicators for the highest qualification achieved, indicators for subjects at higher education (science, technology, engineering, or mathematics, law, economics or management, others), months of full-time and part-time work experience, squared months of full-time and part-time work experience, the number of different spells of work experience.[7]
- Family formation variables: indicators for whether the young adult is married/cohabiting, and the presence of dependent children.
- Job characteristics: sets of indicators for weekly hours worked (30 h or less, 30 to 45 h and more than 45 h); and for the major group of the Standard Occupational Classification for the cohort member's job; and, for BCS and Next Steps only, a measure of

[3] Next Steps is linked to administrative records of educational attainment. Therefore it includes scores in Reading and Maths from the tests taken by the cohort members at the end Key Stage 2, aged 11–12.

[4] Graduates are defined as individuals who achieved level 4 or 5 of National Vocational Qualifications (NVQ), which include: BTEC Higher Certificate/Diploma, HNC/HND, Professional degree level qualifications, Nursing qualifications, RSA Higher diploma, Non-Vocational Qualification level 4 and 5, Degree, HE diploma, PGCE, other teaching qualifications, higher degree.

[5] The apparent dip in the proportion of male graduates between 1972 and 1981 should be qualified by the consideration that the NCDS (1981) data comes from the lowest age, 23. This drop can be observed among similarly aged respondents in the Labour Force Survey (Figure A2 in Appendix A) and it is possibly driven by the age difference between the cohort members in those two years (respondents were 26 year old in 1972 and 23 years old in 1981). The series disguises an upward trend across all of these cohorts in the proportions having graduate qualifications by their thirties, see Ferri et al. (2003), Table 2.1a. By that age then, NCDS had overtaken the cohort interviewed in 1972 in the proportion graduate qualifications gained post 23.

[6] Figure A2 in Appendix A plots employment and higher education status by gender for the Labour Force Survey respondents of similar age as the respondents in the cohort studies. This comparison shows that the cohort studies are generally representative. Discrepancies may be due to differences in the questionnaires and classifications (in the Labour Force Survey individuals are classified as either employed or in education while in the cohort studies they can be both at the same time); and, in the case of higher education status, to the different ways qualifications are defined and grouped in the data. Wages are available in the Labour Force Survey only after 1996. Table A1 in Appendix A compares Labour Force Survey (log) real hourly wages reported by 25 to 27 years old respondents in 1996 and 2015 to wages in BCS and Next Steps showing that the cohort studies are indeed representative. Furthermore, Table 2 in Neuburger et al. (2011) compares log median hourly wages on men and women in the 1972 and 1981 sweeps of the cohort studies with respondents of similar ages from the Family Expenditure Survey also confirming the representativeness of the earlier two cohort studies.

[7] A detailed description of the cognitive measures available in the cohort studies and used in the analysis are available here: https://closer.ac.uk/cross-study-data-guides/cognitive-measures-guide/.

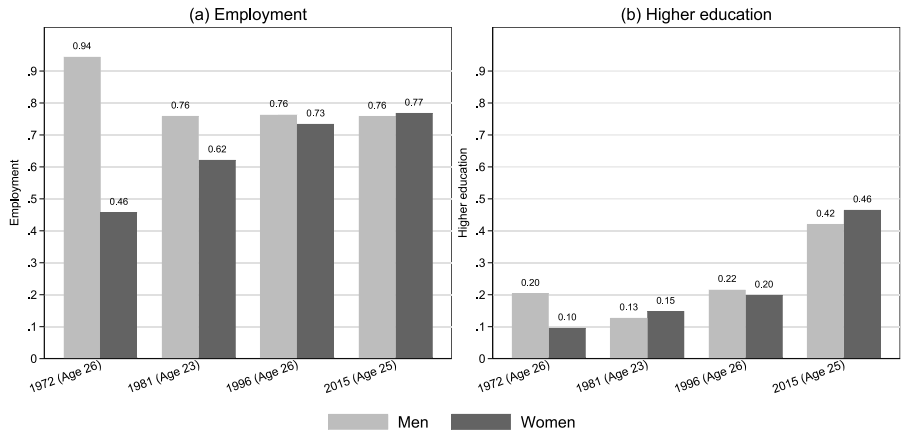


Fig. 1. Employment and higher education rates among young men and women in the British cohort studies.
Note: Authors' calculation based on NSHD data in 1972, NCDS data in 1981, BCS data in 1996 and Next Steps data for England in 2015. Weights are used to account for attrition in all four cohorts and for stratification in NSHD.

occupational segregation, defined specifically as the proportion of women by occupation at four digit level of the Standard Occupational Classification. This last variable is constructed by using the contemporary Labour Force Surveys.

Descriptive statistics for these variables in the estimation samples are reported by higher education status and cohort in Tables B1 to B4 in Appendix B.

3. Empirical methodology

We compute and compare three variants of the gender wage gap: the raw gender wage gap; a covariate adjusted gap; and a covariate adjusted gap that also accounts for selection into employment. We define the raw gender wage gap as the difference in mean log hourly wages between men and women. The adjusted gender wage gap is the estimated coefficient γ of the female indicator variable from the following model:

$$Ln\left(HW_i\right) = \gamma Female_i + X_i\beta + \varepsilon_i \quad (1)$$

Where HW_i is the real hourly wage derived from reported weekly wage and hours worked; X_i include all the variables for early years factors, human capital and family formation presented in Tables B1 to B4 in Appendix B.

To account for the effect of differential selection into employment on the estimates of the gender wage gap we adopt an imputation-based approach similar to Bryson et al. (2020) and Neuburger et al. (2011).[8] Wages are missing for individuals who are in employment without a wage observation, unemployed or economically inactive, self-employed or full-time students. These are imputed with 'potential wages' obtained from individuals with an observed wage who are closer in terms of their propensity to be in waged employment. First we estimate, by gender and cohort, the propensity for waged employment using probits for the (0,1) outcome of having an observed wage at the time of the interview. Tables B5 to B7 in Appendix B reports the list of variables used in the estimation of the probit models.[9] We use propensity score matching to find for each individual without an observed wage, the nearest neighbour in terms of propensity for waged employment within the same sex and cohort. The common support requirement is implemented by dropping respondents whose propensity for waged employment is smaller than the lowest probability in the waged employee sample in each cohort.[10] The wage of the nearest neighbour is then used as 'potential' wage to impute the missing wage.

The selection-adjusted gender wage gap is the estimated coefficient of the female indicator from model (1) where employees' log hourly wages are supplemented with the imputed log hourly wages for individuals for whom we do not observe earnings in the sweep. This imputation-based approach to adjust the estimates of the gender wage gap for non-random selection into employment can only account for selection on observables. It is however possible that the decision to work is also determined by characteristics of the cohort members that are usually unobserved, such as preferences and attitudes. For this reason we explore a robustness check where we include measures of job aspirations and attitudes in the estimation of the propensity to waged employment for two of the four cohorts.

We adjust for sample attrition over time by weighting each individual observation by the inverse probability of responding to each sweep. The inverse probability is obtained as the inverse of the predicted probability estimated with a logit for the probability of responding in the sweep of interest given a valid response in the sweep at age 2 (the first follow up) for NSHD, the sweep at age 10/11 for BCS and NCDS; and in the sweep at age 14 for Next Steps.[11] Tables B8 to B11 in Appendix B report the list of variables used to estimate

[8] Other approaches that deal with non-random selection include the Heckman two-step selection model (Mulligan and Rubinstein, 2008) and estimating bounds for the gender wage gap based on quantiles of the wage distribution (Blundell et al., 2007). However, both these methodologies require a credible exclusion restriction which was not available to us.

[9] Our assumption is that selection into employment over time is based on observed characteristics. Heckman et al. (2006) show that pre-market non-cognitive traits, mostly unobserved, are important drivers for labour market outcomes. Unfortunately, while British cohort studies include some non-cognitive measures, these are not consistent across cohorts and therefore they could not be used for a cross-cohorts comparison.

[10] After implementing the common support restriction the number of observations for men and women in each sample adjusted for selection is the following: 1879 and 1831 in the NSHD, 5911 and 5810 in the NCDS, 3721 and 4341 in the BCS and 3322 and 4085 in Next Steps.

[11] The NSHD consists of a socially stratified target sample of individuals born in one week in March 1946. This longitudinal survey only followed up one in four urban working-class children (whose fathers were in urban, manual occupations at birth) and therefore stratification weights are used for descriptive statistics, sample sizes and for the analysis based on these data.

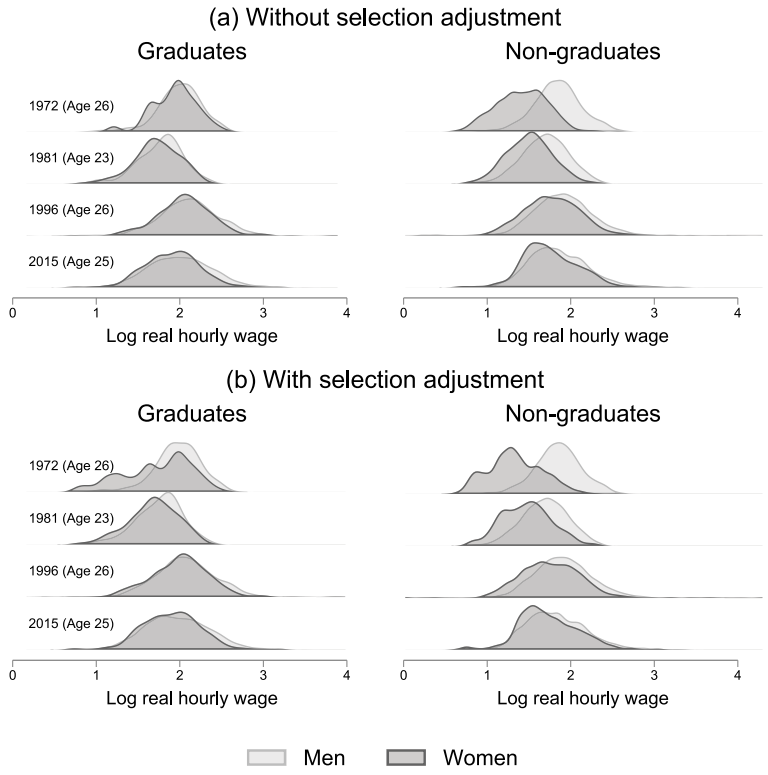


Fig. 2. Distribution of log hourly wages for graduates and non-graduates by cohort.
Note: These plots present the kernel density of observed log hourly wages (top panel) and observed as well as potential log hourly wages (bottom panel) by gender, higher education status and cohort. Potential log hourly wages are obtained from nearest neighbour 'donors', defined as those in waged employment in the same cohort and from the same gender who are closer in their propensity to waged employment to the respondents for whom a wage is not observed.

the probability of responding at the sweep of interest. We deal with item non-response in individual characteristics by creating a binary indicator for the missing observations if the variable of interest is categorical, while if the variable is continuous we impute the missing value with the gender-specific mean of the variable obtained by using non-missing only observations and we include an indicator for the missing observations.

To understand the contribution of each set of variables (early life factors, human capital and family formation) to the gender wage gap we carry out a Kitagawa–Oaxaca–Blinder decomposition (KOB hereafter), Kitagawa (1955), Oaxaca (1973), Blinder (1973), of the covariate adjusted and the selection and covariate adjusted gap. To this end log hourly wages for each gender (g) are estimated using the following equation:

$$Ln\left(HW_{ig}\right) = X_{ig}\beta_g + \varepsilon_{ig} \tag{2}$$

The models for females and males can be subtracted from each other to decompose the mean gender wage gap into mean differences in observed characteristics and differences in returns to these characteristics. To define the differences in returns, one needs to choose reference parameters. Studies suggests results can be sensitive to different choices. We follow the approach suggested by Neumark (2004) and Jann (2008) by selecting as reference parameters the coefficients from regressions pooled over males and females with the addition of a gender indicator. The decomposition is thus:

$$\overline{Ln\left(HW_{iM}\right)} - \overline{Ln\left(HW_{iW}\right)} = \left(\overline{X}_M - \overline{X}_W\right)\beta^* + \overline{X}_M\left(\beta_M - \beta^*\right) + \left(\beta^* - \beta_W\right)\overline{X}_W \tag{3}$$

Where β^* is the reference coefficient estimated in Eq. (1).[12] The first term on the right-hand side of Eq. (3) represents the explained component of the gender wage gap. The second and third terms together represent the unexplained component: the deviations of male parameters from the reference set, weighted by mean male characteristics plus the deviations of female parameters from the reference set, weighted by mean female characteristics. This unexplained component reflects any unequal treatment/discrimination in the labour market, the impact of gendered social norms and preferences, and, of course, of any other factors differentiating the sexes which may not be captured by the characteristics included in the model. In an additional set of analyses we explore the role of job characteristics to explain the gender wage gap among young adults. These decompositions of the gender gap are based on observed wages only because we cannot impute wages where we do not have information on occupations.

Finally, we study the relative importance of differences in observable characteristics and their returns in explaining wage inequalities across the wage distribution, without and with adjustment for selection. To study the unconditional gender gap across the wage distribution and its decomposition between the effect of characteristics and coefficients we rely on the methodology proposed by Chernozhukov et al. (2013). This method decomposes unconditional gender gaps at a given percentile of the distribution of log wages into a component due to the distribution of characteristics (the explained part of the gap) and a component due to different wage functions conditional on characteristics (the unexplained part of the gap). The empirical implementation

[12] We deal with the sensitivity to choice of the reference categories for categorical variables by using the option "normalize" from the Stata command oaxaca (Jann, 2008).

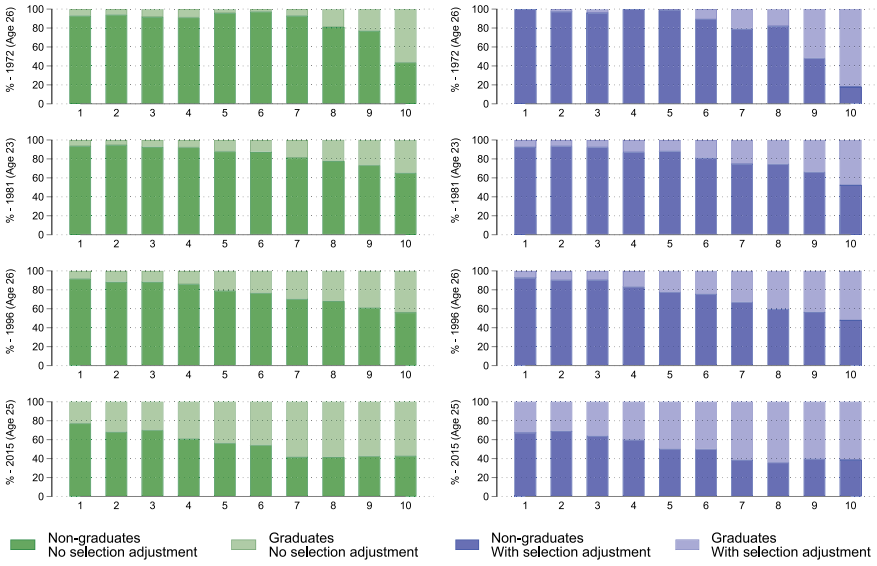


Fig. 3. Percentage of graduates by decile of wages — Women.
Note: this figure plots the percentage of graduates and non-graduates by deciles of young women's log real hourly wages for each cohort study. The green bars report the percentages of graduates and non-graduates by deciles of wage for wage earners. The blue bars report the percentages of graduates and non-graduates by deciles of wages adjusted for selection. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

of this method first requires the estimation of conditional wage distributions by using quantile regressions. Then, unconditional wage distributions are estimated through integration over the distribution of the covariates.

4. Results

4.1. Selection adjustment and wages

After imputing potential wages for individuals without observed earnings, thus adjusting the sample of working individuals for non-random selection into the labour market, we compare the samples with and without selection adjustment. The aim is to show how the main characteristics of young women and men in and out of the labour market have changed over time, bearing in mind that the exact age at observation varies by cohort.

Tables B1-B4 in Appendix B report descriptive statistics, by gender and higher education status, for the early life factors, human capital and family formation variables that will later be included in the estimation of the covariate adjusted gap. The first four columns present descriptive statistics for the sample that includes observed and potential wages while the last four columns report the same for the sample that includes only observed wages. By comparing some of the main individual characteristics, such as cognitive test scores, full-time experience and parental status, between these two samples, by gender and higher education status, we can approximate whether selection into the labour market evolved over time differently between men and women.

Relative to men, graduate women in the first cohort and non-graduate women in the first two cohorts, were positively selected into work in terms of cognitive test scores and months of work experience. In all cohorts working women are much less likely to have children than women in the broader sample while in none of the cohorts did the proportion of men with children differ by whether they are selected into work or not.

By considering only respondents who have an observed wage across cohorts in 1996 and 2015 it is possible to observe how differences in job characteristics for men and women have changed over time. Young women are more likely than men to work in part-time jobs and less likely to work long hours. However, between 1996 and 2015, women narrow the gap in hours, particularly in long hours worked. We do not observe noticeable differences in the proportion of men and women in managerial or professional jobs while the proportion of all young people in these occupations has decreased across the two cohorts.[13] On average both graduate and non-graduate women hold occupations that are more female-intensive than their male counterparts. But the gap between men and women in terms of female concentration in the occupation has narrowed between 1996 and 2015, as more women joined more integrated and male dominated jobs.

Fig. 2 plots, for each cohort study and by higher education status, the distribution of men's and women's log hourly wages before and after imputing potential earnings. This shows that observed wages of women in the labour market in 1972 and, marginally, in 1981 were higher than the potential wages of women who were not in employment, confirming the hypothesis that young women in NSHD (1972) and NCDS (1981) positively selected into employment. Comparison of the distribution of wages before and after imputation for BCS (1996) and Next Steps (2015) shows that the stronger positive selection of young women reduced over time.

Graduates are expected to have higher wages than non-graduates in the population as a whole. This is not so clear-cut amongst young adults in their mid-twenties where, at the same time, we observe graduates at the beginning of their career and non-graduates who may already have several years of experience in the labour market. Furthermore the entry wage for graduates may have changed across cohorts and may differ between men and women. Figs. 3 and 4 plot the proportion of female and male graduates and non-graduates by deciles of log real hourly wages without and with selection adjustment (respectively, green and blue bars). These figures show that graduates are mostly higher up the distribution of wages and this is particularly true for women and in earlier cohorts. As the proportion of graduates increases in the samples (the greatest increase is between 1996 and 2015) we observe more graduates in the lowest deciles of the wage distribution, although the majority of individuals in the lowest deciles are still non graduates. The

[13] Managerial and professional occupations are defined as Major Groups 1 (managers, directors and senior officials) and 2 (professional occupations) in the Standard Occupational Classification 2000 and 2010.

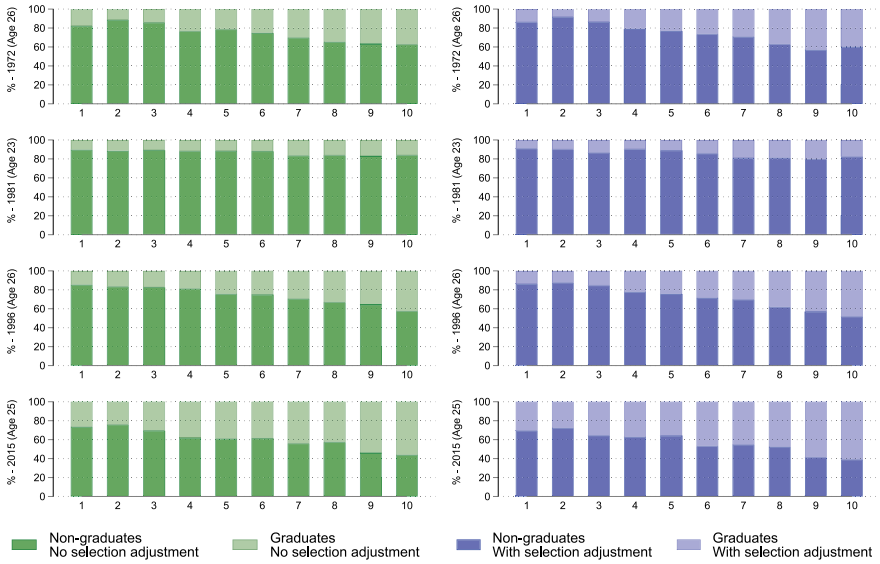


Fig. 4. Percentage of graduates by decile of wages — Men.
Note: this figure plots the percentage of graduates and non-graduates by deciles of young men's log real hourly wages for each cohort study. The green bars report the percentages of graduates and non-graduates by deciles of wage for wage earners. The blue bars report the percentages of graduates and non-graduates by deciles of wages adjusted for selection. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

proportion of graduates across the wage distribution does not change with the imputation of unobserved wages, as we would expect.

4.2. Trends in gender wage gap

A comparison of the raw gender wage gap and the covariate-adjusted gap for waged employees in the four cohorts is presented in Fig. 5. The gender wage gaps reported here are the estimates of the coefficients for the female indicator in the Mincer wage Eq. (1). Results are reported for the whole sample (black) and for graduates (dark grey) and non-graduates (light grey) respectively. For each sample the figure reports the estimates of the raw gap (dot marker) and of the conditional gap (triangular marker) from a model that includes early life characteristics, qualifications, experience, marital and parental status, and the relative 95% confidence intervals. The estimates are negative as they represent the female mean log hourly wage minus the male mean log hourly wage. The full sets of estimated coefficients for these regressions are reported in Tables C1 to C4 in Appendix C.

Measured in points below the zero mark, the raw gap for the whole sample shrinks from 41 log points in 1972 to 10 log points in 1996 and 8 log points in 2015. Adjusting for early life factors, human capital and family formation characteristics only reduces the overall gap significantly in the earliest cohort.

The raw gender wage gap among graduates, represented by the dark grey dots, increases between 1972 and 2015 (from 2 log points to 9.1 log points). The covariate-adjusted gender wage gap among graduates is small or statistically non-significant in the first three birth cohorts, but is sizeable (6.2 log points), and statistically significant, in 2015.

The two light grey estimates in Fig. 5 present the same analysis, for the non-graduate subsample. Among non-graduates the raw gender wage gap dropped between 1972 and 2015: from 46.5 to 8.5 log points. As with the overall sample, adjusting for observed characteristics reduces the gap only in the earliest cohort but does not in any of the subsequent cohorts.

Fig. 6 compares the covariate adjusted gender wage gap (triangular marker) to the selection and covariate adjusted gender wage gap (square marker). The full sets of estimated coefficients for these regressions are also reported in Tables C1 to C4 in Appendix C.

Adjusting for employment selection as well as individual characteristics increases the gender wage gap for the whole sample in a statistically significant way in 1972 (by 8 log points, from 36 to 44) and 1981 (by 1.5 log points from 16.6 to 18.1). The increase in the gap estimate in 1996 of 1.1 log points (from 8.9 to 10) is not statistically significant, while the gender wage gap in 2015 decreases (from 7 to 5.7 log points), but this change is not statistically significant.

Accounting for non-random selection into employment increases the gender wage gap among young graduates in 1972 by 20 log points. Furthermore, it increases the size of the estimated graduate gender wage gap in 1981 by 2 log points (a statistically significant rise from 3.3 to 5.3 log points) while the change in the estimate of the gap in 1996 and 2015 is not statistically significant.

Similarly adjusting for selection increases the gender wage gap for non-graduates in 1972 (from 42 log points to 48) and marginally but significantly (from 19 to 20 log points) in 1981, while it does not vary the estimates of the gap among non-graduates in 1996 and 2015.

In summary, whilst the overall gender wage gap decreases over time this is driven by the decrease in the gap among non-graduates. Gender wage inequality among graduates worsened over time. Adjusting for observed characteristics explains some of the gap for the whole sample in 1972, and only among graduates in 1996 and 2015. Taking into account selection into employment widens the gap among young adults in the earliest cohort, mostly driven by positive selection among graduates, but it has little or no bearing on estimates of the gaps for the most recent cohorts.

There is a striking reduction between the wage gaps in 1972 and 1981, adjusted for covariates (apart from graduates before allowing for selection). Although these were also years of a rapid decline in early childbearing (De Cooman et al., 1987), and age at observation had changed from 26 to 23, the closing of the wage gaps can be assumed to reflect, at least in part, the impact of intervening equal opportunities legislation on the otherwise unexplained female pay penalty. This result is in line with previous studies showing that over the 1970s the

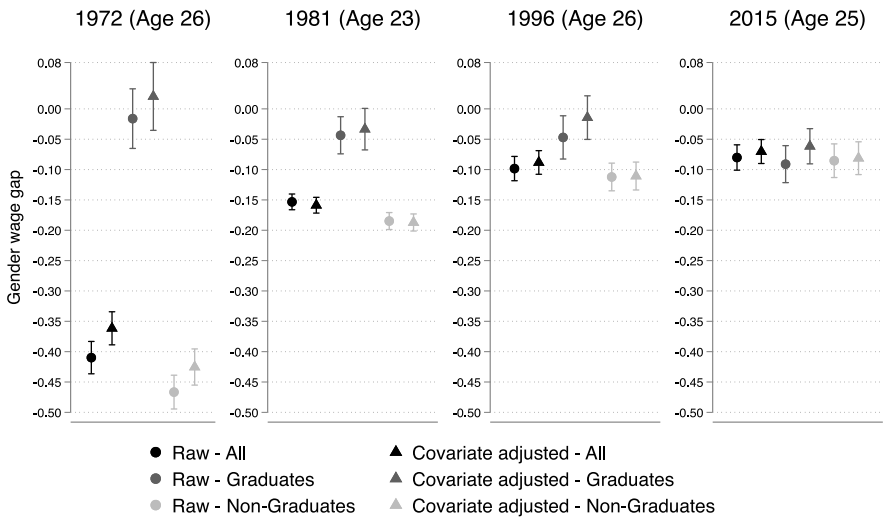


Fig. 5. Raw and covariate-adjusted gender wage gap (with 95% CI).
Note: These plots report, by cohort, the difference between women's and men's log hourly pay as estimated coefficients for the female indicator from Eq. (1) and its 95% confidence interval for different samples and models. Estimates marked in black refer to the whole sample; estimates in dark grey refer to graduates; estimates in light grey refer to non-graduates. For each sample the dot marker reports the estimate of the coefficient for the female indicator in a wage model without additional controls, while the triangular marker reports the estimate of the coefficient for the female indicator in a wage model that includes early life characteristics, qualifications, experience, marital and parental status. Tables B1 to B4 in Appendix B report the full list of variables included in the model for each cohort. Full set of estimates are reported in Appendix C, Tables C1 to C4.

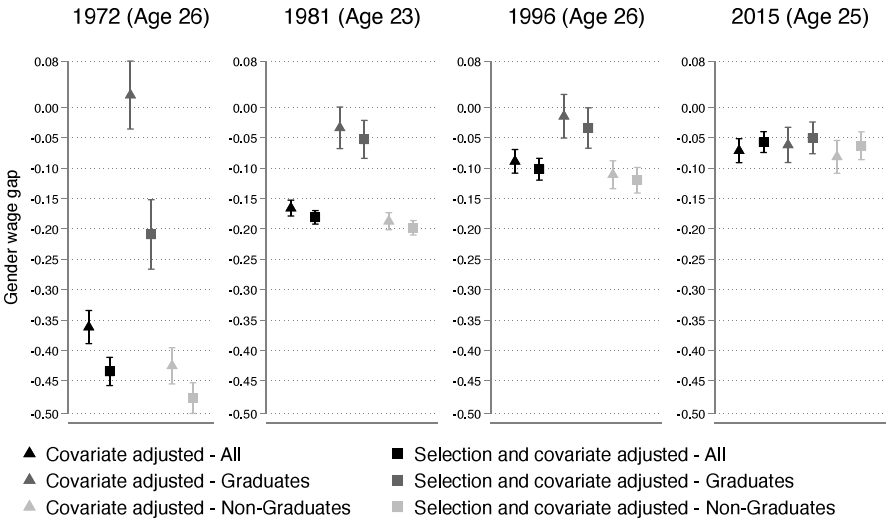


Fig. 6. Covariate-adjusted and covariate- and selection-adjusted gender wage gap (with 95% CI).
Note: These plots report, by cohort, the difference between women's and men's log hourly pay as estimated coefficients of the female indicator from Eq. (1) and its 95% confidence interval for different samples and models. Estimates marked in black refer to the whole sample; estimates in dark grey refer to graduates; estimates in light grey refer to non-graduates. For each sample the triangular markers report estimates based on observed wages while the square markers report estimates based on observed and potential wages. All models include the full set of controls for personal characteristics (early life characteristics, qualifications, experience, marital and parental status). Tables B1 to B4 in Appendix B report the full list of variables included in the model for each cohort. Full set of estimates are reported in Appendix C, Tables C1 to C4.

relative employment and wages of women increased as a consequence of these policy changes (Zabalza and Tzannatos, 1985; Manning, 1996; Harkness, 1996). Manning (1996) argues that the simultaneous increase in both women's employment and pay would have been a reflection of a monopsonistic labour market, which prior to the Equal Pay Act would have deterred participation in labour market. Zabalza and Tzannatos (1985) also point to the simultaneous enactment of anti-discrimination and equal pay legislation as a reason for the rise in women's relative wages after 1975 being sustained without lowering their employment.

The positive selection of graduate and non-graduate 26 year old women into the labour market, in the years before the full implementation of the Equal Pay Act, is therefore likely to have reflected a monopsonistic labour market where women were paid less than their marginal productivity and only women with the highest returns would enter the labour market. Admittedly, a caveat to these results is that the younger age of respondents in 1981 may also have contributed to the sharp gender convergence in hourly earnings that we observe in our data between the first two cohort studies. Observations in 1981 among 23 year olds

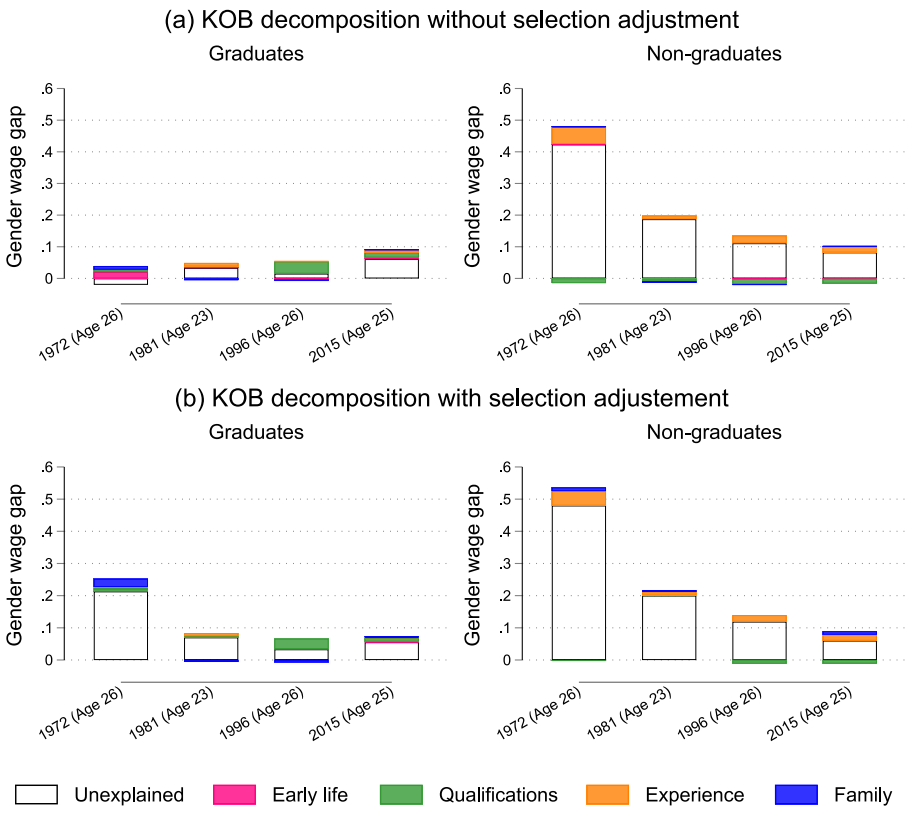


Fig. 7. Decomposition of the mean gender wage gap for graduates and non-graduates by cohort: early life factors, human capital, marital and parental status.
Note: The figure reports the Kitagawa–Oaxaca–Blinder decomposition of the log wage gap among graduates and non-graduates without selection adjustment (top panel) and with selection adjustment (bottom panel). The gap is reported as men's minus women's' log hourly wage. The full decomposition with confidence intervals for each cohort is reported in Tables C5 and C6 in Appendix C. Results from gender specific regressions are available upon request. The variables used in the decomposition are those reported in Tables B1 to B4 in Appendix B. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

could underestimate the gap among individuals aged 25 in the same year.

4.3. Decomposition of the mean gender wage gap (Kitagawa–Oaxaca–Blinder decomposition)

Fig. 7 presents the KOB decomposition of the gender wage gap for graduate (on the left) and non-graduate (on the right) cohort members. This figure reports how the main sets of characteristics account for the gender wage gap. The full decomposition estimates with standard errors for each cohort and by higher education status are reported in Appendix C, Tables C5 and C6. The top two plots in Fig. 7 report the decompositions of the gender wage gap without selection adjustments in the sample as the gap is reported as men's minus women's'. The bottom two plots present the decompositions of the gender wage gap that accounts for non-random selection into employment. Negative numbers represent differences in characteristics and returns that work in favour of women. Results of the decompositions without selection adjustment show that in 1972 the small gap among graduates was entirely explained by differences in early life characteristics, qualifications (that include cognitive test scores and subjects studied), and parental status. Differences in experience explain a small portion of the gap in 1981 while differences in qualifications explain some of the gap among graduates in 1996 and 2015.

Among non-graduates differences in experience explain a small proportion of the gender wage gap across the four cohorts while qualifications, that for this group include indicators for secondary school levels as well as cognitive test scores, represent consistently a small negative component of the gap, reflecting the higher attainment of women in these sub-samples. Accounting for non-random selection increases the unexplained component of the gender wage gap in the earlier two cohorts for both graduates and non-graduates. The unexplained component of the gender wage gap decreases over time among non-graduates. Among graduates instead the unexplained component is greater in the most recent cohort compared to the previous two.[14]

The major component of each wage gap is not accounted for by differences in human capital, and must therefore reflect other sources of differential rewards for women. Apart from omitted covariates, there could be unequal remuneration due, for example, to discrimination, unequal bargaining power, different hours, different risks, sexual harassment or sex stereotyping of tasks, which may be inter-related and reinforcing (Azmat and Petrongolo, 2014; Blau and Kahn, 2017). In Section 4.5 we will consider job characteristics, and in particular job segregation, as additional covariates. These can be interpreted as adding to the explanation of the gender gap or as the channel through which the unequal treatment of human capital characteristics operates. Which of these dominates depends on the net addition the job characteristics make to the explanation of the pay gap and how far it absorbs differentials associated with human capital.

[14] When the samples are pooled across those with and without higher education (Tables C7 and C8 in Appendix C) the contribution of qualifications to explaining the gender gap is apparent in the 1946 cohort. The small contribution of human capital to explaining the gender wage gap reaches 7.5 log points in this cohort (in the selection adjusted model) with the higher qualifications of men accounting for 2.3 point of the gap. In the three later cohorts the explained gap remains negligible, as does the estimate for qualifications, reflecting the convergence of educational attainments. The contribution of qualifications to the overall pay gap is sometimes negative, depending on the cohort and model estimated, but it was not always so.

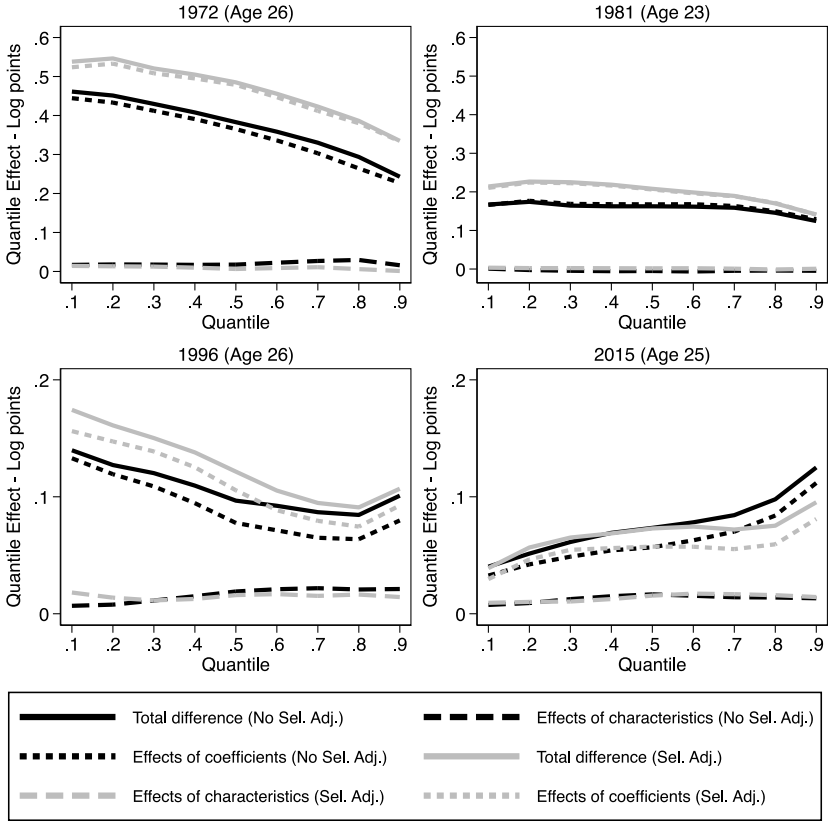


Fig. 8. Decomposition across the wage distribution by cohort: early life factors, human capital, marital and parental status.
Note: The four plots report the decomposition between the 1st and 9th decile of the unconditional difference between the male and the female log wage distribution for the sample of observed wages (black lines) and for that of imputed and observed wages (grey lines) for each cohort study. The variables used in the decomposition are those included in the KOB decomposition reported in Fig. 7 and listed in Appendix B. The solid lines show the total difference between the unconditional distributions of log hourly wages of men and women at every decile. The *y*-axis shows the wage differential in favour of men expressed in log points. The long-dashed line represents the portion of the total gap explained by observed characteristics of the individuals in the sample. The short-dotted line shows the residual portion of the gap that is not explained by differences in observed characteristics. Confidence intervals are reported in Figure C1, Appendix C. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

4.4. Decomposition across the wage distribution

Previous research has highlighted barriers in the labour market that affect women at different points of the wage distribution: women may be prevented from escaping low-paid jobs or progressing to the highest paid positions. These barriers are conventionally referred to as the sticky floor and glass ceiling and they may change differentially over time as shown by Blau and Kahn (2017) for the US. Our study of the gender wage gap at the mean among young adults has shown that the convergence of wages for non-graduates has been consistently faster than that for graduates. In order to understand whether the gender wage gap evolved differently among individuals with lower and higher wages we now investigate the gender wage gap across the wage distribution.

Fig. 8 reports the decomposition across the wage distribution for the sample of observed wages (black lines) and for that of imputed as well as observed wages (light grey lines) using the method proposed by Chernozhukov et al. (2013). The variables used in the decomposition are those in the KOB decomposition reported in Fig. 7. The plot presents the results of the decomposition between the 1st and 9th decile of the unconditional distribution of wages for each cohort study. Confidence intervals for the total gap and the effect of characteristics are reported in Figure C1, Appendix C.

The solid lines show the total difference between the unconditional distributions of log hourly wages of men and women at every decile. The *y*-axis shows the wage differential in favour of men expressed in log points. The long-dashed line represents the portion of the total gap explained by observed characteristics of the individuals in the sample. The short-dotted line shows the residual portion of the gap that is not explained by differences in observed characteristics: it captures the additional log hourly wages women would receive if their attributes were rewarded in the same way as men with those same attributes.

As observed in the decomposition of the mean gender wage gap, male/female differences in observed characteristics explain very little at every decile for the three later cohorts. Amongst cohort members in 1972 the raw gender wage gap, reported in the first plot on the top left, declines as we move up the wage distribution (as indicated by the downward sloping black line), that is it is highest for the low paid at the 10th percentile, as confirmed by the confidence intervals reported in Figure C1. The decline is mostly accounted for by gender differences in the returns while differences in characteristics explain a small component of the gap for low and high wages. Allowing for non-random selection increases the gap for every decile.

The plot at the top right shows the raw gender wage gap declining only marginally after the 8th decile in 1981, however this decline is not statistically significant. The gender wage gap is wholly accounted for by differences in returns to characteristics. Selection adjustment increases the gap among lower wages however does not increase the explained component.

The decomposition of the gender wage gap for 26 year olds in 1996 is shown in the bottom left plot. The total gender wage gap is not significantly different from the gap for 23 year olds in 1981 for the lowest and highest deciles. The gender wage gap declines as we move up the wage distribution. Differences in human capital characteristics contribute only to a very small portion of the gap. Once again, nearly all of the gap is unexplained by the differences in the attributes of men

and women. Instead, it is the gendered returns to those attributes that account for the gap. Accounting for selection does little to the estimates of the gap across the distribution of wages, however it determines narrower confidence intervals: as a consequence the selection-adjusted gender wage gap among lower wages is significantly greater than the gap among higher wages.

Results for the cohort members in 2015 appear at bottom right. The total gender wage gap is significantly reduced among the lowest wages compared to the BCS. In contrast to NSHD, NCDS and BCS, the gap among the highest wages is significantly greater than the gap among the lowest wages, in line with our findings about graduates, who are likely to be among the highest earners. The closing of the gender gap at low wages is plausibly attributable to the introduction of the National Minimum Wage in 1999, from which low paid women were the main beneficiaries (Stewart and Swaffield, 2002; Connolly and Gregory, 2002). Again, most of the gap is unaccounted for by differences in men's and women's characteristics. Accounting for non-random selection increases the level of uncertainty of the total gap among the highest wages leading to a non-statistically significant difference in the gap between the bottom and the top of the wage distribution.

In summary, the analysis of the gender inequalities across the wage distribution has some important findings. First, as already observed in the decomposition at the mean, differences in observable characteristics account for little if any of the gender wage gap across the wage distribution in all cohorts. This result is in line with the outcome of the decomposition at the mean. Second, accounting for selection into the labour market increases the total gender gap across the wage distribution in the first two cohorts. Third, gender inequalities in wages for lower paid jobs decrease over time. Finally, the gap at the top of the wage distribution decreases dramatically only between 1972 and 1981 while in the most recent cohort it is higher than for low paid jobs.

4.5. Decomposition of the gender wage gap with job characteristics

So far our results show that little of the wage gap is explained by differences in human capital and other personal characteristics/family formation. However, as discussed earlier, a growing literature points to the importance of job characteristics in explaining the gap. To study the role of job characteristics, the analysis of the gender wage gap has to focus on individuals who are in the labour market and hold a job. As a consequence, our imputation-based approach to adjust the estimates of the gender wage gap for non-random selection cannot be used.

Detailed occupations are available and comparable only for the two most recent cohorts interviewed in 1996 and 2015. In addition, results reported in Sections 4.2 and 4.4 show that estimates of the gender wage gap among young adults in these two cohorts are not affected by the selection adjustment introduced with our empirical strategy. For this reason, results of the decompositions that include occupational characteristics for respondents interviewed in 1996 and 2015 only are reported in this section in Figs. 9 and 10. To reiterate, these decompositions do not account for selection into employment as we do not have information about occupations for individuals who are not in the labour market.

Estimates of the wage equation that includes occupational characteristics are reported in Table C9 in Appendix C while results for the KOB decomposition with standard errors are reported in Table C10.

The decomposition shows gender differences in hours worked reduce the gender wage gap in 1996 cohort among graduates, and in both cohorts among non-graduates. Table C9 in Appendix C shows that this result, ceteris paribus, stems from the negative estimated returns to hours worked, conditional on type of occupation and occupational gender segregation.[15] Returns to longer hours become less negative among graduates over time while they become more negative among non-graduates. At the same time the proportion of young women working long hours increases, particularly among graduates. Negative returns to hours worked are not surprising: Denning et al. (2022) find that wage returns to mean hours worked across occupations are positive while wage returns are very small or negative at individual level within occupations. Among non-graduates we observe that type of occupation and qualifications represent further negative components of the gap across both cohorts. This implies that women are in broad types of occupations and have qualifications with higher returns compared to men, once the sex composition of the occupation is taken into account.

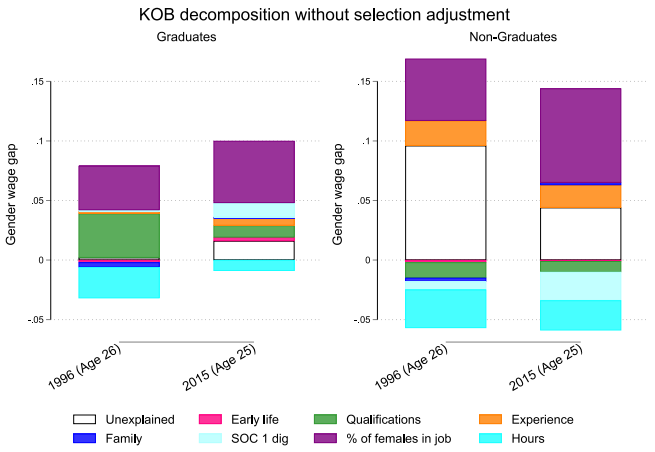


Fig. 9. Decomposition of the mean gender wage gap for graduates and non-graduates by cohort: early life factors, human capital, family formation and job characteristics. Note: The stacked bar charts report, by cohort, the KOB decomposition of the log wage gap among graduates (left panel) and non-graduates (right panel) based on observed wages. The gap is reported as men's minus women's' log hourly wage. Results from the gender specific regressions are available upon request. Estimates of the coefficients for job characteristics are reported in Table C9 in Appendix C.

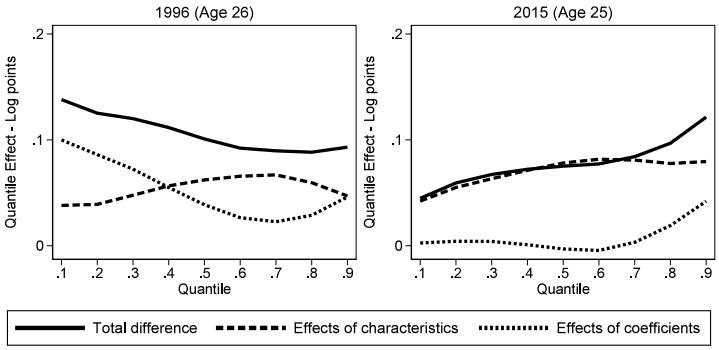
Differences in the proportion of females in the job nationally, ceteris paribus, explains a sizeable and growing component of the gender wage gap across both cohorts. This result implies that women (and men) employed in more female-dominated occupations are worse paid than other workers. Descriptive statistics in Tables B3 to B4 in Appendix B and estimates of the wage equation in Table C9 in Appendix C show that young women tend more than young men to work in female-dominated occupations while the returns to the proportion of women in the occupation becomes more negative over time.

One main takeaway of this analysis is that differences in the nature of jobs undertaken by men and women play an important role in explaining the otherwise unexplained part of the gender wage gap, but different aspects of those jobs have offsetting effects. Over time however, occupational segregation accounts for a rising proportion of the total gap for both graduates and non-graduates.

Turning to the full distribution of wages, the top two plots in Fig. 10 report the overall raw gap and the explained and unexplained component by deciles when type of occupation and proportion of females in the occupation at national level are added to the initial set of individual characteristics in the decomposition proposed by Chernozhukov et al. (2013). The two plots at the bottom of Fig. 10 include additionally hours worked. Confidence intervals for the total gap and the effect of characteristics are reported in Figure C2, Appendix C. By comparing Figs. 10 and 8 we can recover what component of the gap across the wage distribution is attributable to differences in job characteristics. The comparison shows that over time more of the gap across the wage distribution is explained by occupational segregation, while hours worked reduces the explained component of the gap in both cohorts.

[15] The excluded category for hours worked in the regression is 'less than 30 hours'.

(a) Early life factors, human capital characteristics, marital and parental status, indicators for major group of occupations, proportion of female in the occupation



(b) Early life factors, human capital characteristics, marital and parental status, indicators for major group of occupations, proportion of female in the occupation and weekly hours worked

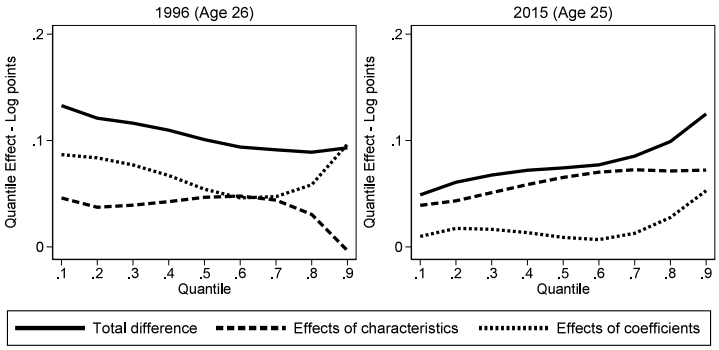


Fig. 10. Decomposition across the wage distribution.
Note: The two plots in the top panel report the decomposition between the 1st and 9th decile of the unconditional difference between the male and the female log wage distribution for the sample of observed wages in 1996 and 2015. The variables used in the decomposition include early years characteristics, qualifications, experience, parental and marital status, indicators for broad types of occupations (major groups in Standard Occupational Classification 2000 and 2010) and proportion of women at national level in detailed occupations (4 digits Standard Occupational Classification 2000 and 2010). The two figures in the bottom panel report the same decomposition where the controls additionally include hours worked. Confidence intervals are reported in Figure C2, Appendix C.

A visual inspection of the decomposition in the gap in 1996 (left plots) shows that differences in job characteristics account for less than half of the gender wage gap until the 7th decile. Between the 7th and 9th decile their role in explaining the gap is reduced and coefficients for individual and job attributes explain most of the gap. The decomposition in Next Steps is reported in the right plot: here differences in job characteristics explain most of the gap up to the 6th decile. After the 6th decile of the distribution of wages differences in coefficients of job characteristics begin to account for a small but rising component of the gap while the greatest part of the gender wage gap is still explained by differences in job characteristics, in particular by occupational segregation. These results are in line with findings from Levanon et al. (2009) and Blau and Kahn (2017) showing that female-dominated occupations pay less than male occupations for similar workers with the consequence that, over time, occupational segregation has become the main factor accounting for gender disparities in earnings as women caught up with men in terms of qualifications and experience.

4.6. Robustness

The results presented in this study account for non-random selection of young men and women into employment. However our empirical strategy can only account for selection on observables. Even with the broad set of variables available in the cohort studies and that are used in the matching procedure, selection on unobservables may still play an important role in the decision to work. For this reason we extend our analysis to include characteristics that are normally 'unobserved', such as measures of job aspirations and attitudes, in the estimation of the propensity to waged employment for two of the four cohorts, NCDS and Next Steps, whose members were interviewed in 1981 and 2015. Given the focus of this study on the gender wage gap, we consider factors for which there are gender differences that have already been established in the labour economics literature. We follow Fortin (2008) and consider three factors: self esteem, the attitude towards money and job progression and the value given to helping people and society for future career.

The NCDS and Next Steps collected comparable information during respondents' adolescence about self-esteem and the importance of money and caring for people for their future careers. We use these three sets of information to create three scores (self-esteem, money score and people score) following Fortin (2008) and add these new measures to the probits used to estimate the propensity for waged employment. An updated propensity score matching based on the nearest neighbour methodology is then implemented to find for each individual without a wage a new 'donor'.

Figure D1 in Appendix D shows the new results that differ only marginally from the main results reported in the paper in Fig. 4 for these two cohorts. This additional analysis shows that including individual characteristics that are normally unobserved in the estimation of the propensity to waged employment does not affect the main results obtained with an assumption of selection on observables.

A different issue is whether the component of the gap captured by job characteristics is partly reflecting differences in aspirations and other usually unobserved traits between men and women. To disentangle the contribution of non-cognitive traits as omitted variables from that of job characteristics, we focus on the most recent cohort, for which we observe both sets of variables. We first decompose the gap by adding three non-cognitive measures, as well as early years characteristics, qualifications, experience, parental and marital status. We finally add job characteristics, that is indicators for broad types of occupations (major groups in Standard Occupational Classification 2000 and 2010), proportion of women at national level in detailed occupations (4 digits Standard Occupational Classification 2000 and 2010). The results for these two additional decompositions across the wage distribution are presented in Figure D2 showing that differences in job preferences and self-esteem do explain some of the gap across the wage distribution. The remaining gap is then explained by job characteristics. By comparing these new results to the fourth panel in Fig. 8 and the second panel in Fig. 10 in the paper it emerges that the component explained by job characteristics is biased as it is also likely to capture the partial correlations with some unobserved characteristics such as job aspirations and self-esteem.

5. Discussion and conclusion

We argue that to better understand the convergence of men's and women's earnings it is important to study the gender wage gap at the beginning of the working life and how it evolved across cohorts through socio-demographic changes and the implementation of new policies.

This paper documents the gender wage gap among young adults across four cohorts born between 1946 and 1990. It compares the gender wage gap among graduates and non-graduates over time, and across the distribution of wages. It explores the role of changing women's selection into employment over time in explaining the convergence of women's and men's earnings. Lastly it presents the relative contribution of early life factors, human capital, family formation and job characteristics in accounting for the gender wage gap among young adults over two decades.

The main results show that the overall raw and covariate adjusted gap, as expected, narrows over time. What was not well known for Britain is that this shift is driven by convergence in non-graduate men's and women's wages. The gender wage gap among graduates increases between 1972 and 2015. This first evidence seems to confirm that graduate young women starting when their participation in higher education was particularly low and positive selection into employment was high, had advantages relative to young graduate men which were not sustained for those coming later. Unlike the graduates, non-graduates benefited considerably the equal opportunity measures of the 1970s and then from the National Minimum Wage legislation. In the last cohort we observe this reduced gender wage inequalities for very low paid jobs, as also shown by Amadxarif et al. (2020).

We find that non-random selection into the labour market affects relatively more graduate women than men in the earliest cohort, where they were least numerous. It also affects marginally the estimates of the gender wage gap for the cohort observed in 1981 while it does not significantly affect the estimates in other cohorts.

In line with previous literature, the findings from the decomposition of the gender wage gap show that differences in traditional human capital characteristics and marital and parental status do not explain any significant component of the gap among young workers or potential workers- except perhaps workers in 1972. It is perhaps too early in their life course for these factors to make much difference. The wage gap decreases over historical time and is mainly driven by a decrease in the unexplained component.

Looking across the wage distribution, we observe that gender inequalities were greater for those with the lowest wages in the first two cohorts. The opposite happens in the latest cohort where differences in wages are smallest among the lowest paid while the gap widens for higher wages. This last result is consistent with the trend in the mean log gender wage gap among graduates, who are likely the highest earners in the samples, and non-graduates.

The proportion of women in the occupation accounts for a sizeable part of the gender wage gap among young adults in the two most recent cohorts, and it accounts for most of the gap for graduates in the later cohort. This result suggests that women, and in particular high-skilled women, concentrate in occupations that are likely to have lower pay than more mixed occupations or occupations mainly employing men.

Overall, these findings show that young non-graduate females' and males' wages have been converging at a faster rate than the pay of graduate men and women. Changing selection of women into higher education may have contributed to these different rates of convergence. Changes in the returns to higher education are determined by shifts in the demand and supply of skilled labour as well as by changes in the quality of graduate workers (Blundell et al., 2007; Carneiro and Lee, 2011; Cavaglia and Etheridge, 2020).[16] Over the period considered in this paper the demand for skilled labour has risen, driven by technological changes. At the same time more women relative to men have accessed higher education and the employment rate of women, and in particular graduate women, has increased vis-a-vis that of men. These relative changes in the proportion of graduates and labour market participation have driven gender differences in the composition and quality of the skilled workforce that may in part explain the slower gender convergence of wages among graduates over the last two decades.

Occupational segregation is the main contributor to the lingering wage gap among young adults in recent years, as shown in this paper by the decomposition of the gender wage gap that includes job characteristics. This result is in line with Blau and Kahn (2017) who show that occupational segregation explains an increasing proportion of the gender wage gap over time in the US as women catch up with men in terms of human capital. Occupational segregation could be reinforced by women who perceive the greater risk of sexual harassment in mixed or predominantly male workplaces and therefore choose female segregated occupations (Folke and Rickne, 2022; Batut et al., 2021). As such, harassment should be tackled by employers with more determined efforts to change workplace culture. Occupational segregation could also be the result of women making pre-market choices that have lower returns: they may choose jobs that are more flexible and with lower wages, even before forming a family, with the expectation of needing less rigid working patterns (see for instance Chevalier, 2007; Wiswall and Zafar, 2018). In addition, the post-COVID-19 world is likely to introduce changes towards flexibility and teleworking for a broader set of occupations. Future research will have to establish whether these changes are likely to favour women and increase their number in jobs where they are less represented.

CRediT authorship contribution statement

Francesca Foliano: Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Alex Bryson:** Writing – original draft, Funding acquisition, Conceptualization. **Heather Joshi:** Writing – review & editing,

[16] As higher education enrolment expands, the overall quality of graduates is likely to decline, given that the marginal graduate is expected to be of lower quality compared to the average graduate (Carneiro and Lee, 2011).

Writing – original draft, Conceptualization. **Bożena Wielgoszewska:** Data curation. **David Wilkinson:** Data curation.

Data availability

The authors do not have permission to share data.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.labeco.2024.102614.

References

Albanesi, S., Olivetti, C., Petrongolo, B., 2022. Families, Labor Markets, and Policy. Tech. Rep., National Bureau of Economic Research.

Albrecht, J., Björklund, A., Vroman, S., 2003. Is there a glass ceiling in Sweden? J. Labor Econ. 21 (1), 145–177.

Amadxarif, Z., Angeli, M., Haldane, A., Zemaityte, G., 2020. Understanding pay gaps. Staff Working Paper No. 877, Bank of England.

Arulampalam, W., Booth, A.L., Bryan, M.L., 2007. Is there a glass ceiling over Europe? Exploring the gender pay gap across the wage distribution. ILR Rev. 60 (2), 163–186.

Azmat, G., Ferrer, R., 2017. Gender gaps in performance: Evidence from young lawyers. J. Polit. Econ. 125 (5), 1306–1355.

Azmat, G., Petrongolo, B., 2014. Gender and the labor market: What have we learned from field and lab experiments? Labour Econ. 30, 32–40.

Batut, C., Coly, C., Schneider-Strawczynski, S., et al., 2021. It's a Man's World: Culture of Abuse,# MeToo and Worker Flows. Università Bocconi, The Dondena Centre.

Bayard, K., Hellerstein, J., Neumark, D., Troske, K., 2003. New evidence on sex segregation and sex differences in wages from matched employee-employer data. J. Labor Econ. 21 (4), 887–922.

Benny, L., Bhalotra, S.R., Fernández, M., 2021. Occupation Flexibility and the Graduate Gender Wage Gap in the UK. Tech. Rep., ISER Working Paper Series.

Bertrand, M., Goldin, C., Katz, L.F., 2010. Dynamics of the gender gap for young professionals in the financial and corporate sectors. Am. Econ. J. Appl. Econ. 2 (3), 228–255.

Blau, F.D., Kahn, L.M., 2017. The gender wage gap: Extent, trends, and explanations. J. Econ. Liter. 55 (3), 789–865.

Blinder, A.S., 1973. Wage discrimination: reduced form and structural estimates. J. Human Res. 436–455.

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

Bryson, A., Joshi, H., Wielgoszewska, B., Wilkinson, D., 2020. A short history of the gender wage gap in Britain. Oxf. Rev. Econ. Policy 36 (4), 836–854.

Bütikofer, A., Jensen, S., Salvanes, K.G., 2018. The role of parenthood on the gender gap among top earners. Eur. Econ. Rev. 109, 103–123.

Carneiro, P., Lee, S., 2011. Trends in quality-adjusted skill premia in the United States, 1960–2000. Amer. Econ. Rev. 101 (6), 2309–2349.

Cavaglia, C., Etheridge, B., 2020. Job polarization and the declining quality of knowledge workers: Evidence from the UK and Germany. Labour Econ. 66, 101884.

Chernozhukov, V., Fernández-Val, I., Melly, B., 2013. Inference on counterfactual distributions. Econometrica 81 (6), 2205–2268.

Chevalier, A., 2007. Education, occupation and career expectations: determinants of the gender pay gap for UK graduates. Oxf. Bull. Econ. Stat. 69 (6), 819–842.

Combet, B., Oesch, D., 2019. The gender wage gap opens long before motherhood. Panel evidence on early careers in Switzerland. Eur. Sociol. Rev. 35 (3), 332–345.

Connolly, S., Gregory, M., 2002. The national minimum wage and hours of work: implications for low paid women. Oxf. Bull. Econ. Stat. 64, 607–631.

Cortes, P., Pan, J., 2018. Occupation and gender. In: The Oxford Handbook of Women and the Economy. Oxford University Press New York, pp. 425–452.

Costa Dias, M., Joyce, R., Parodi, F., 2020. The gender pay gap in the UK: children and experience in work. Oxf. Rev. Econ. Policy 36 (4), 855–881.

De Cooman, E., Ermisch, J., Joshi, H., 1987. The next birth and the labour market: A dynamic model of births in England and Wales. Popul. Stud. 41 (2), 237–268.

Denning, J.T., Jacob, B.A., Lefgren, L.J., vom Lehn, C., 2022. The return to hours worked within and across occupations: Implications for the gender wage gap. ILR Rev. 75 (5), 1321–1347.

Dolton, P., O'Neill, D., Sweetman, O., 1996. Gender differences in the changing labor market: the role of legislation and inequality in changing the wage gap for qualified workers in the United Kingdom. J. Hum. Resour. 549–565.

Ferri, E., Bynner, J., Wadsworth, M., 2003. Changing Britain, Changing Lives. Institute of Education Press.

Folke, O., Rickne, J., 2022. Sexual harassment and gender inequality in the labor market. Q. J. Econ. 137 (4), 2163–2212.

Fortin, N.M., 2008. The gender wage gap among young adults in the united states the importance of money versus people. J. Hum. Resour. 43 (4), 884–918.

Francesconi, M., Parey, M., 2018. Early gender gaps among university graduates. Eur. Econ. Rev. 109, 63–82.

Goldin, C., 2021. Career and Family. Princeton University Press.

Goldin, C., Katz, L.F., 2008. Transitions: Career and family life cycles of the educational elite. Amer. Econ. Rev. 98 (2), 363–369.

Groshen, E.L., 1991. The structure of the female/male wage differential: Is it who you are, what you do, or where you work? J. Hum. Resour. 457–472.

Harkness, S., 1996. The gender earnings gap: evidence from the UK. Fiscal Stud. 17 (2), 1–36.

Heckman, J.J., Stixrud, J., Urzua, S., 2006. The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. J. Labor Econ. 24 (3), 411–482.

Jann, B., 2008. The Blinder–Oaxaca decomposition for linear regression models. Stata J. 8 (4), 453–479.

Kitagawa, E.M., 1955. Components of a difference between two rates. J. Am. Stat. Assoc. 50 (272), 1168–1194.

Kleven, H., Landais, C., Posch, J., Steinhauer, A., Zweimüller, J., 2019. Child penalties across countries: Evidence and explanations. In: AEA Papers and Proceedings, vol. 109, American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203, pp. 122–126.

Kunze, A., 2018. The gender wage gap in developed countries. In: The Oxford Handbook of Women and the Economy. Oxford University Press New York, pp. 369–394.

Levanon, A., England, P., Allison, P., 2009. Occupational feminization and pay: Assessing causal dynamics using 1950–2000 US census data. Social Forces 88 (2), 865–891.

Machin, S., 1994. Changes in the Relative Demand for Skills in the UK Labour Market. Tech. rep., CEPR Discussion Papers.

Manning, A., 1996. The equal pay act as an experiment to test theories of the labour market. Economica 191–212.

Manning, A., Swaffield, J., 2008. The gender gap in early-career wage growth. Econ. J. 118 (530), 983–1024.

Medical Research Council, 2016. National Survey of Health and Development. http://dx.doi.org/10.5522/NSHD/Q101, MRC Unit for Lifelong Health and Ageing at UCL.

Mulligan, C.B., Rubinstein, Y., 2008. Selection, investment, and women's relative wages over time. Q. J. Econ. 123 (3), 1061–1110.

Neuburger, J., 2010. Trends in the Unequal Pay of Women and Men Across Three British Generations (Ph.D. thesis). Institute of Education, University of London.

Neuburger, J., Kuh, D., Joshi, H., 2011. Cross-cohort changes in gender pay differences in Britain from 1972 to 2004: accounting for selection into employment using wage imputation. Longitud Life Course Stud 2 (3), 260–285.

Neumark, D., 2004. Employers' discriminatory behavior and the estimation of wage discrimination. In: Sex Differences in Labor Markets. Routledge, pp. 163–177.

Oaxaca, R., 1973. Male-female wage differentials in urban labor markets. Int. Econ. Rev. 693–709.

Stewart, M.B., Swaffield, J.K., 2002. Using the BHPS wave 9 additional questions to evaluate the impact of the national minimum wage. Oxf. Bull. Econ. Stat. 64, 633–652.

University College London, 0000. National Child Development Study [data collection] UK Data Service, SN: 2000032, https://beta.ukdataservice.ac.uk/datacatalogue/series/series?id=2000032,

University College London, 0000. 1970 British Cohort Study [data collection] UK Data Service, SN: 2000001 https://beta.ukdataservice.ac.uk/datacatalogue/series/series?id=200001,

University College London, 0000. Next Steps (previously the Longitudinal Study of Young People in England (LSYPE1)) UK Data Service, SN: 2000030 https://beta.ukdataservice.ac.uk/datacatalogue/series/series?id=2000030,

Wiswall, M., Zafar, B., 2018. Preference for the workplace, investment in human capital, and gender. Q. J. Econ. 133 (1), 457–507.

Zabalza, A., Tzannatos, Z., 1985. The effect of Britain's anti-discriminatory legislation on relative pay and employment. Econ. J. 95 (379), 679–699.