Explanations for the Gender Earnings Gaps

By JIANI KANG

Student No. 1902013

Supervisor MARCO FRANCESCONI

Abstract: Using the UK Labour Force Survey data in 2018 and 2020, I assess the extent of the gender wage differentials at the mean by the OLS estimators, and the results indicate an increasing trend over this period. Results from the quantile regressions confirm the 'glass ceilings' phenomenon for the higher-paid women. The single most prominent aspect that explains the gap is the part-time work pattern, and above half of the gap remains unexplained, suggested by the decomposition outcomes.

UNIVERSITY OF ESSEX

APRIL 2021

Table of Contents

1. Introduction

The wage differential between women and men has been a controversial issue. Existing literature suggests a remarkable convergence in the wage gap since the last century in many countries, a large portion of which is due to the decline in the differences in the wage determinants, while the disparities in the returns to the wage determinants remain statistically high. This paper investigates the link between gender and labour market earnings and examines how much of the gap can be explained by the wage determinants included in the model.

In what follows, I consider education, industry, experience, the number of children in a household, marital state, workplace region, and ethnicity as the wage determinants to assess the extent of the gender differentials. Besides, quantile regressions are applied to observe how the earnings gaps differ across the wage distribution, and the phenomenon of 'glass ceilings' is verified. Moreover, I examine the most critical contributor in explaining the gap by decomposition analysis and find that the part-time working pattern plays the most salient role. Significant results also indicate that the gap is driven mainly by the coefficient effects rather than the endowment effects, indicating the existence of labour market discrimination against females. However, due to omitted variable bias and potential simultaneity, the coefficients of female and decompositions suffer from endogeneity problem, leading to failure in inferring causality and concluding the whole unexplained part as discrimination.

2. Literature Review

My paper relates to the literature on the drivers of gender wage gaps. Established literature has typically found that even when women have identical labour market attributes to men, women do not earn the same amount as men. Most studies have produced estimates of an explained portion and an unexplained portion to explain the gender differentials. The explained part results from differences in the measurable wage determinants, such as human capital, occupations, and family labour division. Whereas differences in aspects that lack solid quantitative evidence, like psychological attributes, gender identity norms, and returns to the wage determinants, are cited as potential sources of the unexplained part. Overall, there has been a noticeable decline in the earnings gap across many countries over the last century. Estimates from the US and the UK suggest that there has been a 'grand convergence' of the explained wage gap across cohorts, meaning that the educational levels, participation rates, working hours, experience, and occupations of men and women in the labour market have converged dramatically (Goldin, 2014; Manning and Swaffield, 2008). Therefore, the contribution made by the unexplained part has become relatively more responsible for the remaining gap (Blau and Kahn, 2017). Moreover, the unexplained part is even more pronounced for highly educated and paid females in many countries, which refers to the existence of the ‘glass ceilings’ (Albrecht et al., 2003; Arulampalam et al., 2007).

2.1 Human capital

The Neoclassical economic theory of human capital states that wage differentials between individuals can be explained by human capital differences (Becker, 1980). Analysing data from the US over the 1980-2010 period, Blau and Kahn (2017) suggest that recently a situation has been reached where there are few differences in terms of human capital between women and men. This reduction can be primarily accounted for by a reversal of the gender differences in educational levels and a substantial reduction in the gap of labour market experience.

2.1.1 Education

Goldin et al. (2006) and Goldin and Katz (2007) document that women have not only been advancing to higher education at higher rates than men, eventually overtaking their male counterparts, but that women have also started majoring in fields that are more rewarded in the labour market, especially since the mid-1970s in the US. This reversal in educational attainment has contributed significantly to closing the gap when men and women just straight out of university. However, although the obstacle of reaching the same level of education has now been removed, there is still a tiny gap at the outset of their careers. Evidence from Germany suggests that the essential proximate sources may be the field of study at university, with the largest early gap emerging among graduates from economics, business and STEM subjects (Francesconi and Parey, 2018).

2.1.2 Labour market experience

Blau and Kahn (2017) provide various reasons for the rapid increase in women's employment rates from 1947 to the 1990s, including a rise in real wages, increasing educational attainment, greater availability of market substitutes for housework and improvements in household technology, the development and dissemination of the birth control pill, and demand shifts in occupations where women were well represented. However, a plateau was reached, even for university graduate females, in the 1990s. Blau and Kahn (2013) suggest that this plateau is related to policies on the length of paid maternity leave. Such policies lead to more career disruption on the labour supply side and discourage women from having full-time jobs. From the labour demand perspective, employers are more likely to anticipate that women will take advantage of the maternity leave opportunities, resulting in discrimination during the employment selection process.

Using LFS data from 1993 to 2016, Costa Dias et al. (2016) observe that the gender pay gap expands over the life cycle. By tracking people longitudinally across education groups, they find that women are more likely to leave paid employment or move to part-time jobs to care for children, which leads to less accumulated job experience, more career interruptions, shorter working hours, and ultimately a substantial decline in earnings, especially for the highly educated women. Conversely, men's work pattern after the birth of a child is essentially unaffected. Moreover, evidence shows that the gap remains very persistent as women are significantly more likely to stay in jobs where they work fewer paid hours, except for lower-educated women, because they have less wage progression to miss out on or fewer skills to depreciate. However, due to potential selection bias (women who work fewer hours may have experienced slower wage growth even if they work longer), the correlation between women's wages and their career patterns (i.e., employment and choice of hours) does not imply a causal relationship. The same pattern has been observed in the US for female MBAs, except for those with lower-earning husbands (Bertrand et al., 2010).

2.2 Occupations

The occupation remains a salient factor in explaining the gap. Bertrand et al. (2010) find that female MBAs appear to have a more difficult time balancing career and family than female physicians, PhDs, and lawyers across all of these BA classes, which indicates that the differences in production technologies and the organisation of work may result in greater costs in the business and corporate sectors than in medicine or academia in terms of discontinuous experience and more flexible hours, and hence the disparities in the wages they are paid.

Goldin (2014) finds a common feature in the large-wage-gap occupations, where persistence and the amount of continuous time committed are vital for achieving higher earnings. Apart from longer working hours, Goldin (2014) also emphasises the role played by flexibility. For instance, the business and financial sectors value those not only able to work long hours, but also at whatever times necessary and those who are prepared to accept an on-call schedule. Besides, these jobs usually compensate employees for this disamenity by paying higher wages per hour. However, men tend to work jobs with these features, and consequently, female employees earn less than their male counterparts. Therefore, Goldin (2014) suggests that if firms do not have the incentive to reward employees who work long and specific hours disproportionately, the substantial costs of women's temporal flexibility, or controlling men's hours, would decrease, so that the gap would be considerably reduced or even eliminated. Thus, policies designed to create effective teams of substitutes to make all hours worked equally valuable should be implemented.

2.3 Family labour division

Becker (1981) argues that, the whole family would benefit from the better division of labour, and whoever is more productive in housework should be the homemaker. However, choosing to devote a substantial proportion of time to one's family and housework can be economically risky for the 'homemaker', which women usually undertake, even they are just as productive as men in doing housework (Bergmann, 1981). With the rapid development of the market and technology, much housework can now be outsourced or performed by robots, which mitigates the trouble of weighing family against work. However, the market itself can exacerbate existing inequalities between those who can access the increasing standard of living afforded by employment and those who cannot earn enough to meet their caring responsibilities through the market (Himmelweit, 2007). Furthermore, not everything can be contracted out, like being parents. Therefore, one person within a couple needs to be willing to work the more flexible, less-remunerative job, which is typically shouldered by women again. Consequently, there is a 15-per cent wage penalty for having more than one child, known as the motherhood wage penalty (Anderson et al., 2002). When analysing the earnings gap results from occupations and family labour division, a key question is why women are more likely to take the types of jobs that allow them to be at home at certain times and days and why women tend to devote more hours to their families than men do even when they have the same home-activities productivity? Evidence manifests that inherited and traditional norms are essential factors in explaining the question.

2.4 Gender identity norms

As discussed in the section on occupations and labour division, family 'rational choice' can interfere with gender identity norms. Akerlof and Kranton (2002) proposed a model where women's utility would decrease when deviating from the social norms. Social norms, such as 'men should work in the labour force, and women should work in the home', 'men should have more right to work than women when jobs are scarce', and 'a working mother cannot build as warm and secure a relationship with her children as a mother who does not work' drive the variation of women's labour force participation as well as occupational segregation (Fortin, 2005). In addition, evidence from the US suggests that the women's sense of self about these norms does not appear to be significant in their labour market outcomes. By contrast, it is the man's attitudes in the median position of the wage distribution that counts for women's labour market outcomes (Charles et al. 2009). Moreover, if the norm ‘a man should earn more than his wife’ is violated in a household, the wife would compensate the utility loss via reducing labour supply or conducting more household chores and thus earns less than her potential. Furthermore, under marriage patterns like these, the couples are less happy, report more significant strife in their marriage, and are ultimately more likely to get a divorce (Bertrand et al., 2015).

2.5 Psychological attributes

Fuchs (1988) illustrated that women give up possible income to do housework and care for children because they derive more utility as compensation, which implies that women have greater natural preferences in housework and caring for children. Thus, the gendered division of labour does not necessarily induce negative consequences. However, the preference for caring may be developed during caring for others (Ferguson, 1989), and this explanation cannot be justified with real-world market failures. For example, women may get paid wages below what they would have obtained under competitive equilibrium due to the collusive behaviour of men, which forces women into lower-paid jobs, so women have no choice (Bergmann, 1971). It has also been claimed that women may receive positive externalities, i.e., psychic income. However, Thurow (1978) documents that a higher psychic income does not necessarily compensate for lower wages due to the lack of trade between the two.

Niederle and Vesterlund (2007) and Croson and Gneezy (2009) suggest that women have less desire to compete and negotiate at the top levels of financial and business occupations and are also more risk-averse. Nevertheless, most of the results of these studies were obtained from laboratory or field experiments, and therefore questions still exist about how representative they are of the real world (Harrison and List, 2004). In addition, negotiation skills cannot explain why the wage gap is smallest when both women and men are at the beginning of their careers, since these interpersonal skills would not be lost over time. It cannot explain why there is still a gap in 'winner-take-all' jobs (for example, partner in a firm, tenured professor, top manager) as these jobs heavily reward competition. Besides, psychological attributes can affect conventional variables. For instance, risk-averse can impact women's choice in occupations, suggesting endogeneity problems when doing regression analysis.

3. Data

I use data from the UK Labour Force Survey in October-December 2018 and 2020. The sample contains roughly 9500 observations aged 16-69, and the data are mainly collected by telephone and face-to-face interviews.

3.1 Variable construction

1) lwage

The continuous dependent variable ‘lwage’ is constructed by taking the logarithm of the average gross hourly pay variable in the original dataset since I want to interpret coefficients of the regressors as percentage change to gauge the sizes of effects. Zero and negative wages are dropped from the sample.

2) female

The gender dummy equals 2 for female and 1 for male.

3) education

‘education’ represents the highest qualification respondents have obtained. There are six types in ascending order: No qualification; Other qualification; GCSE grades A*-C or equivalent; GCE A-level or equivalent; Higher education diplomas; Degree or equivalent. In the following analysis, I will always use respondents who have acquired the 'Degree or equivalent' qualification as the reference for interpretation. If the base category changes, the statistic numbers change, but the implications stay the same.

4) Potential experience

In some economic literature I have reviewed, working hours is used to represent working experience. However, working hours could be endogenous, resulting in biased OLS estimators. Because firstly, working hours is very likely to be correlated with the error term in the model, secondly, working hours can also be a function of wage, and thirdly I lack instrumental variable data. Hence, I construct a potential work experience predictor using age – education – 5.

5) industry

'Industry' suggests the industry sector of the respondents' main jobs. There are nine categories: Agriculture & fishing; Energy & water; Manufacturing; Construction; Distribution, hotels & restaurants; Transport & communication; Banking, finance & insurance; Public admin, educ & health; Other services. 'Banking, finance & insurance' is used as the base category.

6) partime

This binary variable equals 1 if respondents work full time and 0 part-time.

7) flexibility

This dummy variable demonstrates employment pattern, which equals 1 if employees can have variable start and finish times on their own and carry over debit and credit hours into another accounting period over an accounting period (usually four weeks or a calendar month). Additionally, it equals 0 if employees work under more intense patterns, such as annualised hours contract, term-time working, job sharing, nine-day fortnight, four-and-a-half-day week, zero-hours contract, on-call working schedule.

8) children

This variable displays the number of dependent children in a family aged under 19. Besides, no respondent replies to this question according to 2020 data, so this variable is not included in the 2020 OLS regression.

9) marital status

There are six types of marital status: Single, never married; Married, living with a spouse; Married, separated from a spouse; Divorced; Widowed; Currently or previously in a civil partnership. I choose 'Single' as the reference.

10) region

'region' shows the regions of the workplace of the respondents. There are thirteen categories: North East; North West; Yorkshire and Humberside; East Midlands; West Midlands; East of England; London; South East; South West; Wales; Scotland; Northern Ireland; Workplace outside the UK. I will use London as the base group in the following analysis.

11) ethnicity

White is used as the base category for 'ethnicity'. There are eight other dummy variables: Mixed/Multiple ethnic groups; Indian; Pakistani; Bangladeshi; Chinese; Any other Asian background; Black/African/Caribbean/Black British; Other ethnic groups (respondents in Northern Ireland identifying themselves as 'Irish Traveller' and respondents in all UK countries identifying themselves as 'Arab').

4. Model

My model is a variant on the human capital model initially proposed by Becker (Becker, 1980), and my model takes the form

$$ln(Y_{it}) = \beta_{it}X_{it} + \epsilon_{it}$$

where $ln(Y_{it})$ is the natural logarithm of the average gross hourly wage for observation *i* in year *t*, X_{it} is the vector of the independent variables mentioned above, β_{it} is the corresponding coefficient, and ϵ_{it} represents the error term. I assume the model satisfy the Gauss-Markov assumptions.

5. Estimations and Findings

I proceed in three steps. I begin with analysing the wage gaps at the mean in 2018 and 2020, using the OLS estimators, and then examining gender effects across the wage distribution in 2018 via quantile regressions. Finally, I perform a decomposition of the gap at the mean in 2018 so that I can better assess the role played by each control.

5.1 OLS regression analysis

Likelihood-ratio test	LR chi2(3) = 270.13
(Assumption: m1 nested in m2)	Prob > chi2 = 0.0000

Ramsey RESET test using powers of the independent variables
Ho: model has no omitted variables
F (16, 9708) = 37.89
Prob > F = 0.0000

The Likelihood-ratio test suggests introducing non-linear terms age2, age3 and children2. The Ramsey test suggests the endogeneity problem. Therefore, the Gauss-Markov assumptions are failed, and the OLS estimators are biased.

Table 1: 2018 OLS regression

	(1) Raw gap	(2) Pooled lwage	(3) Female lwage	(4) Male lwage
female	-.177***	-.111***		
	(.011)	(.011)		
Degree		0	0	0
Higher education diplomas		-.163***	-.187***	-.125***
		(.02)	(.025)	(.031)
GCE A level		-.191***	-.228***	-.141***

	(.024)	(.032)	(.035)
GCSE grades A*-C	-.223***	-.234***	-.191***
	(.033)	(.044)	(.049)
Other qualification	-.24***	-.226***	-.228***
	(.043)	(.057)	(.064)
No qualification	-.241***	-.275***	-.177**
	(.052)	(.071)	(.078)
Agriculture & fishing	-.399***	-.337***	-.427***
	(.055)	(.087)	(.072)
Energy & water	.041	.088	.034
	(.037)	(.076)	(.045)
Manufacturing	-.09***	-.122***	-.073***
	(.018)	(.03)	(.024)
Construction	-.09***	-.03	-.105***
	(.025)	(.051)	(.031)
Distribution, hotels & restaurants	-.261***	-.286***	-.235***
	(.016)	(.022)	(.025)
Transport & communication	-.096***	-.066*	-.1***
	(.023)	(.038)	(.029)
Banking & finance	0	0	0
Public admin, educ & health	-.177***	-.191***	-.149***
	(.014)	(.018)	(.023)
Other services	-.25***	-.213***	-.291***
	(.024)	(.032)	(.036)
Potential experience	.068***	.066***	.075***
	(.009)	(.013)	(.014)
partime	-.17***	-.141***	-.206***
	(.012)	(.014)	(.026)
flexibility	-.09***	-.085***	-.098***
	(.015)	(.019)	(.023)
children	.032	.021	.084***
	(.023)	(.014)	(.016)
children2	-.016***	-.008*	-.024***
	(.003)	(.004)	(.005)
Single, never married	0	0	0
Married, with spouse	.106***	.083***	.125***

	(.013)	(.016)	(.02)
Married, separated	.037	.031	.049
	(.031)	(.039)	(.05)
Divorced	.017	.006	.027
	(.02)	(.025)	(.034)
Widowed	-.016	-.045	.001
	(.042)	(.047)	(.082)
Civil partnership	.239***	.162	.305**
	(.088)	(.132)	(.119)
North East	-.326***	-.269***	-.373***
	(.026)	(.035)	(.04)
North West	-.308***	-.24***	-.375***
	(.021)	(.027)	(.031)
Yorkshire and Humberside	-.304***	-.257***	-.347***
	(.022)	(.029)	(.032)
East Midlands	-.301***	-.233***	-.364***
	(.023)	(.03)	(.034)
West Midlands	-.292***	-.24***	-.339***
	(.022)	(.029)	(.033)
East of England	-.246***	-.21***	-.278***
	(.021)	(.029)	(.032)
London	0	0	0
South East	-.225***	-.178***	-.265***
	(.02)	(.026)	(.029)
South West	-.305***	-.261***	-.345***
	(.021)	(.028)	(.032)
Wales	-.351***	-.234***	-.461***
	(.026)	(.035)	(.04)
Scotland	-.248***	-.196***	-.297***
	(.022)	(.03)	(.034)
Northern Ireland	-.358***	-.282***	-.434***
	(.026)	(.034)	(.04)
Workplace outside the UK	-.062	.177	-.158
	(.109)	(.219)	(.131)
White	0	0	0
Mixed ethnic groups	-.013	.013	-.038

		(.051)	(.068)	(.075)
Indian		-.118***	-.08*	-.15***
		(.032)	(.042)	(.049)
Pakistani		-.192***	-.251***	-.14*
		(.051)	(.072)	(.073)
Bangladeshi		-.314***	-.231*	-.383***
		(.08)	(.132)	(.102)
Chinese		.04	.06	-.002
		(.062)	(.075)	(.104)
Other Asian background		-.179***	-.138**	-.224***
		(.048)	(.064)	(.071)
Black/African/Caribbean/Black British		-.218***	-.154***	-.267***
		(.032)	(.043)	(.049)
Other ethnic groups		-.086*	-.087	-.079
		(.044)	(.061)	(.064)
age2		-.001***	-.001***	-.001***
		(0)	(0)	(0)
age3		0***	0**	0*
		(0)	(0)	(0)
_cons	2.83***	2.763***	2.546***	2.622***
	(.018)	(.08)	(.105)	(.124)
Observations	9957	9766	5149	4617
R-squared	.024	.367	.351	.363

Standard errors are in parentheses

**** p<.01, ** p<.05, * p<.1*

Table 2: 2020 OLS regression

	(1)	(2)	(3)	(4)
	Raw gap	Pooled lwage	Female lwage	Male lwage
female	-.184***	-.133***		
	(.012)	(.011)		
Degree		0	0	0

Higher education diplomas	-.127***	-.16***	-.08**
	(.022)	(.028)	(.034)
GCE A level	-.144***	-.183***	-.088**
	(.026)	(.034)	(.041)
GCSE grades A*-C	-.13***	-.159***	-.071
	(.036)	(.046)	(.057)
Other qualification	-.134***	-.143**	-.085
	(.049)	(.064)	(.075)
No qualification	-.124**	-.158**	-.039
	(.059)	(.076)	(.092)
Agriculture & fishing	-.148**	-.308***	-.071
	(.061)	(.11)	(.075)
Energy & water	.029	.102	-.002
	(.043)	(.078)	(.054)
Manufacturing	-.063***	-.046	-.075***
	(.02)	(.032)	(.027)
Construction	-.061**	-.068	-.063*
	(.028)	(.054)	(.034)
Distribution, hotels & restaurants	-.253***	-.257***	-.254***
	(.018)	(.024)	(.027)
Transport & communication	-.097***	-.08**	-.104***
	(.024)	(.039)	(.031)
Banking & finance	0	0	0
Public admin, educ & health	-.155***	-.149***	-.15***
	(.015)	(.019)	(.023)
Other services	-.199***	-.206***	-.185***
	(.025)	(.033)	(.04)
Potential experience	.087***	.073***	.11***
	(.01)	(.013)	(.016)
partime	-.123***	-.102***	-.15***
	(.013)	(.014)	(.028)
flexibility	-.092***	-.068***	-.122***
	(.015)	(.019)	(.024)
Single, never married	0	0	0
Married, with spouse	.114***	.057***	.179***
	(.013)	(.017)	(.021)

Married, separated	.017	-.005	.028
	(.035)	(.042)	(.06)
Divorced	.064***	.024	.125***
	(.022)	(.026)	(.038)
Widowed	.08*	0	.232**
	(.046)	(.052)	(.092)
Civil partnership	.092*	.093	.071
	(.056)	(.066)	(.102)
North East	-.329***	-.337***	-.327***
	(.029)	(.036)	(.046)
North West	-.283***	-.301***	-.259***
	(.022)	(.029)	(.034)
Yorkshire and Humberside	-.295***	-.295***	-.299***
	(.024)	(.031)	(.037)
East Midlands	-.308***	-.323***	-.295***
	(.024)	(.031)	(.037)
West Midlands	-.284***	-.329***	-.238***
	(.024)	(.031)	(.036)
East of England	-.246***	-.254***	-.236***
	(.023)	(.03)	(.036)
London	0	0	0
South East	-.213***	-.263***	-.153***
	(.021)	(.027)	(.032)
South West	-.302***	-.333***	-.261***
	(.023)	(.03)	(.035)
Wales	-.331***	-.359***	-.306***
	(.029)	(.038)	(.045)
Scotland	-.268***	-.276***	-.259***
	(.025)	(.033)	(.039)
Northern Ireland	-.366***	-.338***	-.4***
	(.026)	(.034)	(.041)
Workplace outside the UK	-.134	.005	-.15
	(.12)	(.267)	(.141)
White	0	0	0
Mixed ethnic groups	.025	.014	.029

		(.051)	(.061)	(.086)
Indian		-.02	-.018	-.024
		(.035)	(.046)	(.054)
Pakistani		-.168**	-.078	-.245**
		(.071)	(.103)	(.098)
Bangladeshi		-.098	-.035	-.163
		(.102)	(.147)	(.144)
Chinese		.016	.048	-.007
		(.078)	(.109)	(.113)
Other Asian		-.177***	-.261***	-.054
background		(.055)	(.071)	(.085)
Black/African/Caribbean/Black		-.177***	-.168***	-.213***
British		(.04)	(.049)	(.066)
Other ethnic groups		-.13**	-.081	-.194**
		(.053)	(.07)	(.081)
age2		-.001***	-.001***	-.002***
		(0)	(0)	(0)
age3		0***	0*	0***
		(0)	(0)	(0)
_cons	2.955***	2.612***	2.451***	2.323***
	(.019)	(.09)	(.113)	(.146)
Observations	9466	9349	4990	4359
R-squared	.024	.315	.298	.309

Standard errors are in parentheses

*** *p<.01,* ** *p<.05,* * *p<.1*

Table 1 and Table 2 present estimates of the gender wage gap in 2018 and 2020, which suggests that the inclusion of the observables substantially reduces, but does not eliminate, the gender earnings gap. In 2018 it drops from 17.7% to 11.1%, suggesting a decline of more than one third. In 2020, it goes down from 18.4% to 13.3%, indicating an increasing trend over this period.

The results show that, in 2018, females faced higher wage loss at almost every qualification level relative to their male counterparts, except for the ‘Other qualification' level, in which women gained a slight advantage of 0.2% over men. The largest difference occurs in the 'No qualification' group, revealing that females without any qualifications would, on average, earn 9.8% less than males. In 2020, the relative differences between females and males ascend at every qualification level, especially for low-level qualifications holders, although the positive impact of having qualifications on wages has decreased for both groups in absolute terms over this time.

Based on the banking and finance industry, except for the energy and water sector, the other industries pay lower wages to women and men alike in both years. Additionally, the significant statistics indicate that females are particularly disadvantaged in the manufacturing sector, with a gap of 4.9% in 2018. The larger magnitude of the impact reconciles the study by Olivetti and Petrongolo (2016), which emphasises the role of the shift in industry structure, involving a change from manufacturing to services, which might have enhanced women's employment and reduced the earnings gap. Interestingly, people working in the agriculture and fishing industry on average earn 39.9% less than bankers in 2018 and just 14.8% less in 2020. Meanwhile, females previously enjoy a 9% wage advantage but conversely face a 23.7% disadvantage in 2020.

Work experience plays an increasingly important role for males than females, with a gap of 0.9% in 2018 and 3.3% in 2020. In terms of employment patterns, part-time workers experience less wage loss in 2020 than 2018, and the data suggests that part-time women, on average, perform better than comparable part-time men. Regarding job flexibility, it appears that asking for more flexibility does reduce wages for both groups and women on average bear lower-wage losses than men.

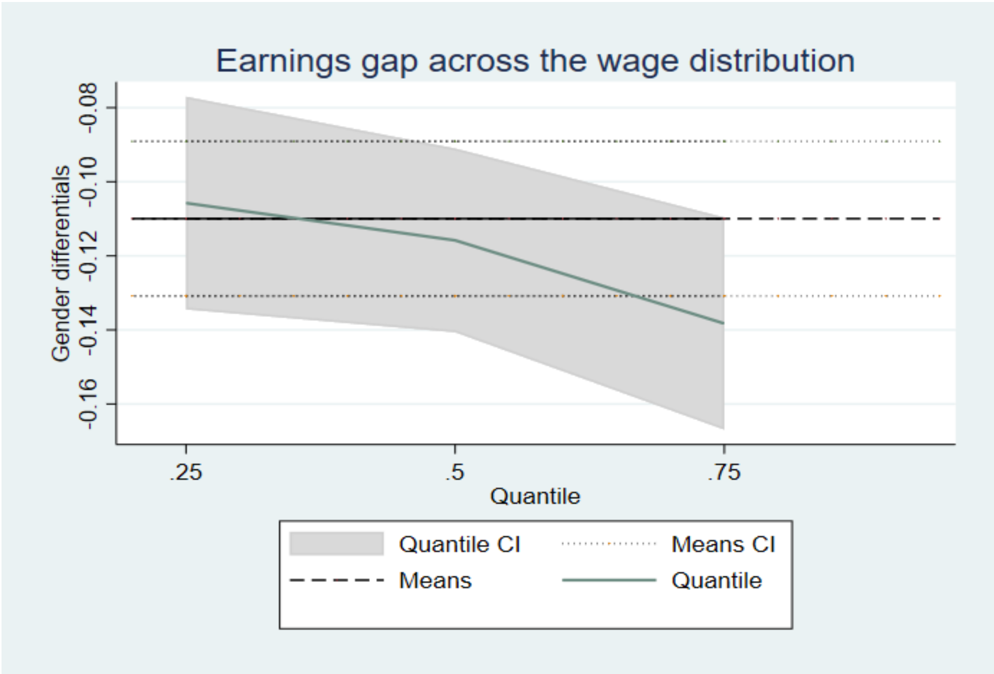
The estimates show that the coefficient of children is positive and the coefficient of $children^2$ is negative for both groups, which suggests that wages increase with having one more child but declines after reaching a certain threshold. Introducing the interaction term children & educ, I find that highly educated women experience wage penalties for having children. In contrast, highly educated men experience a wage premium, which is consistent with the work interruption effects illustrated by Costa Dias et al. (2016). Concerning marital status, it is suggested that being single is only better than being widowed in 2018 and worse than being widowed in 2020 in terms of wages. In 2018, both groups – men and women - enjoy the highest wage in a civil partnership, but the gap is also the largest for women (14.3%). Conversely, women in a civil partnership earn 2.2% higher wages than men in 2020. Additionally, divorced women encounter an expanding gap relative to legally married women who have separated from their spouse, rising from 2.1% in 2018 to 10.1% in 2020.

Regarding workplace location, London is the 'golden' place to work in because it offers wages that are higher than all the other areas. Furthermore, on average, women earn higher wages in all locations in 2018, especially in Wales, where women have a wage advantage of 22.7%. However, in 2020, only women working in Yorkshire and Humberside, Northern Ireland, and outside the UK earn more than comparable males, with the largest earnings gap of 11% in the South East.

In terms of ethnicity, in 2018, on the one hand, Indian, Pakistani, Bangladeshi, Other Asian backgrounds, and Black/African/Caribbean/Black British females earn higher wages than their male counterparts, with the most significant advantage of 15.2% for Bangladeshi women. On the other hand, the earnings gap was largest for Pakistani women (11.1%). Furthermore, the average gap between white people and people from other Asian backgrounds and Black/African/Caribbean/Black British people has narrowed during this period, which mainly results from the decreasing gap among males, but the gaps among females from the same backgrounds have risen by 12.3% and 1.4% respectively.

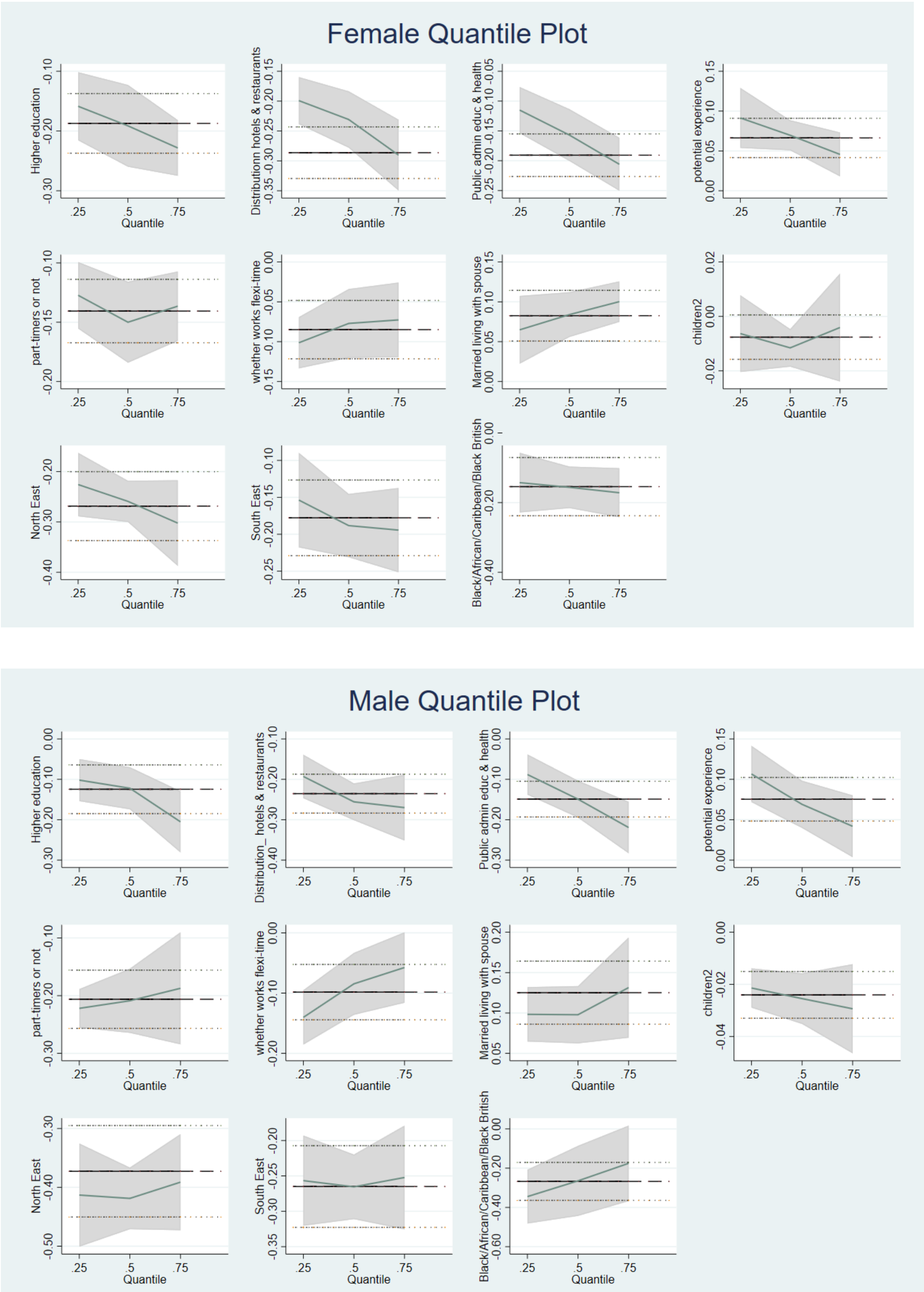
5.2 Quantile regression analysis

Figure 1: Gender differences across the wage distribution



To measure the earnings gap across the wage distribution, I classified the observations into three quantiles based on their wage levels: the bottom 25%, the median 50%, and the upper 75%. Figure 1 shows that the gap expands from 10.6% to 13.8% along the wage distribution, which is in line with studies from Sweden that women and men wages become extremely unequal among top earners (Albrecht et al., 2003).

Figure 2: Quantile plots of significant controls for both groups



Note: This figure reports estimates that are significant at 90% confidence level.

The estimates suggest that the impact of obtaining higher education diplomas is most prominent for females and males in the upper quantile. Additionally, females without higher education diplomas experience a greater wage loss than males in every quantile, with the largest gender gap of 7.6 log points found in the median 50%.

Regarding industry, the wage loss gradient increases for both women and men along their wage distributions in the 'Public administration, education & health' sector. However, the wage gap decreases across the quantiles, starting from 3.5 log points for the bottom 25% to 0.9 log points for the median 50%, and the upper quantile women have a 1.5% wage advantage compared with their male counterparts. Concerning 'Distribution, hotel & restaurant' sector, the gradient of the negative effect is stronger for higher-paid women, while for lower-paid men, and it is notable that women at the lower and upper quantile experience more severe loss than males, with the largest gap of 2.1% for women at the higher wage distribution.

The positive effect of work experience decreases from 9.3% to 4.5% for female employees and from 10.5% to 4.2% for males. Besides, accumulating work experience becomes relatively more important for women than men at the top of the wage distribution. The negative impact of working part-time on wage descends for males as wage increases, but increases for females at first, and then declines to a lower level than their comparable males. Furthermore, females in the 25% quantile, while males in the 75% quantile experience the lowest loss, compared with other quantiles in each group. Overall, females in each quantile experience less wage loss than comparable male employees, which presumably is due to the employers' biases from social norms like 'women are not penalised for part-time work because they are expected to be second earners' as presented in Akerlof and Kranton (2002), but the advantage goes down from 10.3 log points to 5.0 log points. The negative effect of working flexibly decreases for both groups along the wage distribution, which may arise from the non-substitutable feature of highly paid jobs. The negative impact falls from 10.1% to 7.1% for women and drops from 13.5% to 5.5% for males. Hence, the evidence suggests that women working in highly paid jobs bear greater wage losses for not being able to work to an on-call schedule, and this finding corroborates those in the literature, as illustrated by Goldin (2014).

One pair of significant results comes from the median 50% concerning the number of children in a household. Males experience a 1.4% higher wage loss due to having one more child in the household. In addition, getting married and living with a spouse has a positive effect on both groups. However, in every quantile, women benefit less from it, especially the upper 75% quantile of women, who benefit by 3.4% less than comparable males.

The negative effect of belonging to the 'Black/African/Caribbean/Black British' group is more significant for females than males along the wage distribution, increasing from 13.8% to 18%. Conversely, the effect for men decreases, ranging from 32.2% to 17.4%. Furthermore, females in the lower 25% and median 50% positions are less affected by ethnic characteristics, but the reverse is true for the upper 75% of Black/African/Caribbean/Black British females, who earn 0.6% less than males from the same background.

5.3 Oaxaca-Blinder decomposition analysis

5.3.1 Methods description

Following methods described in Fortin and Lemieux (2011), I find that intuitively, OB decomposition methods decompose observed differences of wage between females and males into an explained part and an unexplained part. The explained component is associated with characteristics differences between women and men, which means differences in wage determinants. The unexplained part results from coefficients differences between the groups, indicating different returns to wage determinants for men and women or potential discrimination against women.

However, since one observed wage structure, not the non-discriminatory wage structure, is used as a counterfactual for another group in OB decomposition. Thus, OB decomposition methods inherently follow a partial equilibrium, which gives rise to the endogeneity problem, and the results cannot fully infer causal relationships.

Technically, OLS regression results of the male group (M) and the female group (F) from Table 1 and Table 2 can be represented as follows

$$\bar{Y}_M = \hat{\beta}_M \overline{X'}_M \quad \bar{Y}_F = \hat{\beta}_F \overline{X'}_F$$

The potential non-discriminatory wage structure can be shown as follows

$$Y^* = X\beta^* + \epsilon$$

The difference of the wages at the means can be decomposed as follows

$$\bar{Y}_M - \bar{Y}_F = \hat{\beta}_M \overline{X'}_M - \hat{\beta}_F \overline{X'}_F$$

$$=\hat{\beta}_M \overline{X'}_M - \hat{\beta}^* \overline{X'}_M + \hat{\beta}^* \overline{X'}_M - \overline{X'}_F \hat{\beta}^* + \overline{X'}_F \hat{\beta}^* - \hat{\beta}_F \overline{X'}_F$$

$$= \hat{\beta}^* (\overline{X'}_M - \overline{X'}_F) + [\overline{X'}_M (\hat{\beta}_M - \hat{\beta}^*) + \overline{X'}_F (\hat{\beta}^* - \hat{\beta}_F)]$$

The first term is the explained part, the second term is the unexplained part, which can be further divided into discrimination in favour of males and discrimination against females.

5.3.2 Basic OB decomposition

The difference between the primary OB and the omega weighted or pooled weighted OB is how the non-discriminatory coefficients β^* is determined. Modern methods are trying to make β^* a more accurate indicator of the potential non-discriminated wage structure.

Blinder (1973) and Oaxaca (1973) assumes that there is only one reference group that represents the non-discriminated base, which corresponds to either $\beta^{*} = \beta_{M}$ or $\beta^{*} = \beta_{F}$. Thus, the result depends on choosing women or men as the counterfactual. I performed two basic decompositions and found that the results are different, so I turned to weighted decompositions.

5.3.3 Omega weighted and pooled weighted OB decomposition

Oaxaca and Ransom (1994) and Cotton (1998) apply a combination of β_{M} and β_{F} to determine β^{*}. Reimers (1983) proposed using the average coefficients over both groups to estimate the non-discriminatory parameter vector ($\hat{\beta}^{*} = 0.5\hat{\beta}_{M} + 0.5\hat{\beta}_{F}$). Cotton (1988) suggests weighting the coefficients by the group sizes, n_{M} and n_{F} ($\hat{\beta}^{*} = \frac{n_{M}}{n_{M}+n_{F}}\hat{\beta}_{M} + \frac{n_{F}}{n_{M}+n_{F}}\hat{\beta}_{F}$). I applied the method mentioned by Neumark (1988), using the coefficients from a weighted matrix omega ($Y = X\beta$) over both groups as my reference coefficients ($\hat{\beta}^{*} = (X'X)^{-1}(X'Y)$).

However, as mentioned before, this OB method can inappropriately transfer parameters in the unexplained parts of the differential into the explained part, so there is omitted variable bias problem in my decomposition. Therefore, to address the endogeneity problem, I used a more modern method proposed by Jann(2008) and Fortin(2011) as well, which suggests estimating the pooled regression over both groups but controlling a gender dummy variable D ($Y = \beta^{*} + \delta D + \epsilon$). In this case, my reference coefficient now becomes $\hat{\beta}^{*} = ((X, D)'(X, D))^{-1}\ (X, D)'Y$, and the unexplained part is δ, the coefficient of the gender dummy variable in the pooled regression.

5.3.4 Decomposition analysis

I first applied the basic OB decomposition method with male and female as the reference, respectively, and the results suggest that the explained parts account for 34.3% and 37.1% for the gender wage differential. Therefore, I also used omega matrix OB and pooled OB methods. The former shows that above half of the raw wage gap in 2018 is accounted for by my observables, and the latter demonstrates that 37.6% of the differential can be explained. I base my analysis on the pooled OB decomposition in the following.

Table 3 Decomposition results

	(1) Male reference	(2) Female reference	(3) Omega weighted	(4) Pooled weighted
Overall				
Female group	2.480***	2.480***	2.480***	2.480***
	(330.50)	(330.50)	(330.79)	(330.79)
Male group	2.657***	2.657***	2.657***	2.657***
	(299.87)	(299.87)	(300.16)	(300.16)
Observed gender gap	-0.178***	-0.178***	-0.178***	-0.178***
	(-15.29)	(-15.29)	(-15.31)	(-15.31)
Explained	-0.00753	-0.0298***	-0.0650***	-0.0378***
	(-0.47)	(-3.58)	(-8.65)	(-4.87)
Unexplained	-0.170***	-0.148***	-0.113***	-0.140***
	(-9.35)	(-12.77)	(-12.54)	(-12.58)
Part explained by				
education	0.00467	0.00495*	0.00533**	0.00556**
	(1.10)	(2.30)	(2.82)	(2.88)
industry	-0.00255	-0.00183	-0.0118***	-0.00237
	(-0.27)	(-0.43)	(-4.08)	(-0.79)
experience	0.00528	0.00530	0.00515	0.00509
	(0.20)	(0.20)	(0.20)	(0.20)
part-time	-0.0185	-0.0496***	-0.0778***	-0.0611***
	(-1.57)	(-11.13)	(-18.92)	(-14.99)

flexibility	0.00155	0.00149*	0.00142*	0.00144*
	(1.61)	(2.17)	(2.23)	(2.24)
children	-0.00306	0.00112	0.00368*	0.00341*
	(-1.48)	(1.31)	(2.29)	(2.27)
children2	0.000636	-0.000425	-0.00101	-0.000992
	(0.69)	(-0.77)	(-0.84)	(-0.84)
marital	-0.00348	-0.000691	-0.000294	0.000442
	(-1.51)	(-0.74)	(-0.37)	(0.57)
region	0.0000577	0.0000663	0.0000737	0.0000711
	(0.20)	(0.20)	(0.20)	(0.20)
ethnicity	-0.0000640	0.000186	0.000274	0.000316
	(-0.17)	(0.84)	(1.06)	(1.10)
age2	0.0326	0.0299	0.0274	0.0266
	(0.70)	(0.72)	(0.72)	(0.72)
age3	-0.0247	-0.0203	-0.0174	-0.0162
	(-1.08)	(-1.17)	(-1.17)	(-1.17)
N	9766	9766	9766	9766

Notes:

1. This table reports decomposition results, based on the procedure developed by Blinder (1973) and Oaxaca (1973); Neumark (1988); Jann (2008) and Fortin (2011).

2. The dependent variable is log hourly wage.

3. t statistics in parentheses:

** p<0.05, ** p<0.01, *** p<0.001*

4. unexplained part's specific statistics are not reported for the sake of brevity.

The fourth column of the table 3 shows that the wage determinants in the model explain nearly a quarter of the gap. Working part-time is the most crucial factor among all the other controls for driving the gap, which accounts for 6.1% of the explained part. Education as the prominent human capital accumulation means, especially the difference in acquiring the GCE A level qualification, contributes 0.6% of the gap. Manufacturing & Construction and, Public administration & Education, and health are the most significant components in explaining the gap regarding the working industry. As for the marital status, choosing to get married and live with a spouse is the most crucial characteristic in explaining the gender wage differential. The decomposition also shows that workplace regions and ethnic backgrounds are not the main drivers of the gap, and the unexplained part remains statistically significant.

6. Conclusion

The extent of the earnings gap expands from 11.1% in 2018 to 13.3% in 2020, after controlling for wage determinants in the model. Additionally, the gender wage gap is exceptionally high for women at the top of the wage distribution. Besides, job interruptions brought from having children induce wage penalties for highly educated women, while the reverse is true for men. Aspects like obtaining higher education diplomas, accumulating work experience, being able to work in a less flexible schedule are more important for highly paid females than their male counterparts. Part-time employment contributes most in explaining the gap, which involves disproportionately more women than men who work part-time, even though, on average, women perform better than comparable part-time men accordingly to the OLS estimators. Other characteristics like marital state, workplace region, and ethnicity do not significantly associate with the explained portion. Furthermore, since not all wage determinants are included in the model, the gaps represented by OLS estimators are biased, and the coefficient effects indicated by the decompositions cannot be referred to as discrimination.

7. References

1. Akerlof, G.A. and Kranton, R.E., 2002. Identity and schooling: Some lessons for the economics of education. *Journal of economic literature, 40*(4), pp.1167-1201.

2. Albrecht, James, Anders Bj¨orklund, and Susan Vroman. 2003. "Is There a Glass Ceiling in Sweden?" *Journal of Labor economics, 21*(1): 145–177.

3. Anderson, D.J., Binder, M. and Krause, K., 2002. The motherhood wage penalty: Which mothers pay it and why?. *American economic review, 92*(2), pp.354-358.

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

5. Becker, G.S., 1980. *Human capital* (Vol. 3). Chicago: University of Chicago press.

6. Becker, G.S., 1981. Altruism in the Family and Selfishness in the Market Place. *Economica, 48*(189), pp. 1-15.

7. Bergmann, B.R., 1971. The effect on white incomes of discrimination in employment. *Journal of Political Economy, 79*(2), pp. 294-313.

8. Bergmann, B.R., 1981. The economic risks of being a housewife. *The American economic review*, *71*(2), pp. 81-86.

9. Bertrand, M., Goldin, C. and Katz, L.F., 2010. Dynamics of the gender gap for young professionals in the financial and corporate sectors. *American economic journal: applied economics*, *2*(3), pp. 228-55.

10. Bertrand, M., Kamenica, E. and Pan, J., 2015. Gender identity and relative income within households. *The Quarterly Journal of Economics, 130*(2), pp.571-614.

11. Blau, F.D. and Kahn, L.M., 2013. Female labor supply: Why is the United States falling behind?. *American Economic Review*, *103*(3), pp. 251-56.

12. Blau, F.D. and Kahn, L.M., 2017. The gender wage gap: Extent, trends, and explanations. *Journal of Economic Literature, 55*(3), pp. 789-865.

13. Blundell, R., Costa Dias, M., Meghir, C. and Shaw, J., 2016. Female labor supply, human capital, and welfare reform. *Econometrica*, *84*(5), pp. 1705-1753.

14. Charles, Kerwin, Guryan, Jonathan, Pan, Jessica, 2009. *Sexism and women's labor market outcomes*. Working Paper, University of Chicago.

15. Costa Dias, M., Elming, W. and Joyce, R., 2016. The gender wage gap. *IFS Briefing Note*.

16. Costa Dias, M., Joyce, R. and Parodi, F., 2018. *The gender pay gap in the UK: children and experience in work* (No. W18/02). IFS Working Papers.

17. Croson, R. and Gneezy, U., 2009. Gender differences in preferences. *Journal of Economic literature*, *47*(2), pp. 448-74.

18. Ferguson, A., 1989. Blood at the root: Motherhood, sexuality and male dominance.

19. Fortin, N., Lemieux, T. and Firpo, S., 2011. Decomposition methods in economics. In *Handbook of labor economics* (Vol. 4, pp. 1-102). Elsevier.

20. Fortin, N.M., 2005. Gender role attitudes and the labour-market outcomes of women across OECD countries. *oxford review of Economic Policy, 21*(3), pp.416-438.

21. Francesconi, M. and Parey, M., 2018. Early gender gaps among university graduates. *European economic review, 109*, pp.63-82.

22. Fuchs, V.R., 1988. *Women's quest for economic equality*. Harvard University Press.

23. Goldin, C., 2006. The quiet revolution that transformed women's employment, education, and family. *American economic review*, *96*(2), pp. 1-21.

24. Goldin, C., 2014. A grand gender convergence: Its last chapter. *American Economic Review, 104*(4), pp. 1091-1119.

25. Goldin, C. and Katz, L.F., 2007. *The race between education and technology: The evolution of US educational wage differentials, 1890 to 2005* (No. w12984). National Bureau of Economic Research.

26. Goldin, C., Katz, L.F. and Kuziemko, I., 2006. The homecoming of American college women: The reversal of the college gender gap. *Journal of Economic perspectives*, *20*(4), pp. 133-156.

27. Harrison, G.W. and List, J.A., 2004. Field experiments. *Journal of Economic literature*, *42*(4), pp. 1009-1055.

28. Himmelweit, S., 2007. The prospects for caring: economic theory and policy analysis.
Cambridge Journal of Economics, 31(4), pp. 581-599.

29. Manning, A. and Swaffield, J., 2008. The gender gap in early-career wage growth. *The Economic Journal*, *118*(530), pp. 983-1024.

30. Niederle, M. and Vesterlund, L., 2007. Do women shy away from competition? Do men compete too much?. *The quarterly journal of economics*, *122*(3), pp. 1067-1101.

31. Olivetti, C. and Petrongolo, B., 2016. The evolution of gender gaps in industrialised countries. *Annual review of Economics, 8*, pp.405-434.

32. Thurow, L.C., 1978. Psychic income: useful or useless? *The American Economic Review, 68*(2), pp. 142-145.