An Economic Analysis of the Gender Gap in Top Professional Jobs

Paloma Ruiz De Dulanto

Prof: Marco Francesconi

Department of Economics, University of Essex

EC831: Project Economics

26/04/2019

Word count: 7969

Abstract

Based on cross-sectional data of 30,000 households in the United Kingdom for two time periods, 2007 and 2016/18, this study provides new empirical evidence on the extent of the gender wage gap, which declined considerably during this decade. An Ordinary Least Square regression reviews the factors that drive gender convergence. It concludes that conventional human capital variables such as education, number of own children, race or marital status explain little in both time periods of the gender gap. Despite this, the results suggest that the wage gap between men and women working full-time would not disappear even if these women received the same returns to their characteristics as men do.

CONTENTS

1. INTRODUCTION

While the gender wage gap has now been intensively investigated both in the United Kingdom and internationally, it remains an area of active and innovative research. One might wonder about the justification for yet another study on the subject, and the answer is that this paper focuses on the legacy that the financial crisis might have had on the gender wage gap after ten years of its occurrence. To test whether this analysis supports a causal relationship between

relative wage and sex of the individual, this study uses the most current cross-sectional data for the United Kingdom and compares these data with the gender pay gap in 2007. The report applies wage regressions for the years 2007 and 2016/18 to examine the difference in the pay gap over time. In the aggregate, the pay gap has moved from 20% to 15% of men´s wages over this decade.

According to the literature, the major advancements in the female labour market has led to the convergence of human capital investment and employment prospects throughout the past century. However, an unexplained gap remains even after controlling for male and female characteristics.

This report uses a variety of explanatory variables that account for women´s lower pay. In the statistical models it is explained that a large portion of the gender pay gap is due to differences in male and female employment. Adding controls to the models decreases the negative wage rate differential of women, to the point that number of hours normally worked in a week is less than 1% significant in both time periods, having children is not statistically significant and individuals with a white background only earn 10% more than those with non-white background. Employing an Oaxaca decomposition, the results show that removing worker heterogeneity does not minimize women´s tendency to earn less. The unexplained component remains a key factor in the current analysis.

The structure of the paper is as follows. Section 2 reviews the theoretical and empirical literature on the gender wage gap, both for United Kingdom and internationally. Section 3 presents the data used in the empirical analysis, a description of the variables and summary statistics. Section 4 presents the paper´s methodology, an Oaxaca-Blinder decomposition and assesses findings with reference to the current empirical literature. Section 5 presents the limitations of the study, and the last section sets out the conclusions and policy implications.

2. LITERATURE REVIEW

Female involvement in the labour market has been the focus of a vast and still growing strand of literature on its causes, characteristics and consequences. Existing research has disclosed supply-side explanations for gender drifts, including medical advances, human capital investment, technological changes, equal pay legislation and evolving social norms. Most of the

studies on gender wage gap are for the United States, but recently, more data for the labour market from both administrative sources and experiments has become available for high-income countries. This has drawn economists' attention towards international comparisons and other frameworks (see Bertrand 2011 for a comprehensive survey).

At the same time, previous studies focused on conventional economic variables (education, husband´s or nonlabour income, demographic variables…) for periods prior to the 1990s (i.e. periods of rapid increase in female participation rates) which cannot fully explain the observed increase in female labour-participation. Altonji and Blank (1999) investigate differentials by race and gender in the labour market. They discuss that the main sources of gender gaps in wages, hours worked, and job allocation lie in discrimination and human capital accumulation. Anticipated discrimination or discrimination may cause men and women with equal abilities to hold different positions. This suggests a significant role on shifts for preferences and other unmeasured variables. Cotter et al. (2011) provide little evidence on attitudes, although it might prove to be challenging to establish causation as people may adjust their attitude according to outcomes and labour-force behaviour. Flory et al. (2015) investigate workplace attitude when competing and find that women are less likely to both opt for performance pay and to thrive in competitive environments (Gneezy et al. 2003). Moreover, a recurrent finding from lab evidence is that women are more risk averse than men, (e.g. Croson and Gneezy 2009) and weaker negotiators (Rigdon 2012).

US women´s labour-force participation rates increased considerably in the years following World War II (from 31.8% in 1947 to 57.2% in 2013). This was mainly for married women who were more able to make joint decisions with their husbands (Blau and Kahn 2017). This increase inspires what Goldin (2006) has named the "quiet revolution" in gender roles (from late 1970s – right after the resurgence of feminism – to the present). It underlies the process in narrowing labour market differences. Before the revolution, women formed their adult identity after marriage, whilst after the revolution, women married late enough so that their adult identity was already formed. The increase in labour-force participation rates creates rising returns to skill for highly educated women compared to the less educated men (Blau et al 2014, figure 6-6).

Extensive explanations have been proposed for the increase in female labour supply. Medical advances have improved maternal health (reducing post birth disability), provided substitutes to maternal lactation (Albanesi & Olivetti 2016) and reduced fertility via the introduction of the

birth control pill (Bailey 2006). These advances allowed women to plan their careers in advance. The need of labour input in the household has been reduced by market substitutes such as household technological progress, and greater availability of child care *(Attanasio et al. 2008).*

Technological progress in the workplace has also raised female wages and participation rates by increasing the value of nonmanual skills relative to manual skills. Heathcote et al. (2010) interpret the female labour upward demand shift as a result of the growth in wages and hours worked for women, and in their model such a shift explains part of the rise in female hours in the US. The transition to the post-industrial phase is associated with a growth of the service sector. This expansion created jobs particularly suited for women´s skill set and preferences, rising female employment rates and consequently causing changes in social norms (Ngai and Petrongolo 2017). As a result, women had longer horizons, increased formal schooling investment, continued to graduate from professional schools and acquired doctorates in far greater numbers.

Others have observed changes in preferences and attitudes as driving forces to the increase in female participation in the labour market. Fernández et al. (2004) investigate men´s attitude changes. In their model, sons of working mothers prefer working wives, thus more women enter the labour market. Fogli et al. (2007) argue that each generation updates their parents´ beliefs by observing the children of employed women, therefore there are changes in women´s own sense of self. Consequently, the stronger presence of women in the labour-market made the gender stereotypes to weaken (Fernández et al. 2004). Olivetti and Petrongolo (2016) propose policies that would accelerate a change in social norms.

Olivetti and Petrongolo (2016) document gender convergence in main labour market indicators for a sample of 19 high-income countries for the period 1850-2008 using comparable data on employment, working hours and wages. Most countries in this sample share comparable trends for female outcomes even though the female participation rate differs widely across countries. Despite women´s progress in the labour-market, they conclude that there is no evidence of fully closing gender gaps. The reallocation of human capital from manufacturing into services is associated with an important variation in hours worked by women, as women choose “male” over “female” occupations. According to England (2010), “male” occupations tend to pay more on average than jobs which are female-dominated, so there are powerful economic incentives for women to join business-related fields.

Women´s role in childbearing acts as a disadvantage for high-rewarding careers, often known by long, inflexible working hours and strong nonlinear pay agendas relative to time worked (Goldin 2014a). Studies performed by Bertrand et al (2010) found unambiguous evidence of a negative relationship on relative earnings and having children, especially among the highly educated. Bertrand et al (2010) executed a web-based survey studying the careers of MBAs and career dynamics difference by gender for people who graduated between 1990 and 2006 from a top US business school. At the outset of their careers, female and male MBAs have nearly identical labour income, however, their earnings soon diverge. In their paper, they identify three main reasons that explain 84 percent of the 31-log point gender gap pooling across all years upon MBA graduation: "differences in business school courses and grades; differences in career interruptions; and differences in weekly hours" (Bertrand et al. 2010, pg 230). They sustain that the presence of children for female MBAs is the main reason for fewer job experience, shorter work hours, greater career discontinuity, and substantial earnings decline, unless they have lower-earning husbands. The competitive nature of corporate and financial companies and the up-or-out nature of major firms and academic institutions contributes to the cost of career interruptions.

Blau and Kahn (2017) use Panel Study of Income Dynamics microdata over the 1980-2010 period, providing new empirical evidence on the extent of and trends in gender gap and documented the improvements in women's education, experience and occupational representation. Blau and Kahn conclude that women´s work force interruptions and shorter hours remain significant in high-skilled occupations. They discuss that by 2010 the pay gap was most noticeable at the top of the wage distribution, even after accounting for measured characteristics. They showed that women played a key role in the reduction of the pay gap. Their main argument is that gender differences lie in occupations, roles, industries and preferences, mentioning as well that discrimination cannot be discounted. Noncognitive skills and physiological attributes comprise a very small portion of the gender pay gap. They also highlight the diminished role of human-capital factors accounting for gender differences over time – due to the narrowing of the experience gap and to the reversal of the education gap between women and their counterparts. Moreover, by 2010 conventional human capital variables (labour-market experience and education) explain little of the aggregate gender wage differentials. Furthermore, the union status role has virtually disappeared in accounting for gender differences. An unexplained gender wage gap remains.

In conclusion, human capital investment and employment prospects have converged due to the major advancement in the female labour-market throughout the past century. However, there is still an unexplained wage gap that remains even when accounting for measured characteristics. It is noticeable that the unexplained gap declined more slowly at the top of the wage distribution, where the labour-market favoured highly-skilled men over women. So far, quantitative evidence has shown that occupation and industry preferences continue to be an important part of explaining gender gap. Similarly, even though education and experience continue to have salience, when taken together they explain little of the gender gap in the aggregate, as we have seen that women have weakened the gender experience and education gap. Recent research highlights the importance of work force interruptions and fewer hours worked in high-skilled careers.

3. DATA

3.1. The Study Population

This research draws on data from wave B17 (year 2007) of Understanding Society, the British Household Panel Survey (BHPS) and wave 8 (years 2016-2018) of the UK Household Longitudinal Study (UKHLS). The BHPS is an annual longitudinal survey of household and individuals living in the UK. It started in 1991 with 5,000 households selected at random within Great Britain. The UKHLS draws heavily on the success of its predecessor, BHPS. The UKHLS is a multi-focus multi-topic longitudinal household panel survey that started in 2009 with a nationally representative, stratified, clustered sample of 28,000 UK households and includes an ethnic minority boost sample of around 4,000 households. Field-work takes place over a period of 24 months, with a random sample of households selected for interview each month. Detailed information about Understanding Society is available on the study website http://www.understandingsociety.ac.uk/ and in its quality profile (Lynn and Knies 2015).

The reason why wave B17 from BHPS (2007) and wave 8 from UKHLS (2016/18) were chosen for this analysis is because the former occurred right before the financial crisis started and the latter, ten years later. This comparison should allow us to see the legacy of the financial crisis on gender differentials after a decade of its appearance. The sample size for year 2007 is of 5,201 whilst for years 2016/18 the number of observations is 12,272.

The BHPS and UKHLS are particularly suited for our analysis. It oversamples members of the UK and retrieves information from over 30,000 households which were selected from more than

2,640 primary sampling units and stratified by region. Combined with the possibility of linking the study members over the years, this allows us to investigate with greater statistical power how the gender gap co-varies over different time periods.

3.2. Variable Description

The outcome variable, log hourly wage, is the natural logarithm of average gross hourly earnings. It is derived from gross monthly pay at last payment and total weekly hours (including paid overtime). As we do not want the measure of earnings growth to include aggregate wage growth, gross monthly income is deflated using the yearly consumer price indices provided by the Office for National Statistics taking as base year 2015. We exclude from the analysis individuals with zero or negative income. The information is collected from adults in the household and accounts for item non-response, see Knies (2015).

For this analysis, workers in top professional jobs are defined as adults (age 16+) with a permanent paid employment, full-time workers and working more than 20 hours a week, who lived in England at the time of the interview. Mothers on maternity leave were excluded, as well as full-time students, retired people, sick or disabled and others whose economic activity was not paid employment. The self-employed were also left out due to the difficulty in measuring their earnings. In total, about 68.29% of individuals are excluded due to these restrictions or data problems.

We include the following individual characteristics, which are expected both to differ across groups and be associated with hourly wage: age, sex, family context (i.e., marital status, number of own children in the household), employment status, occupational classification of current job (current job refers to being in paid employment during the week before question was ask), education (i.e., highest education), job satisfaction, ethnicity group (i.e., white, non-white) and health (i.e., whether the respondent presents a longstanding illness or disability). These are standard measures for wage gap also included in studies of Stanley and Jarrell (1998) and Weichselbaumer and Winter-Ebmer (2005).

The variable white includes British, English, Scottish, Welsh, Northern Irish, Irish, Gypsy, Irish Traveller or any other white background. Non-whites are defined as having a mixed, Indian, Asian or black background or any other ethnic group not mentioned previously. For marital status, three different groups are created. Firstly, single, never married or never had a civil partner. Secondly, married, legal civil partner or separated legally married. Lastly, divorced,

widowed or separated from civil partner. Every other refusal or missing data has been excluded. Job satisfaction is collected in the self-completion questionnaire, where respondents are asked to report how satisfied they are with their job overall on a fully labelled scale running from 1 “completely dissatisfied” to 7 “completely satisfied”. The inclusion of the polynomials age squared, age cubic and age quartic allows us to model more accurately the effect of age, which may have a non-linear relationship with the dependent variable. Age is derived from the sample individual´s age at time of the interview. The *highest qualification* variable is a categorical variable composed of six different levels: degree, other higher degree, A-level or equivalent, GCSE or equivalent, other qualification or no qualification. The individuals were asked the highest educational or school qualification they have obtained.

The Government Office Region variable is a code frame variable expanded to cover Wales, Scotland and Northern Ireland as applicable. It is derived from the individual’s postcode. Occupational dummies are created from the Standard Occupational Classification 2000 (SOC2000) of current job. For this paper, occupations have been grouped into 9 different categories (taking the first digit of the 4-digit version for Special Licence). The groups are: operatives, sales related occupations, elementary occupations, service occupations, traders, administrative occupations, associate professionals, professionals and managers. These subgroups are unique to this study, therefore contributing to the literature as previous research only takes into consideration one occupational group and not all of them (i.e., Bertrand et al, 2010 only looks at MBA graduates).

The variable long-standing illness or disability stands for any long-standing physical or mental impairment, illness or disability. In other words, anything that could have troubled the individual over a period of at least 12 months. This dummy variable could only be found for the second wave, years 2016/18. In the first wave, year 2007, this question was asked more specifically and divided into subcategories, which made it impossible to create a dummy that would capture the same meaning as for the other wave. For exact question wording consult the study questionnaires which are provided on the study homepage, www.understandingsociety.ac.uk.

3.3. Descriptive Statistics

Table 1 and 2 present summary statistics for the variables used in our analysis for each time period.

Table 1. Summary Statistics for both periods

	2007		2016/18	
	mean	sd	mean	sd
Log of Hourly Wage	2.385647	.5047859	2.589945	.5131092
Long-standing illness or disability	-	-	.2247392	.4174275
Married	.5298981	.4991533	.5682855	.4953353
Single, never married	.3895405	.487693	.3319752	.4709413
Divorced, widowed or separated	.0805614	.2721866	.0972947	.2963706
Operatives	.0870986	.2820067	.0714635	.257608
Sales Related Occupations	.0517208	.2214841	.0513364	.2206919
Elementary Occupations	.0851759	.2791702	.0753748	.264006
Service Occupations	.0680638	.2518796	.0910202	.2876496
Traders	.1188233	.3236116	.0664113	.2490099
Administrative Occupations	.1274755	.3335369	.1071545	.3093222
Associate Professionals	.1680446	.3739418	.1892927	.3917569
Professionals	.1538166	.3608075	.1954042	.3965277
Managers	.1703519	.3759778	.1872555	.3901324
Female	.4270333	.4946948	.4611247	.4985067
White	.9596231	.1968604	.8133963	.3896089
Age derived	39.33032	12.05186	42.29963	11.92095
Age squared	1692.094	971.3837	1931.356	1004.282
Age cubed	78112.21	63707.27	93534.81	68344.13
Age quartic	3806586	3966997	4738635	4399610
Highest qualifications	3.267614	1.923747	2.503844	1.600111
Government Office Region	8.031534	3.414702	6.532002	3.144328
Number of own children in household	.5423957	.8826209	.5867014	.9215743
Job satisfaction: overall	5.378197	1.191725	5.284585	1.360309
No. of hours normally worked per week	38.33782	6.179393	37.92953	5.748555

After the restrictions imposed (full-time workers, with permanent employment, etc) and processing the data by eliminating outliers and negative values, only 40.40% of the individuals for year 2007 meet the requirements for this study, and 31.24% for years 2016/18.

For the year 2007, the mean of the variable “number of own children in the household” is 0.54. As this sample is only for full-time employees, the low number of children can be due to individuals choosing to have less children and focusing on their career, allowing less time to raise children. For years 2016/18, 65.45% of the population have no children, 2% lower than in the previous decade. This could be explained by the increasing price of having children (childcare is too expensive) or to the financial instability present over the last 10 years (since the financial crisis started).

The percentage of those employed who are non-white increases from 4.04% to 18.66% across waves. The increase could be attributed to the Equality Act 2010 which legally protects people

from discrimination in the workplace and in wider society (GOV.UK, 2019). The Affirmative Action programs and quotas were also designed to end discrimination in the workplace because of creed, colour or national origin. This as well as the increased proportion of non-white college graduates contributed to the higher number of non-white workers in the UK.

4. ANALYSIS AND RESULTS

4.1. Empirical Specification

The main question addressed in this study is whether empirical data supports a casual relationship between relative wage and sex of the individual. Most authors have adopted the human capital model as the theoretical basis for the earnings function (Becker, 1962; Mincer, 1958). Similarly, this study also adopts the human capital model. To examine the association between hourly wage and several explanatory variables representing the individual characteristics, the following model for both time periods (wave B17 and wave 8) is used:

$$\ln(y_{it}) = \beta_0 + \beta_1 female_{it} + \beta_2 hiqual_{it} + \beta_3 region_{it} + \beta_4 jbsat_{it} + \beta_j X_{it} + \varepsilon_{it} \qquad (1)$$

with $i = \overline{1, 12272}$ denoting individuals and $t \in \{2007, 2016/18\}$ denoting time. The outcome variable y_{it} denotes log hourly wage for individual *i* at time *t.* It depends upon a female dummy variable, a *highest qualification*, *region* and *job satisfaction* vector variables and nineteen independent explanatory variables (i.e., marital status, job characteristics, age…) and corresponding coefficient parameters β_{5-19}, a constant term β_o, and an unknown component comprised of a random error term ε_{it}. The model assumptions are that ε_{it} is normally distributed with a mean zero and a standard deviation σ_W and uncorrelated with X_{it}.

This paper adopts the central multiple linear regression (MLR) model which is estimated using ordinary least squares (OLS) during our analysis. The specification of this model allows us to investigate the size, direction and significance of variables in determining the gender gap in top professional jobs within our sample. Recurrent variables in the literature have been included, alongside the model-specific variables mentioned in section 3.2; supporting both the unique contribution of this paper to the literature and boosting the explanatory power of this model.

Before starting to estimate the model outlined in this section, preliminary work was undertaken to ensure that the model is properly specified. OLS was chosen to estimate this model as it is the best linear unbiased estimator (BLUE) among the class of all linear estimator, so long as the

Gauss-Markov assumptions are met (Wooldridge, 2012). Conditional on Gauss-Markov assumptions holding (MLR.1-5), the model will provide unbiased ($E(\hat{\beta}) = \beta$) and efficient estimates of the parameters with the lowest variance. Combined with the assumption of normality (MLR.6), the model will now produce accurate confidence intervals and conduct hypothesis tests on the estimated parameters.

On one hand, to check whether the variance of the errors from the regressions are dependent on the values of the independent variables, a Breusch-Pagan test is needed. The results are consistent in rejecting the null hypothesis of homoskedasticity, implying that heteroskedasticity is present in both our models. As a result, the regressions are run with heteroskedastic robust errors. On the other hand, to check whether there are any neglected nonlinearities in the model, a Ramsey test is conducted. The results are consistent in rejecting the null hypothesis of correct specification, therefore we take age squared, age cubic and age quartic as new independent variables and the log-level model of this regression.

Table 2. Test Statistics	2007		2016/18	
Heteroscedasticity[1]	$chi2(1) = 44.78$	Prob > chi = 0.0000	$chi2(1) = 163.93$	Prob > chi = 0.0000
Omitted Variable Bias[2]	$F(3, 5059) = 11.44$	Prob > F = 0.0000	$F(3, 12025) = 57.67$	Prob > F = 0.0000

1. Breusch-Pagan test H_0: Constant Variance Variables: fitted values of ln hourly wage
2. Ramsey RESET test using powers of the fitted values of ln hourly wage H_0: model has no omitted variables

4.2. Empirical Findings

4.2.1. BASE LINE RESULTS

Table 3 and 4 present the estimation results, incorporating the statistical adequacy amendments outlined in the empirical specification section. Each model takes Log Hourly Wages as the dependent variable. Table 3 includes the regressions for the first year – 2007 (wave B17) and estimates three models: the first model is a pool regression, including both female and male individuals, (2) is concerned only with males and (3) displays results for females. Table 4 includes the regressions for years 2016/18 (wave 7) in the same format as Table 3.

Table 3. Regressions for year 2007: population characteristics of individuals by sex

	(1) Log Hourly Wage - Pool		(2) Log Hourly Wage - Men		(3) Log Hourly Wage - Women	
Married	0.0376**	(2.81)	0.0877***	(4.87)	-0.0442*	(-2.29)
Divorced, widowed or separated	-0.0411	(-1.95)	-0.00278	(-0.09)	-0.0985***	(-3.48)
Sales-Related Occup	0	(.)	0	(.)	0	(.)
Elementary Occupations	-0.214***	(-8.58)	-0.252***	(-8.00)	-0.186***	(-3.93)
Operatives	-0.0164	(-0.65)	-0.0705*	(-2.25)	-0.0284	(-0.54)
Service Occupations	-0.174***	(-6.55)	-0.320***	(-6.35)	-0.0857**	(-2.62)
Traders	-0.0356	(-1.48)	-0.0617*	(-2.06)	-0.278***	(-5.73)
Administrative Occup	-0.0362	(-1.65)	-0.111**	(-3.04)	0.0473	(1.64)
Associate Professionals	0.135***	(6.26)	0.0420	(1.38)	0.247***	(7.99)
Professionals	0.278***	(13.25)	0.193***	(7.15)	0.398***	(11.96)
Managers	0.210***	(10.18)	0.181***	(6.64)	0.250***	(7.85)
White	0.0912***	(3.30)	0.129***	(3.50)	0.0107	(0.27)
female	-0.199***	(-16.15)				
Age	0.279***	(5.36)	0.332***	(5.47)	0.197**	(2.70)
Age Squared	-0.00886***	(-4.21)	-0.0109***	(-4.48)	-0.00546	(-1.87)
Age Cubed	0.000130***	(3.59)	0.000162***	(3.95)	0.0000700	(1.40)
Age Quartic	-0.000000730**	(-3.28)	-0.000000931***	(-3.71)	-0.000000354	(-1.15)
No. of hours normally worked per week	-0.00782***	(-6.66)	-0.00855***	(-6.21)	-0.0101***	(-4.50)
No. of own children in household	0.0107	(1.48)	0.0175	(1.92)	-0.0181	(-1.47)
Highest qualification						
Degree	0	(.)	0	(.)	0	(.)
Other higher degree	-0.102***	(-4.95)	-0.111***	(-3.90)	-0.0838**	(-2.94)
A-level etc	-0.225***	(-13.29)	-0.206***	(-8.74)	-0.230***	(-9.63)
GSCE etc	-0.288***	(-16.84)	-0.265***	(-11.21)	-0.286***	(-11.68)
Other qualification	-0.331***	(-15.00)	-0.288***	(-9.73)	-0.378***	(-11.73)
No qualification	-0.456***	(-17.20)	-0.416***	(-12.91)	-0.466***	(-9.57)
Government Office Region						
North East	0	(.)	0	(.)	0	(.)
North West	0.0390	(1.17)	0.0613	(1.32)	0.00560	(0.13)
Yorkshire & Humber	0.0559	(1.62)	0.119*	(2.55)	-0.0457	(-0.93)
East Midlands	0.0340	(0.99)	0.0832	(1.75)	-0.0416	(-0.87)
West Midlands	0.0210	(0.58)	0.0540	(1.08)	-0.0267	(-0.56)
East of England	0.103**	(2.88)	0.193***	(3.95)	-0.0126	(-0.27)
London	0.242***	(6.27)	0.292***	(5.35)	0.178***	(3.56)
South East	0.127***	(3.95)	0.180***	(3.99)	0.0501	(1.19)
South West	0.0557	(1.68)	0.129**	(2.85)	-0.0401	(-0.87)
Wales	-0.0223	(-0.75)	0.0138	(0.33)	-0.0699	(-1.81)
Scotland	0.0385	(1.28)	0.0780	(1.84)	-0.00686	(-0.18)
Northern Ireland	0.0115	(0.38)	0.0452	(1.05)	-0.0306	(-0.79)
Job Satisfaction						
Not Satis At All	0	(.)	0	(.)	0	(.)
Mostly dissatisfied	0.0742	(1.37)	0.0767	(1.09)	0.0569	(0.69)

Somewhat Dissatisfied	0.117*	(2.23)	0.0962	(1.42)	0.124	(1.52)
Not satis/dissatis	0.104*	(2.06)	0.129*	(1.99)	0.0588	(0.73)
Somewhat Satisfied	0.159**	(3.28)	0.165**	(2.63)	0.138	(1.82)
Mostly Satisfied	0.185***	(3.86)	0.205***	(3.30)	0.154*	(2.05)
Completely Satis	0.173***	(3.39)	0.221**	(3.27)	0.108	(1.38)
Constant	-0.727	(-1.56)	-1.305*	(-2.35)	-0.0106	(-0.02)
Observations	5103		2930		2173	
R-squared	0.5077		0.4871		0.5435	

Robust *t* statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4. Regression for years 2016/18: Population characteristics of individuals by sex

	(1) Log Hourly Wage - Pool		(2) Log Hourly Wage - Men		(3) Log Hourly Wage - Women	
Long-standing illness or disability	-0.0433***	(-5.05)	-0.0487***	(-3.76)	-0.0361**	(-3.26)
Married	0.0628***	(6.79)	0.103***	(7.77)	0.0232	(1.82)
Divorced, widowed or separated	0.00505	(0.35)	0.0645**	(2.72)	-0.0398*	(-2.27)
Sales-Related Occup	0	(.)	0	(.)	0	(.)
Elementary Occupations	-0.196***	(-11.47)	-0.233***	(-10.32)	-0.162***	(-5.91)
Operatives	-0.0440*	(-2.39)	-0.0916***	(-4.01)	-0.0385	(-0.96)
Service Occupations	-0.188***	(-11.89)	-0.285***	(-8.75)	-0.109***	(-5.62)
Traders	0.0428*	(2.22)	0.0110	(0.48)	-0.0948*	(-2.17)
Administrative Occup	-0.0273	(-1.81)	-0.107***	(-4.57)	0.0591**	(3.03)
Associate Professionals	0.152***	(10.45)	0.0742***	(3.49)	0.246***	(12.59)
Professionals	0.265***	(19.50)	0.195***	(10.84)	0.361***	(17.80)
Managers	0.205***	(15.34)	0.167***	(9.23)	0.257***	(13.03)
White	0.116***	(11.26)	0.172***	(11.44)	0.0621***	(4.50)
Female	-0.154***	(-18.87)				
Age	0.200***	(5.48)	0.248***	(5.08)	0.0907	(1.62)
Age squared	-0.00583***	(-4.30)	-0.00740***	(-4.09)	-0.00177	(-0.85)
Age cubed	0.0000815***	(3.77)	0.000104***	(3.64)	0.0000163	(0.49)
Age quartic	-0.000000453***	(-3.61)	-0.000000577***	(-3.50)	-7.16e-08	(-0.37)
No. of hours normally worked per week	-0.00843***	(-9.92)	-0.0106***	(-9.06)	-0.00758***	(-6.07)
No. of own children in household	-0.000386	(-0.08)	0.00532	(0.86)	-0.0145*	(-2.00)
Highest Qualification						
Degree	0	(.)	0	(.)	0	(.)
Other higher degree	-0.144***	(-12.63)	-0.135***	(-7.74)	-0.148***	(-10.02)

A-level etc	-0.204***	(-19.29)	-0.195***	(-13.08)	-0.211***	(-14.21)
GCSE etc	-0.295***	(-25.08)	-0.305***	(-18.17)	-0.270***	(-16.55)
Other qualification	-0.328***	(-19.86)	-0.327***	(-14.77)	-0.317***	(-12.84)
No qualification	-0.433***	(-16.55)	-0.428***	(-11.85)	-0.414***	(-11.92)
Government Office Region						
North East	0	(.)	0	(.)	0	(.)
North West	0.0313	(1.62)	0.0621*	(2.27)	0.00272	(0.10)
Yorkshire & Humber	0.0308	(1.55)	0.0518	(1.82)	0.0149	(0.55)
East Midlands	0.0267	(1.31)	0.0415	(1.41)	0.0230	(0.84)
West Midlands	0.0853***	(4.20)	0.107***	(3.64)	0.0611*	(2.22)
East of England	0.124***	(5.93)	0.143***	(4.74)	0.101***	(3.61)
London	0.246***	(12.28)	0.255***	(8.83)	0.237***	(8.68)
South East	0.134***	(6.82)	0.153***	(5.42)	0.111***	(4.13)
South West	0.0577**	(2.75)	0.0548	(1.83)	0.0645*	(2.23)
Wales	0.0412*	(2.00)	0.0318	(1.08)	0.0572*	(2.03)
Scotland	0.0982***	(4.94)	0.104***	(3.61)	0.0960***	(3.59)
Northern Ireland	0.0149	(0.72)	-0.00916	(-0.31)	0.0469	(1.65)
Job Satisfaction						
Completely Dissatisfied	0	(.)	0	(.)	0	(.)
mostly dissatisfied	0.0714*	(2.10)	0.0585	(1.17)	0.0784	(1.88)
Somewhat dissatisfied	0.0961**	(3.08)	0.0831	(1.76)	0.0972**	(2.66)
Not satis/dissatis	0.0954**	(3.10)	0.0883	(1.89)	0.0932**	(2.59)
somewhat satisfied	0.0911**	(3.08)	0.0843	(1.86)	0.0845*	(2.51)
mostly satisfied	0.126***	(4.30)	0.142**	(3.15)	0.0962**	(2.87)
completely satisfied	0.0878**	(2.88)	0.114*	(2.42)	0.0586	(1.69)
Constant	0.0625	(0.18)	-0.424	(-0.89)	0.966	(1.77)
Observations	12070		6498		5572	
R-squared	0.4383		0.4222		0.4614	

Robust *t* statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

4.2.2. OAXACA DECOMPOSITION

A commonly used methodology to investigate differences in wages by groups (i.e., sex, race, so on) is to decompose mean differences based on linear regression models in a counterfactual manner. The procedure is developed by Blinder (1973) and Oaxaca (1973) and allows productive characteristics of men and women to be rewarded differently. Hourly wages are estimated separately for individuals i of different groups g:

$$y_{gi} = \beta_g X_{gi} + \varepsilon_{gi} \tag{2}$$

where $g = (m, f)$, representing male and female; y_{gi} is the log wage, and X_{gi} the control characteristics of an individual i of group g.

The wage differential between male and female can then be decomposed into a part explained by differences in characteristics and an unknown residual. The difference in mean wages can be expressed as:

$$\bar{y}_m - \bar{y}_f = \left(\bar{X}_m - \bar{X}_f\right)\hat{\beta}_m + \left(\hat{\beta}_m - \hat{\beta}_f\right)\bar{X}_f \equiv E + U \tag{3}$$

where $\bar{y}_g$ and $\bar{X}_g$ denote the mean log hourly wages and control characteristics of group g and $\hat{\beta}_g$ represents the estimated parameter from equation (2). The endowment effect E stands for the effect of different productive characteristics, and the unknown residual U is due to differences in the estimated coefficients and is often used as a measure for discrimination. U indicates how much higher the average female wage would be if women´s endowments would be remunerated similarly to men´s (Weichselbaumer and Winter-Ebmer, 2005) but it also subsumes the effects of group differences in unobserved predictors (Jann, 2008).

The results from Table 5 show the estimates of two group-specific regression models and performs a threefold decomposition. For these regressions, only variables *highest qualification, marital status, race, number of own children in the household* and *job occupation* are defined. It reports the mean of log wages, yielding a wage gap of 18 percentage points for 2007 and 14 percentage points for 2016/18.

Table 5. Oaxaca Decomposition for *highest qualification*, *marital status*, *race, number of own children in the household* and *occupation* for both time periods.

	(1) Log Hourly Wage 2007	(2) Log Hourly Wage 2016/18
Differential		
Prediction_1	2.473***	2.661***
	(264.92)	(401.18)
Prediction_2	2.286***	2.520***
	(218.29)	(390.61)
Difference	0.187***	0.142***
	(13.35)	(15.30)
Decomposition		
Endowments	-0.0477***	-0.0487***
	(-4.87)	(-8.25)
Coefficients	0.209***	0.182***
	(18.15)	(22.64)
Interaction	0.0257***	0.00856**
	(5.49)	(2.75)
N	5068	11912

t statistics in parentheses
* p < 0.05, ** p < 0.01, *** p < 0.001

4.2.3. COMMENTS

All analysis is conducted under a ceteris paribus assumption. Using a t-test, there is statistically significant evidence against the null hypothesis, which is the coefficient being indifferent from zero at a given p-value.

The largest single cause of the gender pay gap is simply gender. As expected, wages were 20% lower for women in 2007 and 15% lower in 2016/18 (being statistically significant at the 1% level), even after controlling for explanatory variables. The extent of this pay differential is remarkable given the size of R-squared for each model (0.51 for 2007 and 0.44 for 2016/18), and the number of controls added to the models (see Table 1). The residual presented in the equation is the percentage of the wage level that is explained by the variable *female*. Discussion of the residual is often omitted in the literature, however here we present a decomposition to offer a plausible explanation of the unexplained gender effect.

Neoclassical economists propose that the residual is a gender-patterned productivity effect. They suggest that women seek employment which allows them to balance both the family life and the paid work, even if it means accepting lower wages. In contrast, Olsen et al (2009) account for three components: women´s work is valued less than men´s as employers believe their work is inferior, women´s employers may behave according to a ´breadwinner´ ideology (i.e., man is deemed the principal earner in the household) and finally gender stereotypes. Nevertheless, the residual remains a topic for discussion rather than one which empirical findings can help us interpret.

The next most important driver of the pay gap is occupational segregation. Swaffield (2008) arrives to the conclusion that occupational differences cannot explain much of the gender gap in early-career wage growth. His conclusion can be reconciled with the findings of Machin and Puhani (2003). They both claim that there is a small gender wage gap on entry to the labour market among graduates. In this study, the highest levels of male segregation are in technical occupations and primary industry (65% are men on average), while highest levels for female segregation are in customer services and caring work (77% are women). These results are parallel to the ones presented by Swaffield (2008). Other factors, such as women being more risk averse than men, provide insight into the sources of the unexplained gap. However, this factor may be considered by measured variables: if it results in women avoiding occupations with greater variance in earnings, the variable occupation will adjust for this factor (Blau and Kahn, 2017).

Human capital accumulation before labour market entry, in terms of the quantity of education, explains few gender differences. The extent of education may be different as the subject of degree can explain a portion of the gender gap among college graduates (Machin and Puhani 2003). Tables 3 and 4 include the categorical variable *highest education*. Our sample shows that nearly the same proportion of men and women have a degree (22% for 2007 and 28% for 2016/18). In the two time periods, not holding a qualification impacts negatively both men and women, decreasing wage by approximately 45% compared to having a degree. This is supported by Brown and Corcoran (1997) who found that gender differences in curriculum are larger for graduates than non-graduates for the US. According to Blau and Kahn (2017), in the case of education, there was a dramatic reversal of the gender gap. In 1981, women had lower average levels of schooling than men, but by 2011 women had higher average levels of schooling. In this study there is an increase on the years of schooling for women (23% of

women have a degree in 2007 compared to 41% in 2016/18), although men have higher or equal levels of schooling than women.

In the pay gap equation, we have allowed for age to be a proxy for experience in the labour market. However, age is still associated with pay differences. Increasing age by one unit (years), will increase the hourly wage by 25% for men and 9% for women in 2016/18, which is a significant change. The coefficient of age squared is statistically significant and indicates that the relationship between wage and age is not linear, describing a U-shaped relationship with one inflection point. According to Swaffield (2008), the gender pay gap is lower at all ages for younger cohorts, observing that all the observations of the gender gap in wage growth up to the age of 35 are positive.

Those regions in or near London have higher average wages, 24% higher compared to North East region for both time periods. Scotland, Wales and most of the northern regions of England have small wage differentials, which are not statistically significant. Observing the regional effects, it can be concluded that the BHPS and UKHLS data do not tend to support the idea that regional location contributes to the size of the pay gap. This is in line with Olsen et al (2010) conclusion.

Controlling for marital status and number of own children in the household can be correlated to women´s labour-force decisions. Bertrand et al (2010) point out evidence of an inverse correlation between relative earnings and having children, especially among the highly educated. They maintain that the presence of children is the main reason for shorter work hours and therefore substantial earnings decline. In this study, there is evidence for a negative relationship for women in both time periods, although for 2007 it is not statistically significant and in 2016/18 is only significant at the 5% significance value. This is due to the restricted sample used in the analysis, as these women work full-time, more than 20 hours a week and have a permanent employment. About 71% of women in both time periods don't have any children, which explains why the presence of own children in the household explain so little of the gender gap. Almost 50% of female in our sample are married in both time periods, yielding a decrease of the 4% in log wages for 2007 and an increase of 2% for 2016/18.

On one hand, this study uses a traditional Oaxaca-Blinder decomposition of male/female differences in log wages, explained by *highest qualification, marital status, race, number of own children in the household* and *occupation* for both time periods. The decomposition yields a wage gap of 0.19 for 2007 and 0.14 for 2016/18. The decrease of approximately 0.05 in both

time periods indicates that differences in the explanatory variables mentioned above account for 27% of the wage gap (i.e., if women had the same characteristics as men). On the other hand, 0.21 for 2007 and 0.19 for 2016/18 is the change in women´s wages when applying the men´s characteristics to the women´s characteristics. The interaction term measures the simultaneous effect of differences in endowments and coefficients (Jann, 2008). This means that even if women had the characteristics or the coefficients of men, there is still an unexplained component which many economists call discrimination.

5. LIMITATIONS OF THE STUDY

The interpretation of the estimated "female" coefficient is problematic due to omitted variable bias. Results of a regression analysis might suggest discrimination by gender or race even if the employer didn't use group membership in her decision of how much to pay an employee. It could be that race and gender are correlated with other proxies for productivity that are not observed in this study but might have been observed by the employer.

It has been discussed that less experience, greater time in housework and less investment in on-the-job training of women may be voluntary choices made by women. This might be responsible for the unexplained residual as it is not adequately captured in the data and such variables have not been included in the study due to lack of information. One peculiarity of the BHPS and UKHLS is that its basic measure of job tenure is time in a job with the same employer. This means that within-firm changes of job reset the clock to zero. For this reason, tenure wasn't included in our analysis. Another key explanatory variable that is missing from this study is job experience, as it is not collected in the BHPS or UKHLS. For this reason, age is used as a proxy for experience in the labour market. This study also lacks variables representing attitudes to risk or skills for negotiating, which previous literature employ in explaining the pay gap.

This project could be extended or improved by adding variables such as size of the firm, a dummy for working in the public sector, partners earnings (whether they are higher or lower), work force interruptions due to other than health, such as childcare or home production, etc. However, adding more and more controls to the regression could obscure the interpretation of the evidence. Therefore, it can lead to biased estimators of the coefficients.

6. CONCLUSION

This report examines gender differences using the most recently available UKHLS data for the United Kingdom and compares it with the gender pay gap 10 years before, using BHPS data. The study used cross-sectional data to provide empirical evidence of a causal relationship between relative wage and sex of the individual. The findings are consistent with the existing literature, though this study outlines the importance of a correct interpretation of the unexplained female residual.

The author concludes that many of the traditional explanations continue to have salience for understanding the gender differentials, although some factors have decreased and other have increased in importance. Job design was found to lower women´s earnings as those women who work in male-dominated jobs, especially in construction, manufacturing and trade, earn lower wages than men. In summary, women who work full-time and have a permanent employment in the trade business can expect to be 11% behind equivalent men.

The gender wage gap is found to be largely related to women receiving lower rewards for their characteristics than men. A proportion of employees and employers may hold traditional gender ideologies that are not consistent with equal earnings. Indeed, the results suggest the gender wage differentials between men and women working full-time in United Kingdom would not disappear even if these women received the same returns on their characteristics as men do. This means that removing worker heterogeneity does not remove women´s tendency to earn lower wages.

Future policies should address parental leave provisions and family subsidies (i.e., childcare) and should emphasise their impact in raising female labour force participation in the United Kingdom. Blau and Kahn (2006) point to changes in public policies, specifically welfare and the Earned Income Tax Credit (EITC). Although generous policies are associated with higher female participation rates, such policies may backfire by reinforcing employers´ beliefs and social norms regarding women´s advantage in childcare, home production thus causing job segregation (Francois 1998) or raising the relative cost of hiring women (Blau and Kahn, 2013).

7. CITATION

- Albanesi, S. and Olivetti, C. (2016). Gender Roles and Medical Progress. *Journal of Political Economy*, 124(3), pp.650-695.
- Altonji, Joseph G. and Blank, Rebecca M. (1999). Race and Gender in the Labor Market. *Handbook of labor economics*, Vol. 30. Amsterdam: North-Holland, 1999, pp. 314
- Attanasio, O., Low, H. and Sánchez-Marcos, V. (2008). Explaining Changes in Female Labor Supply in a Life-Cycle Model. *American Economic Review*, 98(4), pp.1517-1552.
- Bailey MJ. (2006). More power to the pill: the impact of contraceptive freedom on women's lifecycle labor supply. *Q. J. Econ*. 121:289–320
- Becker, G. (1962). Investment in Human Capital: A Theoretical Analysis. *Journal of Political Economy*, 70(5, Part 2), pp.9-49.
- Bertrand, M., Goldin, C. and Katz, L. (2010). Dynamics of the Gender Gap for Young Professionals in the Financial and Corporate Sectors. *American Economic Journal: Applied Economics*, 2(3), pp.228-255.
- *Blau FD, Kahn LM. 2013. Female labor supply: Why is the United States falling behind? Am. Econ. Rev. 103(3):251–56*
- Blau, F. and Kahn, L. (2007). Changes in the Labor Supply Behavior of Married Women: 1980–2000. *Journal of Labor Economics*, 25(3), pp.393-438.
- Blau, F. and Kahn, L. (2017). The Gender Wage Gap: Extent, Trends, and Explanations. *Journal of Economic Literature*, 55(3), pp.789-865.
- Brown, C. and Corcoran, M. (1997). Sex-Based Differences in School Content and the Male-Female Wage Gap. *Journal of Labor Economics*, 15(3), pp.431-465.
- Cotter, D., Hermsen, J. and Vanneman, R. (2011). The End of the Gender Revolution? Gender Role Attitudes from 1977 to 2008. *American Journal of Sociology*, 117(1), pp.259-289.
- Croson, R. and Gneezy, U. (2009). Gender Differences in Preferences. *Journal of Economic Literature*, 47(2), pp.448-474.
- England, P. (2010). The Gender Revolution. *Gender & Society*, 24(2), pp.149-166.
- Fernandez, R., Fogli, A. and Olivetti, C. (2004). Mothers and Sons: Preference Formation and Female Labor Force Dynamics. *The Quarterly Journal of Economics*, 119(4), pp.1249-1299.
- Flory, J., Leibbrandt, A. and List, J. (2014). Do Competitive Workplaces Deter Female Workers? A Large-Scale Natural Field Experiment on Job Entry Decisions. *The Review of Economic Studies*, 82(1), pp.122-155.
- Fogli, A. and Veldkamp, L. (2007). Nature or Nurture? Learning and Female Labor Force Dynamics. *SSRN Electronic Journal*.
- *Francois P. 1998. Gender discrimination without gender difference: theory and policy responses. J. Public Econ. 68:1–32*
- Gneezy, U., Niederle, M. and Rustichini, A. (2003). Performance in Competitive Environments: Gender Differences. *The Quarterly Journal of Economics*, 118(3), pp.1049-1074.

- Goldin, C., Katz, L. 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.
- GOV.UK. (2019). *Equality Act 2010: guidance*. [online] Available at: https://www.gov.uk/guidance/equality-act-2010-guidance [Accessed 18 Apr. 2019].
- Heathcote, J., Storesletten, K. and Violante, G. (2010). The Macroeconomic Implications of Rising Wage Inequality in the United States. *Journal of Political Economy*, 118(4), pp.681-722.
- Jann, B. (2008). The Blinder–Oaxaca Decomposition for Linear Regression Models. *The Stata Journal: Promoting communications on statistics and Stata*, 8(4), pp.453-479.
- Knies, Gundi, and Seetha Menon. 2014. "Understanding Society: Waves 1-3, 2009-2012: Special Licence Access, Geographical Accessibility. The UKHLS-Accessibility Data File User Guide 21 " http://doc.ukdataservice.ac.uk/doc/7533/mrdoc/pdf/7533_ukhls_accessibility_userg uide.pdf
- Lynn, Peter L., and Gundi Knies. 2015. "Understanding Society Quality Profile." in Understanding Society Quality Profile, edited by Peter L. Lynn and Gundi Knies. Colchester: ISER University of Essex.
- Manning, A. and Swaffield, J. (2008). The Gender Gap in Early-Career Wage Growth. *The Economic Journal*, 118(530), pp.983-1024.
- Mincer, J. (1958). Investment in Human Capital and Personal Income Distribution. *Journal of Political Economy*, 66(4), pp.281-302.
- Ngai, L. and Petrongolo, B. (2017). Gender Gaps and the Rise of the Service Economy. American Economic Journal: Macroeconomics, 9(4), pp.1-44.
- Olivetti, C. and Petrongolo, B. (2016). The Evolution of Gender Gaps in Industrialized Countries. *Annual Review of Economics*, 8(1), pp.405-434.
- *Olsen, W., Gash V., Grimshawe, D., Walthery P. and Vandecasteele L. (2009). The gender pay gap in the UK 1995–2007: Can SEM reveal causality? Online working paper, URL http://www.ccsr.ac.uk/werg/presentations/SEMaboutWagesProposedStudyCausality.pdf, accessed Jan. 2010.*
- Puhani, P. and Machin, S. (2003). Subject of Degree and the Gender Wage Differential Evidence from the UK and Germany. *SSRN Electronic Journal*.
- Rigdon, M. (2012). An Experimental Investigation of Gender Differences in Wage Negotiations. *SSRN Electronic Journal.*
- Spitze, G., Blau, F. and Ferber, M. (1987). The Economics of Women, Men, and Work. *Contemporary Sociology*, 16(3), p.337.
- Stanley, T. and Jarrell, S. (1998). Gender Wage Discrimination Bias? A Meta-Regression Analysis. *The Journal of Human Resources*, 33(4), p.947.
- University of Essex. Institute for Social and Economic Research, NatCen Social Research, Kantar Public. (2018). *Understanding Society: Waves 1-8, 2009-2017 and Harmonised BHPS: Waves 1-18, 1991-2009*. [data collection]. *11th Edition.* UK Data Service. SN: 6614, http://doi.org/10.5255/UKDA-SN-6614-12
- Weichselbaumer, D. and Winter-Ebmer, R. (2005). A Meta-Analysis of the International Gender Wage Gap. *Journal of Economic Surveys*, 19(3), pp.479-511.
- Wooldridge, J. (2012). Introductory Econometrics: A Modern Approach, Fifth Edition. Mason, Ohio: South-Western Cengage Learning.