Introduction

The aim of this research was to assess the effectiveness of the UK's Government gender pay gap regulation that came into effect in April 2018. This regulation requires firms with more than 250 employees to publicly disclose their gender pay gaps. Using an unbalanced set of panel data derived from the UK Government's gender pay gap reporting site I ran fixed effects regression models using the statistical software Stata which produced various coefficient estimates which I used to answer the key research questions of this investigation (see Appendix Table A1). Additionally, I explained the effect that I believe measurement issues are likely to have on my findings and I assessed the statistical and practical significance of my findings. My results indicate that between 2018 to 2020 the gender pay gap as measured by the main variable of interest from the data set; 'DiffMeanHourlyPercent' (defined in Appendix Table A1) decreased by -0.92 (2dp) percentage points for firms in the data set on average holding other factors constant. However, when considering outliers (defined in Appendix Table A1), the decrease in DiffMeanHourlyPercent between 2018 to 2020 was on average -0.85 (2dp) percentage points. These results are statistically significant but practically insignificant.

Literature Review

Introduction

The gender pay gap is a longstanding issue in society and has encouraged various researchers to investigate the causes of its origin as well as policies that have been implemented to try and address the issue. This literature review is based on carefully selected sources that provide insight into the following themes related to the research area of gender pay gaps:

- The economic theory explaining gender gaps
- The effect of regulation on gender pay gaps
- The impact of measurement issues on research into gender pay gaps

Economic theory used to explain gender pay gaps

The human capital model is a well-documented model used by Economists to explain gender pay gaps. Grybaite (2006) explains that this model proposes that gender pay gaps are due to men having higher human capital than women. The gender differences in human capital stem from the typical division of labour amongst households in which men work more than women. This results in women accumulating less experience in the labour market than men. Additionally, because women anticipate more intermittent participation in the labour market, they tend to have a lower incentive than men to invest their time in activities that increase their human capital (Grybaite, 2006).

However, Grybaite (2006) refers to data observed on European women which shows that they are often more educated and have more work experience than their male counterparts but typically work in low paid jobs. This shows that differences in human capital amongst males and females may not be the only factor that determines differences in earnings amongst males and females (Grybaite, 2006). This highlights an important limitation of the human capital model; it proposes that gender pay gaps are solely due to differences in human capital whereas there are other factors such as differences in the typical jobs held by males and females that affect gender pay gaps.

Another explanation for gender pay gaps is labour market discrimination. Grybaite (2006) states that "The portion of the pay gap that is not due to gender differences in qualifications is generally presumed to be due to labour market discrimination". There are various types of labour market discrimination including statistical discrimination which occurs where employers assess the suitability of women for roles based on the typical characteristics of women such as taking time off work to look after children (Grybaite, 2006). This leads to employers anticipating lower productivity from women compared to their male counterparts resulting in women being paid less than men (Grybaite, 2006).

However, according to Farrell (2004) as cited by Grybaite (2006) women are paid less than men as the typical lifestyles of women do not permit them to take up the risky and demanding careers that men typically pursue. Farrell (2004) as cited by Grybaite (2006) recommends that to address gender pay gaps women should simply choose a risky, demanding, and well-paid career.

Grybaite (2006) highlights the main theories used by Economists to explain gender pay gaps: the human capital model and labour market discrimination. Grybaite (2006) provides detailed explanations of the human capital model and labour market discrimination theories by frequently referring to and complementing the ideas and observations made by other Economists. However, I believe the main limitation of this article is a lack of integration between the theoretical explanations it provides, and the empirical evidence presented.

The effect of regulation on gender pay gaps

Research into the effect of firms providing transparency in relation to their gender pay gaps have has had a limited impact on gender pay gaps (Bennedsen et al, 2019). Legislation enacted in Denmark in 2006 that required firms with more than 35 employees to disclose "salary data broken down by gender for employee groups" reduced the gender pay gap between males and females by approximately two percentage points mainly by reducing the growth of male wages (Bennedsen et al, 2019).

Bennedsen et al (2019) have comprehensively shown the effect of regulation on the gender pay gaps in Denmark via the use of "employee-employer administrative data" and "difference-in-differences and difference-in-discontinuities" designs. However, the main limitation of this study is that it is focused on the estimated effects of regulation on gender pay gaps in Denmark. This means that the research of Bennedsen et al (2019) cannot be used to understand the effect of regulation on gender pay gaps in a broader set of countries such as countries in the EU.

Smith (2012) asserts that although there has been several years of EU legislation aimed at addressing gender wage inequalities, the gender pay gap in the EU have remained 'remarkably resilient' and is on average 16%. Smith (2012) highlights how the EU has deferred to using more soft mechanisms in recent years to address the gender pay gap. This approach involves social partners in the EU (Smith, 2012). Social partners according to Eurofound (2019) are "employer organisations and trade unions that are engaged in European social dialogue".

However, social partners have had an insignificant impact on gender pay gaps in the EU and this is due to several reasons (Smith, 2012). For example, many social partners have a lack of authority over members limiting their ability to encourage members to enact reforms to address gender pay gaps. Additionally, many social partners are dominated by male members, which make it difficult for gender pay reforms to be eagerly enacted (Smith, 2012).

Lastly, some social partners believe that it is the government's responsibility to address gender pay gaps and there is only a handful of countries that have social partners that actively try and encourage employers to increase the wages of women and to create plans to address their gender pay gaps (Smith 2012).

Smith (2012) effectively articulates the structural issues that the EU face by evaluating the hard and soft law approaches the EU has taken to address gender pay gaps. Smith (2012) focuses heavily on explaining the activities of social partners in the EU and EU regulations that have been aimed at addressing gender pay gaps.

The impact of measurement issues on research into gender pay gaps

Valet, Adriaans and Liebig (2018) highlight that research conclusions on the extent of earnings inequalities may be "systematically biased" due to misreporting, non-reporting, and errors in data collection. Surveys are the popular mode of data collection amongst researchers investigating earnings inequalities and this is a significant issue as earnings data is sensitive. Valet, Adriaans and Liebig (2018) note that researchers have found that the sensitivity of earnings data causes some individuals particularly those with low and high income not to report their earnings (i.e. non-reporting of earnings is explained via a u-shaped pattern).

Two theories are used to explain misreporting and non-reporting of earnings: the cognitive model and the rational choice model. The cognitive model suggests that respondents may fail to report their earnings because they "do understand the question at hand" or they lack accessibility to the information required to answer the question at hand. Additionally, the cognitive model suggests that misreporting occurs due to information bias; some respondents have more information than others in relation to gross earnings. For example, some respondents do not know the difference between gross and net income, and this causes them to misreport their earnings (Valet, Adriaans and Liebig, 2018).

The rational choice model on the other hand suggests that non-reporting occurs due to individuals deeming non-reporting to be a course of action that would maximise their utility. This is the case when respondents deem their earnings to be socially undesirable. Additionally, the rational choice model states that misreporting is due to respondents particularly those who have low or high incomes wanting to appear socially desirable. Low-income respondents tend to overreport their income and high-income respondents are likely to underreport their income (Valet, Adriaans and Liebig, 2018).

Misreporting and non-reporting can be decreased by the presence of an interviewer as the interviewer can provide additional information to give respondents more clarity in relation to survey questions (Valet, Adriaans and Liebig, 2018). However, if interviewers ask questions that probe into respondent's financial situation non-reporting and misreporting could increase due to the sensitivity of such questions (Valet, Adriaans and Liebig, 2018).

Based on the rational choice and cognitive models both misreporting and non-reporting can lead to biased conclusions about earnings inequalities by researchers. However, there is evidence to show that non-reporting results in a notable underestimation of earnings inequalities whereas misreporting does not bias "mean estimates on the population level" (Valet, Adriaans and Liebig, 2018).

Valet, Adriaans and Liebig (2018) provide a detailed account about the role of misreporting, non-reporting, and data collection errors on earnings inequality research by consistently referring to supporting evidence from other studies and economic theory. Valet, Adriaans and Liebig (2018) logically show how the research they have studied informed the hypotheses they proposed for their own research.

However, a shortcoming of the research of Valet, Adriaans and Liebig (2018) is that they excluded a subset of the German population who they described are in unstable jobs. They admit that the direct effect of this is an underestimation of earnings inequalities based on the administrative data they gathered. This is a clear example of how data collection processes can impact researchers' findings.

Conclusion

The human capital model would predict that regulation would not be effective at reducing the gender pay gap as men have higher human capital than women. Furthermore, the human capital model would suggest that there is no justification for the wages of women to increase unless women increase their human capital.

The literature I have reviewed on equal pay regulation in the EU would suggest that regulation in the UK may not have a significant effect on gender pay gaps. However, the EU faces a unique set of issues that may be the cause of the ineffectiveness of their equal pay policies. The heterogeneity of member states of the EU makes it difficult for uniform policies to be established and the creation of social partners to address the heterogeneity problem of the EU has led to inconsistent actions across the bloc and a gender pay gap that remains virtually unchanged.

The literature I have reviewed does not look at different factors that may affect the validity and reliability of the gender pay gaps firm report. My research will overcome this issue and will specifically look at how gender pay gaps are influenced by; the sex of firms' gender pay gap reporters and gender pay gap reporters that change between reporting years.

Methodology

The data set I was provided with is a secondary data set as it has been obtained from the UK government's public gender pay gap reporting site. The data set is also a panel data set as it contains a time series for each cross-sectional member (each firm). The most important variable is 'the main outcome of interest': DiffMeanHourlyPercent (defined in Appendix Table A1). The other important variables in the data set include the panel identifier variable that I choose 'CompanyNumber' (defined in Appendix Table A1) and the explanatory variables of interest (defined in Appendix Table A1).

My approach to the analysis

I conducted my analysis using the statistical software Stata. I used Stata to assess the following Ordinary Least Squares (OLS) regression assumptions: linearity in the parameters, sample variation in the explanatory variables, no multicollinearity, the expected value of the error term given the explanatory variable(s) is 0, no heteroskedasticity and normally distributed error terms. There two instances in which the no heteroskedasticity assumption was not met and to address this I computed robust standard errors for the relevant OLS estimates.

The key research questions 1 – 3 (see Appendix Table A1) required me to estimate the effect of the explanatory variables of interest (see Appendix Table A1) on gender pay gaps using regression models. The regression models that I could have used include but are not limited to fixed effects regression models and first differences regression models. However, I choose to use fixed effect regression models rather than first differences regression models. This was because the panel data set was unbalanced and a fixed effects regression model can be easily implemented regardless of whether a panel data set is unbalanced or balanced. I ran fixed effects regression models using Stata once with outliers and once without any outliers. This resulted in two sets of results from Stata. I then assessed the statistical significance and practical significance of my findings.

To answer the last key research question (see Appendix Table A1) I used my knowledge of the findings of the literature paper by Valet, Adriaans and Liebig (2018) that investigated the impact of measurement issues in earnings inequalities research. This literature paper helped me to identify and discuss the most likely potential measurement issues that might affect my interpretation of my findings and helped me to identify and discuss whether firms with larger or smaller gender pay gaps are likelier to misreport.

<u>Results</u>

Note: All figures stated in the Results section (apart from in tables) have been rounded to 2dp unless stated otherwise

At the 5% level of significance, the only coefficient estimates presented in this section that are not statistically significant from 0 are the coefficient estimates from the fixed effects regressions of DiffMeanHourlyPercent on Year2019 and LargeFirmXYear2019 (these regressions models included outliers).

How was the gender pay gap changed at the sampled set of firms between 2018, 2019 and 2020?

Table 1: Fixed effects regression of DiffMeanHourlyPercent on the independent dummy variables 'Year2019' and 'Year2020'

Note: This model includes outliers

The gender pay gap as measured by the dependent variable DiffMeanHourlyPercent including outliers on average for the sampled set of firms was 14.41 percent in 2018 holding other factors constant. In 2019, the gender pay gap as measured by DiffMeanHourlyPercent including outliers was on average -0.18 percentage points lower than in 2018 for a given firm in the sampled set of firms holding other factors constant. In 2020, the gender pay gap as measured by DiffMeanHourlyPercent including outliers was on average -0.92 percentage points lower than in 2018 for a given firm in the sampled set of firms holding other factors constant.

DiffMeanHourly Percent	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]
Year2019	-.181	.101	-1.79	.074	-.379	.017
Year2020	-.919	.134	-6.87	0	-1.181	-.656
Constant	14.406	.072	198.90	0	14.264	14.548

Mean dependent var	14.144	SD dependent var	15.588
R-squared	0.004	Number of obs	21845.000
F-test	24.180	Prob > F	0.000
Akaike crit. (AIC)	130378.083	Bayesian crit. (BIC)	130402.058

Table 2: Fixed effects regression of DiffMeanHourlyPercent on the independent dummy variables 'Year2019' and 'Year2020'

Note: This model excludes outliers

The gender pay gap as measured by the dependent variable DiffMeanHourlyPercent on average for the sampled set of firms was 14.49 percent in 2018 excluding outliers holding other factors constant. In 2019, the gender pay gap as measured by DiffMeanHourlyPercent excluding outliers was on average -0.26 percentage points lower than in 2018 for a given firm in the sampled set of firms holding other factors constant. In 2020, the gender pay gap as measured by DiffMeanHourlyPercent excluding outliers was on average -0.85 percentage points lower than in 2018 for a given firm in the sampled set of firms holding other factors constant.

DiffMeanHourly Percent	**Coef.**	**St.Err.**	**t-value**	**p-value**	**[95% Conf**	**Interval]**
Year2019	-.256	.076	-3.35	.001	-.406	-.106
Year2020	-.853	.101	-8.44	0	-1.051	-.655
Constant	14.492	.055	264.77	0	14.385	14.599

Mean dependent var	14.213	SD dependent var	14.741
R-squared	0.006	Number of obs	21837.000
F-test	35.618	Prob > F	0.000
Akaike crit. (AIC)	118091.320	Bayesian crit. (BIC)	118115.295

Are larger firms progressing at a different rate than smaller firms? Think about the specification of the model necessary to answer this question?

Table 3: Fixed effects regression of DiffMeanHourlyPercent on the independent dummy variables 'LargeFirmXYear2019' and 'LargeFirmXYear2020'

Note: This model includes outliers

The gender pay gap as measured by the dependent variable 'DiffMeanHourlyPercent' including outliers on average was 14.39 percent in 2018 for a large firm from the sampled set of firms holding other factors constant. In 2019, the DiffMeanHourlyPercent including outliers was on average -0.18 percentage points lower than in 2018 for large firms in the data set compared to small firms in the data set, holding other factors constant. In 2020, the DiffMeanHourlyPercent was on average -0.87 percentage points lower than in 2018 for large firms in the data set compared to small firms in the data set, holding other factors constant.

DiffMeanHourly Percent	**Coef.**	**St.Err.**	**t-value**	**p-value**	**[95% Conf**	**Interval]**
LargeFirmXYear 2019	-.179	.102	-1.75	.081	-.38	.022
LargeFirmXYear 2020	-.87	.135	-6.42	0	-1.136	-.605
Constant	14.387	.071	201.67	0	14.247	14.527

Mean dependent var	14.144	SD dependent var	15.588
R-squared	0.003	Number of obs	21845.000
F-test	21.033	Prob > F	0.000
Akaike crit. (AIC)	130389.341	Bayesian crit. (BIC)	130413.317

Table 4: Fixed effects regression of DiffMeanHourlyPercent on the independent dummy variables 'LargeFirmXYear2019' and 'LargeFirmXYear2020'

Note: This model excludes outliers

The gender pay gap as measured by the dependent variable DiffMeanHourlyPercent excluding outliers on average was 14.47 percent in 2018 for a large firm from the sampled set of firms holding other factors constant. In 2019, the DiffMeanHourlyPercent excluding outliers was on average -0.26 percentage points lower than in 2018 for large firms in the data set compared to small firms in the data set, holding other factors constant. In 2020, the DiffMeanHourlyPercent excluding outliers was on average -0.80 percentage points lower than in 2018 for large firms in the data set compared to small firms in the data set, holding other factors constant.

DiffMeanHourly Percent	**Coef.**	**St.Err.**	**t-value**	**p-value**	**[95% Conf**	**Interval]**
LargeFirmXYear 2019	-.257	.077	-3.31	.001	-.408	-.105
LargeFirmXYear 2020	-.802	.102	-7.83	0	-1.003	-.602
Constant	14.473	.054	268.42	0	14.367	14.578

Mean dependent var	14.213	SD dependent var	14.741
R-squared	0.005	Number of obs	21837.000
F-test	30.704	Prob > F	0.000
Akaike crit. (AIC)	118108.867	Bayesian crit. (BIC)	118132.841

Do firms who list a woman as the person responsible for reporting have wider or smaller gender pay gaps compared to firms which list a man responsible?

Table 5: The fixed effects regression of DiffMeanHourlyPercent on the dummy variables 'ResponsiblePersonFemaleXYear2019' and 'ResponsiblePersonFemaleXYear2020'

Note: This model includes outliers and robust standard errors were computed to address the heteroskedasticity I identified in the model

The gender pay gap as measured by the dependent variable 'DiffMeanHourlyPercent' on average was 14.26 percent in 2018 for a firm which had a female gender pay gap reporter from the sampled set of firms holding other factors constant. In 2019, the DiffMeanHourlyPercent was on average -0.31 percentage points lower than in 2018 for firms in the data set which had a female gender pay gap reporter compared to firms in the data set who did not have a female gender pay gap reporter, holding other factors constant. In 2020, the DiffMeanHourlyPercent was on average -0.87 percentage points lower than in 2018 for firms in the data set which had a female gender pay gap reporter compared to firms in the data set who did not have a female gender pay gap reporter, holding other factors constant.

DiffMeanHourlyPercentt	**Coef.**	**St.Err.**	**t-value**	**p-value**	**[95% Conf**	**Interval]**
ResponsiblePersonFemaleXYear2019	-.306	.122	-2.52	.012	-.545	-.068
ResponsiblePersonFemaleXYear2020	-.868	.166	-5.24	0	-1.193	-.544
Constant	14.264	.027	524.53	0	14.21	14.317

Mean dependent var	14.144	SD dependent var	15.588
R-squared	0.002	Number of obs	21845.000
F-test	13.893	Prob > F	0.000
Akaike crit. (AIC)	130429.724	Bayesian crit. (BIC)	130445.708

Table 6: The fixed effects regression of DiffMeanHourlyPercent on the dummy variables 'ResponsiblePersonFemaleXYear2019' and 'ResponsiblePersonFemaleXYear2020'

Note: This model excludes outliers

The gender pay gap as measured by the dependent variable DiffMeanHourlyPercent excluding outliers on average was 14.33 percent in 2018 for a firm which had a female gender pay gap reporter from the sampled set of firms holding other factors constant. In 2019, the DiffMeanHourlyPercent excluding outliers was on average -0.31 percentage points lower than in 2018 for firms in the data set which had a female gender pay gap reporter compared to firms in the data set who did not have a female gender pay gap reporter, holding other factors constant. In 2020, the DiffMeanHourlyPercent excluding outliers was on average -0.87 percentage points lower than in 2018 for firms in the data set which had a female gender pay gap reporter compared to firms in the data set who did not have a female gender pay gap reporter, holding other factors constant.

The coefficient estimates in Table 5 and 6 indicate that based on the gender pay gap data for the reporting years so far, firms who have a female gender pay gap reporter on average have lower gender pay gaps than firms that do not have a female gender pay gap reporter, holding other factors constant.

DiffMeanHourlyPercent	**Coef.**	**St.Err.**	**t-value**	**p-value**	**[95% Conf**	**Interval]**
ResponsiblePersonFemaleXYear2019	-.306	.122	-2.52	.012	-.545	-.068
ResponsiblePersonFemaleXYear2020	-.868	.166	-5.24	0	-1.193	-.544
Constant	14.333	.027	526.87	0	14.279	14.386

Mean dependent var	14.213	SD dependent var	14.741
R-squared	0.003	Number of obs	21837.000
F-test	13.893	Prob > F	0.000
Akaike crit. (AIC)	118159.098	Bayesian crit. (BIC)	118175.081

What happens at firms where the responsible person changes between years?

Table 7: The fixed effects regression of DiffMeanHourlyPercent on the dummy variable 'responsible_person_changes'

Note: Robust standard errors were computed to address the heteroskedasticity I identified in the model

The DiffMeanHourlyPercent for firms that change the person responsible for reporting their gender pay gaps between years is on average -0.35 percentage points lower than the DiffMeanHourlyPercent for firms that do not change the person responsible for reporting their gender pay gaps between years, holding other factors constant.

```
F test that all u_i=0: F(9674, 12169) = 12.01               Prob > F = 0.0000

. asdoc xtreg DiffMeanHourlyPercent responsible_person_changes, vce(robust) fe
(File Myfile.doc already exists, option append was assumed)

Fixed-effects (within) regression               Number of obs     =     21,845
Group variable: CompanyNum~r                    Number of groups  =      9,675

R-sq:                                           Obs per group:
     within  = 0.0004                                         min =          1
     between = 0.0002                                         avg =        2.3
     overall = 0.0000                                         max =          3

                                                F(1,9674)         =       4.13
corr(u_i, Xb)  = -0.0142                        Prob > F          =     0.0422

                           (Std. Err. adjusted for 9,675 clusters in CompanyNumber)
------------------------------------------------------------------------------------
                           |               Robust
     DiffMeanHourlyPercent |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
---------------------------+--------------------------------------------------------
responsible_person_changes |  -.3496355   .1720468    -2.03   0.042    -.6868833   -.0123877
                     _cons |   14.19907   .0271321   523.33   0.000     14.14589    14.25226
---------------------------+--------------------------------------------------------
                   sigma_u |  15.303948
                   sigma_e |  6.4198047
                       rho |   .8503624   (fraction of variance due to u_i)
------------------------------------------------------------------------------------
```

What potential measurement issues might affect the interpretation of your results? Are firms with larger or smaller gender pay gaps likelier to misreport? Assess what effect this has on your estimates (i.e. under/over estimation)?

The research of Valet, Adriaans and Liebig (2018) suggests that the most likely measurement issues in earnings inequalities research are the following: misreporting, non-reporting, and data collection errors. They also indicate all three of these issues could bias a researcher's findings.

Firms are more likely not to report or to misreport their gender pay gap data when they deem it to be socially undesirable (i.e. largely positive or negative). Since the average gender pay gap in the data set as measured by DiffMeanHourlyPercent is positive (14.14; see Appendix Table A2) I would expect any non-reported gender pay gap data to be positive on average. Also, as the average gender pay gap for firms in the data as measured by DiffMeanHourlyPercent is positive I expect that misreporting is likely to be due to largely positive (socially undesirable) gender pay gaps that firms have and do not want to reveal. Thus, I expect any misreported gender pay gap data to be understated on average. This would furthermore suggests that the coefficient estimates of the explanatory variables I have regressed on DiffMeanHourlyPercent are likely to be biased downwards.

Lastly, data collection errors could mean the values of DiffMeanHourlyPercent reported by firms in the data set for the reporting years so far could be higher or lower than their true values. If there is net understating in the values of DiffMeanHourlyPercent reported by firms then the coefficient estimates of the explanatory variables I have regressed on DiffMeanHourlyPercent are likely to be biased downwards. However, if there is net overstating the coefficient estimates of the explanatory variables, I have regressed on DiffMeanHourlyPercent are likely to be biased upwards.

Discussion

Implication of my analysis

The target population for this investigation was firms in the UK that have reported their gender pay gaps during 2018 - 2020. The sampled set of firms should consist of **all the large firms in the UK** (firms with more than 250 employees) who are by regulation required to report their gender pay gaps. The sample also includes some small and medium sized firms (firms with less than 250 employees) in the UK who are not required to report their gender pay gaps. The sample set of firms I have been provided with is highly reflective of the target population suggest that my findings are highly applicable to the target population.

There were four regressions that I ran of DiffMeanHourlyPercent on the different explanatory variables of interest (see Appendix Table A1). The coefficient estimates for the constant terms from the regressions were positive and were greater than 10 percent (as shown in the results section), indicating that these coefficient estimates are economically significant. On the other hand, the coefficient estimates obtained for the explanatory variables of interest (see Appendix Table A1) were negative and were less than one percentage point, which indicates that they are economically insignificant.

The fixed effect regression of DiffMeanHourlyPercent on the dummy variables Year2019 and Year2020 produced a coefficient estimate for Year2020 that was greater than the coefficient estimate of Year2019. This suggests that on average the firms in the data set made greater progress at reducing their gender pay gaps in 2020 compared to in 2019.

The R squared values computed from all the regressions that I ran for this investigation are less than one percent indicating that less than one percent of the variability in DiffMeanHourlyPercent is explained by the explanatory variables of interest (see Appendix Table A1). However, this is a result that I expected given the fact that I regressed DiffMeanHourlyPercent on a maximum of two indepedent variables. In future research having access to more data on independent variables that are likely to effect gender pay gaps such as years of experience and years of education will enable me to better explain the variability in DiffMeanHourlyPercent.

My findings counter the human capital model as this model predicts that gender pay gaps only change when there are changes in the human capital of men and women. However, as I cannot say with absolute certainty that my results are solely due to the regulation implemented by the UK Government, it is possible that changes in human capital may have been a factor that influenced the observed reduction in gender pay gaps as measured by DiffMeanHourlyPercent between 2018 to 2020 that I found.

Lastly, the UK government's gender pay gap regulation implemented in April 2018 is likely to have been more effective at reducing the gender pay gaps of firms whose workers earn a high level of wages compared to firms whose workers earn relatively low wages. This is because the coefficient estimates for the explanatory variables of interest (see Appendix Table A1) are estimated percentage point changes in DiffMeanHourlyPercent and a given percentage point change in gender pay gaps on average would have induced a larger absolute change in the gender pay gap for firms whose workers earn a high level of wages compared to firms whose workers earn a low level of wages.

Limitations of my analysis

I identified heteroskedasticity in two of the fixed effects regression models that I ran and to address this I computed robust standard errors for the relevant OLS regression coefficients. However, in future to ensure the heteroskedasticity problem is comprehensively addressed I will compute robust standard errors and I will also use the weighted least squares estimation method.

The research of Valet, P. Adriaans,J. Liebig,S. (2018) highlight that measurement issues such as misreporting, non-reporting and data collection errors are issues that cannot be eliminated as factors that may influence earnings inequalities research. For reporting years 2018, 2019 and 2020 there were some firms in the data set I have been provided with that did not report their gender pay gap data. Additionally, it is possible that employees at firms made data collection errors when gathering the gender pay gap data and this may explain why I identified eight outliers from studentised residual plots that I produced using Stata. Lastly, as the gender pay gap data that firms report to the UK Government is publicly disclosed it is possible that employees at firms may have understated their gender pay gaps to prevent their employer from being negatively viewed by the public.

The issues that I have mentioned can be mitigated in research I conduct in the future as assuming the UK government continues to use the gender pay gap transparency regulation to address gender pay gaps there will be more reporting data for firms. This would increase the size of the panel data set that I would have access to and would increase the reliability of findings that I derive. This should consequently undermine the effect of non-reported and misreported data as well as data collection errors.

Conclusion

The UK Government should maintain the current regulation that they have implemented to reduce gender pay gaps as thus far it has resulted in year on year decreases on average in gender pays gaps for firms as measured by the main variable of interest DiffMeanHourlyPercent, holding other factors constant.

Bibliography

1. Grybaite,V (2006). *Analysis of theoretical approaches to gender pay gap. Journal of Business Economics and Management,* 7:2, pp. 85-91
2. Becker, G.S. (1985). *Human Capital, Effort, and the Sexual Division of Labor.* Online at: https://www.jstor.org/stable/pdf/2534997.pdf?refreqid=excelsior%3A51a9d66dac18fe8c44415160a87e0ff6 (accessed 02/01/2020).
3. Bennedsen et al (2019). *Do Firms respond to gender pay gap transparency? Online at:* https://www.nber.org/system/files/working_papers/w25435/w25435.pdf(accessed 03/01/2020).
4. Smith, M (2012). *Social regulation of the gender pay gap in the EU. European Journal of Industrial Relations.* 18(4), pp. 365–380
5. Eurofound (2019). European social partners. Online at: *https://www.eurofound.europa.eu/observatories/eurwork/industrial-relations-dictionary/european-social-partners* (accessed 03/01/2020).
6. Valet, P. Adriaans, J. Liebig,S. (2018). *Comparing survey data and administrative records on gross earnings: nonreporting, misreporting, interviewer presence and earnings inequality. Online at:* https://search.proquest.com/docview/2037858050?pq-origsite=gscholar&fromopenview=true (accessed 03/01/2020).
7. Gov.uk (n.d.). Business population estimates for the UK and regions: 2019 statistical release (HTML). Online at: https://www.gov.uk/government/publications/business-population-estimates-2019/business-population-estimates-for-the-uk-and-regions-2019-statistical-release-html

Appendix Table A1: Key Terminology

Key Term	**Definition**
DiffMeanHourlyPercent	The difference between male base wages and female base wages on average as a percentage of male base wages
Outliers	Data points on studentised residual plots that are >20 or < -20 standard deviations from the mean
Explanatory variables of interest	• Year2019 • Year2020 • LargeFirmXYear2019 • LargeFirmXYear2020 • ResponsiblePersonFemaleXYear2019 • ResponsiblePersonFemaleXYear2020 • Responsible_person_changes
Key research questions	1. How has the gender pay gap changed at the sampled set of firms between 2018,2019 and 2020? 2. Are larger firms progressing at a different rate than smaller firms? Think about the specification of the model necessary to answer this question 3. Do firms who list a woman as the person responsible for reporting have wider or smaller gender pay gaps compared to firms which list a man responsible? What happens at the firms where the responsible person changes between years? 4. What potential measurement issues might affect the interpretation of your results? Are firms with larger or smaller gender pay gaps likelier to misreport? Assess what effect this has on your estimates (i.e., over or under estimation)
CompanyNumber	A unique numerical identifier for each firm

Appendix Table A2: Descriptive statistics for DiffMeanHourlyPercent

Descriptive Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
DiffMeanHourlyPercent	21845	14.144	15.588	-499.9	159