Centre for Economic Performance

Discussion Paper ISSN 2042-2695

No.1750
March 2021

Wage responses to gender pay gap reporting requirements

Jack Blundell

THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

30 YEARS CEP

Economic and Social Research Council

Abstract
In this paper I study a policy in which employers are required to publicly report gender pay gap statistics. Proponents argue that increasing the information available to workers and consumers places pressure on firms to close pay gaps, but opponents argue that such policies are poorly targeted and ineffective. This paper contributes to the debate by analyzing the UK's recent reporting policy, in which employers are mandated to publicly report simple measures of their gender pay gap each year. Exploiting a discontinuous size threshold in the policy's coverage, I apply a difference-in-difference strategy to linked employer-employee payroll data. I find that the introduction of reporting requirements led to a 1.6 percentage-point narrowing of the gender pay gap at affected employers. This large-magnitude effect is primarily due to a decline in male wages within affected employers, and is not caused by a change in the composition of the workforce. To explain this effect, I propose that a worker preference against high pay gap employers induces the closing of pay gaps upon information revelation. Newly-gathered survey evidence shows that female workers in particular exhibit a significant preference for low pay gap employers. In a hypothetical choice experiment, over half of women accept a 2.5% lower salary to avoid a high pay gap employer. I also demonstrate substantial heterogeneity in the interpretation of pay gap statistics across workers, and show that this affects their valuation of jobs at employers with different pay gaps.

Key words: gender pay gap; gender pay gap reporting; transparency; discrimination; information; public policy
JEL codes: J31; J38; J33; J71; J78; J16; D82; D83

This paper was produced as part of the Centre's Labour Markets Programme. The Centre for Economic Performance is financed by the Economic and Social Research Council.

This paper has benefited substantially from discussions with Caroline Hoxby, Nick Bloom, Luigi Pistaferri, Guido Imbens, Paul Oyer, Isaac Sorkin, Ed Lazear, Kathryn Shaw, Barbara Petrongolo, Rebecca Diamond, Matt Gentzkow and Mitch Hoffman. The author thanks Stanford Institute for Economic Policy Research for support, in particular the Dixon and Carol Doll Graduate Fellowship and the George P. Shultz Graduate Fellowship. Survey data gathered for this paper was classified as exempt by Stanford IRB (Protocol 55685). This work contains statistical data from ONS which is Crown Copyright. The use of the ONS statistical data in this work does not imply the endorsement of the ONS in relation to the interpretation or analysis of the statistical data. This work uses research datasets which may not exactly reproduce National Statistics aggregates.

Jack Blundell, Stanford University and Centre for Economic Performance, London School of Economics.

Published by
Centre for Economic Performance
London School of Economics and Political Science
Houghton Street
London WC2A 2AE

1 Introduction

The gender pay gap, broadly defined as the difference between male and female earnings, is one of the most salient aspects of modern-day inequality. As shown in Figure 1, in the UK and the US the gap fell rapidly during the 1970s and 1980s. From the mid 2000s in the US and 2010 in the UK, the rate of decline has fallen. Despite progress in the 20th century, at its current rate the gap among full-time workers in the United States would take another 82 years to close fully.[1] Another much-cited statistic is that to earn the same annual salary as men, women in the US would have to work an extra 39 days per year (Pew, 2018).

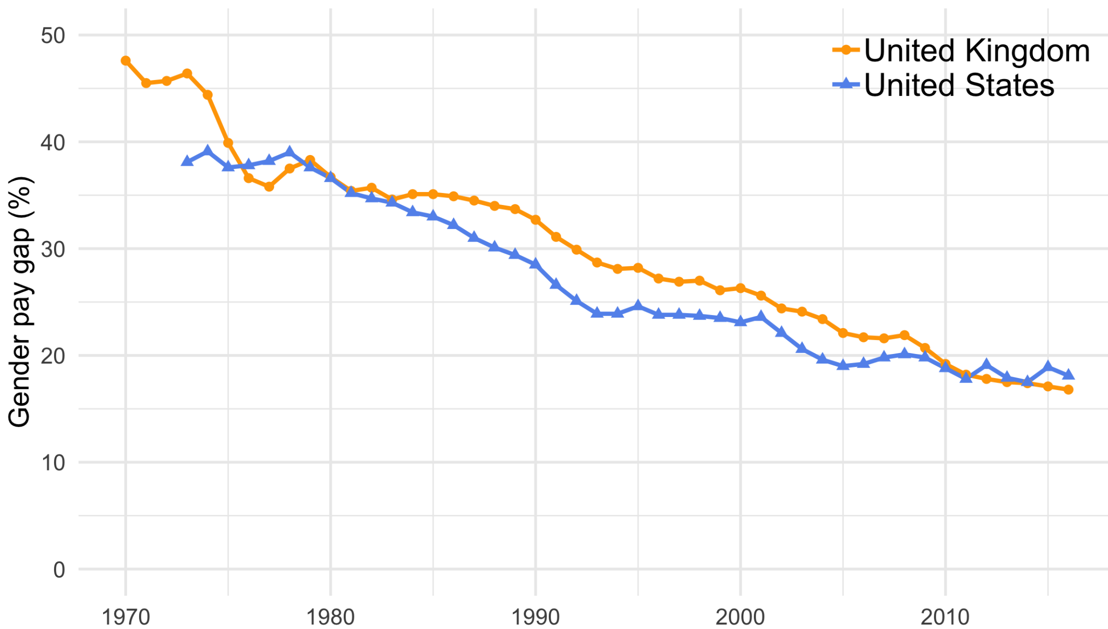


Figure 1: Gender pay gaps in the UK and the US

Notes: Gender pay gap defined as the raw difference between median male and median female earnings. Statistics refer only to full-time workers. Source: OECD

Throughout the 20th century, the United States and other developed economies introduced legislation to equalize opportunities across genders. This has led to convergence across many dimensions. For example, in both the US and the UK, female enrolment in higher education now exceeds that of men. The representation of women in the legislature of both countries has also increased dramatically. Despite these advances, pay gaps have proven to be persistent. Some argue that gaps are the result of men and women having heterogeneous tastes or innate advantages for different jobs. However, others would instead associate the pay gap with discrimination, constrained opportunities or differences in bargaining power and have argued for further interventions in the labor market. Faced with rising public pressure, policy makers are considering increasingly creative policies to tackle the pay gap, many of which are focused on improving wage transparency and shifting accountability towards employers.

This project concerns one such policy in the United Kingdom which explicitly targets the

[1] Author's calculation based on trend since 2006 in OECD data provided in Figure 1.

gender pay gap by requiring large employers to publicly disclose a set of gender pay statistics. Since 2017, all private-sector firms and public-sector organizations with over 250 UK employees must annually declare the raw difference between the average hourly wages of men and those of women, alongside several other statistics. Importantly, these statistics make no adjustment for occupation, part-time status or tenure, all of which statistically explain the gender pay gap. The policy covered more than half of all UK employees and generated ample media attention both in the UK and internationally. My research question is whether this policy was effective in its core goal of reducing wage gaps, on which there has to date been no rigorous evaluation.[2]

To address my primary research question, I use an administrative dataset consisting of payroll records. This dataset represents the premier source of information on earnings in the UK and is the basis for official wage statistics. The records are drawn from a 1% random sample of all private and public sector employees.

My empirical strategy exploits the discontinuous threshold of 250 employees, above which employers are required to report their pay gaps. In this difference-in-difference design, I compare wages of workers at employers above and below the threshold before and after the policy introduction. I find that the introduction of the policy led to a 1.6 percentage point increase in women's hourly wages relative to those of men. The inclusion of worker X firm fixed effects means that the effect is within workers' employment spells at a particular employer. The effect was not driven by a reshuffling of highly paid women into affected employers, but rather due to changes in individual workers' wages. The effect is primarily driven by a fall in male wages. Relative to the baseline gender gap in my estimation sample of 8.6 percentage points, this is a large effect, corresponding to 19% of the total gender gap.

To complement the main analysis, I gather new survey data covering UK workers. The goal of the survey is to test the hypothesis that worker responses drive the main effect identified in the wage data. The survey demonstrates that awareness of the policy is high, with most workers interacting with pay gap information through media reports. Using a hypothetical choice experiment, I find that while male workers are insensitive to pay gap information, over half of women would accept a 2.5% lower salary to avoid the employer with the highest pay gap in their industry. On average, women are prepared to accept 4.9% lower pay to avoid this employer. This is particularly high among younger women. The large magnitudes of this responses, combined with the high awareness of the policy, are consistent with the proposed mechanism.

One key advantage of the survey is that we are able to better understand how workers interpret pay gap information. I demonstrate that workers' interpretation of pay gap information matters. When workers view pay gaps as reflecting differences in an employer's values or in the ethical stance of managers, workers are more averse to high pay gap employers. When workers perceive pay gaps to be due to differences in skills of men and women at employers, they are less sensitive

[2] Concurrent to the development of the current paper, another research team has released a working paper on the policy (Duchini et al., 2020). Though the focus of Duchini et al. (2020) is distinct, their core results on the wage effects of the policy are consistent with those shown here.

to pay gap information.

This paper relates closely to a wide recent literature on the impacts of pay transparency on a variety of outcomes (Card et al. (2012), Breza et al. (2018), Mas (2016), Cullen and Perez-Truglia (2018)). These papers find pay transparency to have strong effects on pay inequality, worker morale and output. Particularly of interest to the current paper is the Mas (2017) finding that exposure to media had a significant effect in restraining top wages in the context of California's public sector pay disclosure policy. Also closely tied to the current paper, Baker et al. (2019) find that public sector pay transparency laws reduced the gender pay gap among university faculty.

There exist a small number of papers directly investigating gender pay gap reporting laws. Bennedsen et al. (2019) find that a pay gap disclosure policy in Denmark successfully reduced the pay gap within affected firms, increasing female earnings by 2 percentage points relative to those of men. Akin to my results, the effect was driven primarily by lower wages for men. They also find that the policy had a negative effect on firm productivity. Gulyas et al. (2020) study a similar law in Austria, finding that the policy led to an increase in the retention rate of workers but no effect on wages.

The policies investigated in these studies are substantively different to the UK policy. In both Denmark and Austria, pay gaps were not made publicly available, but rather were just available to employees within the firm. Pay gap reports were also made specific to occupations, in contrast to the cruder unconditional nature of the UK's reporting policy.

There are a number of papers on the impact of Corporate Social Responsibility (CSR) metrics, which gender pay gap statistics can be considered a form of. A recent related paper in this field is Hedblom et al. (2019), who find that firms advertising CSR see to an uptick in applications from prospective employees.

As the extended literature review given in Appendix A demonstrates, my findings are of interest to multiple academic disciplines. They also constitute a timely contribution to an active policy debate. Many countries are implementing similar pay gap reporting policies, and there is to date very little thorough investigation of whether these policies are effective in their primary goal. In places where the policy already exists, such as the UK, the evidence shown here is valuable in understanding whether the policy should be retracted, or whether it should be extended to broader dimensions of inequality, such as ethnicity or social class, or extended to a wider set of employers.[3]

This paper is structured as follows. In Section 2 I outline the UK policy in detail. Section 3 contains a discussion of the primary datasets used. In Section 4 I outline my difference-in-difference strategy, discussing the key assumptions and threats to identification, before presenting my main results and robustness tests in Section 5. In Section 6 I present evidence on worker responses as a potential mechanism, drawing on the new survey data. In Section 7 I conclude

[3]The analysis has particular relevance for the UK in the short-term, as at the time of writing the reporting policy studied in this paper has been suspended due to the Covid-19 crisis.

and propose avenues for further research. Supplementary analysis and discussion is found in Appendices A through I.

2 The UK's gender pay gap reporting policy

In this section I outline the UK policy in detail. An additional overview of international pay gap policies is given in Appendix E.

As was demonstrated in Figure 1, the gender pay gap in the United Kingdom has followed a strikingly close trend to that of the United States and as of 2016 stood at 16.8% for full-time workers.[4] Building on the Equality Act of 2010, in 2017 the UK government introduced new reporting requirements designed to tackle the pay gap. Since 2017, all firms and public sector bodies with over 250 UK employees are required each year to report several statistics on their gender pay gap. Organizations that are part of a group must report individually, and foreign-based companies are subject to the regulations if they employ 250 or more employees in the UK. This directly covered approximately 10,000 employers, representing the majority of UK employment. Smaller employers were given the option to report should they choose to, but in practice very few have done so.[5]

The government consulted with firms and experts between 2010 and 2017. Given this, it is plausible that the policy was partially anticipated, though the precise details and implementation year were not clear until close to the policy introduction. More detail on the legal background and development of the policy is given in Appendix C. The majority of employers reported their pay gap just before the first deadline of April 2018 and the releases received ample media attention. Several employers were named and shamed in newspapers as having particularly extreme gender pay gaps. For example, the Independent, a prominent daily newspaper ran a story titled "Gender pay gap: worst offenders in each sector revealed as reporting deadline passes" (Independent, 2018). In this article a championship football club and an airline were revealed as having among the greatest gender pay gaps in the country. In Figure 9 in Appendix B, I plot the number of UK print and online news articles mentioning the term "gender pay gap" each week, showing spikes in interest in April 2018 and 2019 around the deadline dates. Pay gap information is clearly presented on a government website which encourages users to directly compare the pay gaps across employers by providing a side-by-side comparison tool.

Rather than a single summary pay gap statistic, under the policy employers must report several features of their pay distribution for a particular snapshot date each Spring. The first snapshot of pay, due April 2018, refers to early April 2017. Statistics to be reported include the mean and median gap in hourly wages, the equivalent gaps for annual bonus payments and

[4] This is the OECD figure. The reported pay gap varies quite substantially depending on the source, as a number of choices can be made about how the figure is calculated and which data is used.

[5] 285 employers with under 250 employees opted to submit a report in the first round of pay gap reporting, out of over 2 million employers. These smaller employers constitute under 3% of total reports.

finally the shares of men and women in each quartile of the within-employer unconditional pay distribution. At no point are absolute values of pay reported. More detail on precise government guidance on how pay variables are to be calculated and on which employees are covered is provided in Appendix C.

Histograms of median and mean gender pay gaps from the 2017 pay period are plotted in Figure 2. Mean pay gaps tend to be higher than median gaps due to a skewness in the distribution of pay within organizations. The first round of reports revealed over 80% of reporting employers had a median pay gap in favor of men. The median employer had a median pay gap of 9.3% and the mean of the median pay gaps was 11.79%. Typically, median wage gaps have been reported in the press and have been emphasized in government communications.

A crucial detail of the UK policy is that the pay gap statistics do not account for differences in occupation, experience or part-time/full-time status.[6] Given this, employers and commentators have questioned whether employers should be made accountable for these aggregate measures of pay gaps. Pay gaps declared are in many cases only loosely related to the concept of 'equal pay for equal work', to which workers are legally entitled. While media coverage has emphasized the discrimination interpretation of pay gaps, in many cases a high pay gap is quite clearly driven by occupational segregation.[7] A salient example of this is provided in the case study of an airline company in Appendix D. If the target is to measure unequal pay for equal work, the gender pay gap statistics here are at best a noisy measure. The crudeness of the measure leaves ample room for interpretation, which I provide evidence on in my survey in Section 5.

Alongside pay gap statistics, employers also have the option of releasing a narrative statement. These statements tend to include a description of their efforts to reduce their pay gap. They often also included a justification for their pay gap, typically citing occupational segregation, and emphasizing that they abided by Equal Pay legislation. The case study in Appendix D gives an overview of a typical accompanying narrative.

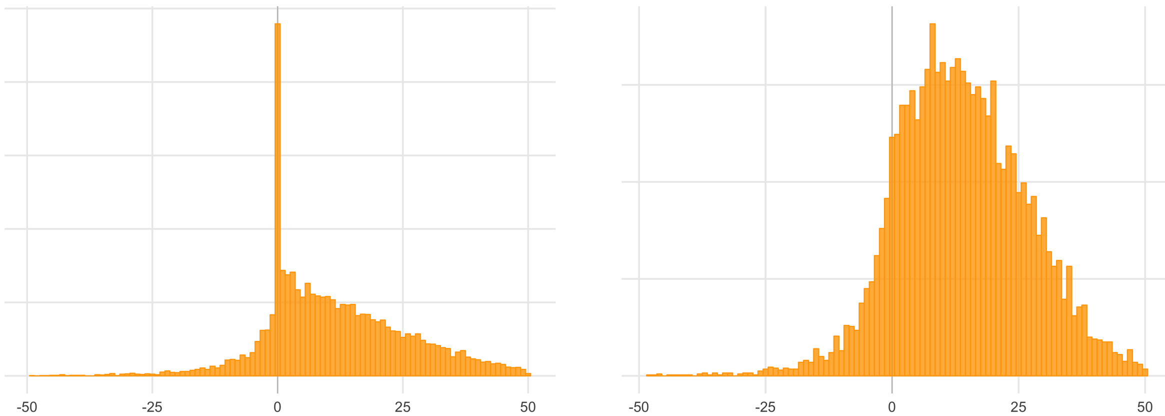
Other than a recent Covid-19 suspension, the policy has been relatively unchanged since its introduction and is due for review in 2022. In the years since the policy was introduced, several issues have been raised concerning its design and implementation. One common argument is that employers often do not have the administrative capabilities required to calculate the statistics correctly. Statistician Nigel Marriot has estimated that between 10 and 15% of employers made errors in their initial gender pay gap reports (Marriott, 2020). Consistent with mis-reporting, there is a clear spike at 0 in Figure 2(a). While a median pay gap of zero is plausible in many large firms with homogeneous workers, such as grocery chains, the magnitude of the spike is suggestive of mis-reporting. Mean pay gap reports exhibit no such spike.

Enforcement of the regulations has been widely criticized. Enforcement powers lie with the

[6]Though the primary pay gap statistics are based on hourly pay, there is a significant part-time hourly-pay penalty in the UK (Manning and Petrongolo, 2008). Since far more women work part-time than men this is an important explanatory factor in the gender pay gap.

[7]Here I take discrimination to mean unequal pay for equal work, which is a narrow definition. Occupational segregation itself could reflect broader notions of discrimination.

Figure 2: Pay gap reports



(a) Median

(b) Mean

Notes: Distribution of reports from initial round of reported, referring to 2017 snapshot. Pay gap reports correspond to how much less women are paid than men as a proportion of male earnings, so that a figure of 20% corresponds to women being paid 80% of what men are paid. Source: UK Gender pay gap service

UK's Equality and Human Rights Commission (EHRC). At a Treasury Select committee meeting in June 2019, it was stated that the EHRC has sent letters to employers who have not complied with reporting requirements, but are yet to employ any harder punishments (Parliament, 2019). Enforcement to date is based on whether or not an employer submits their data at all, rather than concerns over whether or not an employer is submitting false data, either by mistake or to deliberately mislead.

A final issue raised by commentators is that company partners are in general excluded from figures, since they are not technically employees. This particularly matters for business services and law firms and is further complicated by the fact that some of these firms decided to include partners, whereas others did not. Given the high pay of these individuals relative to other workers within their organizations, this can have a significant effect on the reported pay gap. Particularly in these industries then, the reported pay gaps are a noisy measure of the differences in total remuneration between male and female workers.

The target of the policy is explicit. The government seeks to narrow the gender pay gap by increasing the salience and availability of firms' pay gap information to workers and consumers. Though the intuition is clear, in Appendix F I provide a simple partial equilibrium model of the firm to demonstrate how preferences against high pay gap firms on the worker and consumer side can lead firms to narrow their pay gaps. In Section 6 I discuss this further and provide evidence on worker responses as a potential mechanism.

3 Data

My analysis is primarily based on the Annual Survey of Hours and Earnings (ASHE).[8] This is the premier source of earnings information and forms the basis for many official wage statistics in the UK. It is a 1% sample of all employees based on the UK equivalent of Social Security Numbers, with a panel structure which makes it possible to follow workers over time. While the sampling is at the employee level, employers complete the survey from their payroll records, so the information is considered to be highly accurate. The survey covers both the public and private sectors, but as it is administered via employers it excludes the self-employed, who constituted 15% of UK employment as of 2019.

The full survey delivers useable information on 140,000 to 180,000 employees each year. The variables available include detailed wage and hours information for a snapshot week in April, along with a limited measure of annual wages and limited worker demographics. In part due to concerns discussed in Appendix G, I restrict my sample to full-time workers aged 18-55.

Consistent with other academic work using this data, observations are dropped if there is a loss of pay due to absence in the reference period. To be consistent with pay gap reports, our main outcome is hourly wage, deflated by the CPI. Observations are at the job rather than worker level, however the restriction to full-time workers means that in practice, very few workers in our sample hold multiple jobs. Further detail on the ASHE dataset is provided in Appendix G.

To complement the payroll data in ASHE and provide detail on mechanisms, I collected new survey data on UK workers.[9] In Summer 2020, a 7-minute survey was administered to 2,000 UK workers through the Prolific Academic platform, resulting in 1,840 usable responses. Respondents were asked a number of questions on attitudes to work, awareness of the gender pay gap reporting policy and preferences for employers with different pay gaps. Attention checks were placed throughout the survey to validate that respondents were considering their responses carefully.

A full description of the structure of the survey is provided in Appendix H. As is typically the case when using survey data, the sample is not fully representative of the UK working population. Given this, responses are weighted to match population age, sex and education aggregates. Details on the weighting strategy are given in Appendix H.6 along with summary statistics of the unweighted and weighted data. All results reported in the main text use the weighted responses unless otherwise indicated.

I also use two Bureau van Dijk datasets on firms. This is firstly Amadeus,[10] which con-

[8] Office for National Statistics. (2020). Annual Survey of Hours and Earnings, 1997-2020: Secure Access. [data collection]. 17th Edition. UK Data Service. SN: 6689, http://doi.org/10.5255/UKDS-SN-6689-16

[9] Survey data gathered for the first chapter of this dissertation was classified as exempt by Stanford IRB (Protocol 55685).

[10] Amadeus, 2020, Bureau van Dijk Electronic Publishing, accessed through Wharton Research Data Services

tains company accounts records for European companies, and FAME,[11] which is the UK-specific version of this data. These are useful for generating statistics on firm characteristics. Stock price information is obtained from Compustat.[12] Finally, I use the pay gap information itself, as released by the UK Government.

4 Empirical Strategy

4.1 Main specification

The principal goal of this paper is to assess the causal impact of the UK's reporting policy on the pay gap. One simple approach would be to inspect the time trend of the aggregate gender pay gap in the UK. However, this is unlikely to yield useful insights. Male and female workers are found in very different sectors, work in different occupations and hold different levels of seniority within firms. They also differ in their responses to aggregate shocks (Guvenen et al., 2014). This implies that male and female wages will exhibit different trends over short time periods, independent of the pay gap legislation.

Given the above concern, the empirical strategy adopted here exploits the exclusion of small firms from the reporting requirements. This size threshold is a feature shared by most gender pay gap reporting policies internationally, motivated both by the administrative burden reporting arguably places on smaller employers and also by concerns over data confidentiality. In the UK policy, only employers with 250 or more employees were subject to the requirements. This presents a natural control group of employers with under 250 employees, to complement the over-time variation induced by the policy.

An ideal specification would be implemented at the employer level:

$$GPG_{jt} = \beta_0 + \beta_1 \text{Over250}_{jt} + \beta_2 \text{Post}_t + \gamma \text{Over250}_{jt} * \text{Post}_t + \epsilon_{jt} \tag{1}$$

where GPG_{jt} is the gender pay gap of employer j in period t, Over250_{jt} indicates whether an employer has over 250 employees, Post_t indicates whether the period is before or after the policy introduction and ϵ_{jt} an unobserved error term.

Specification (1) is not feasible. It requires employer-level estimates of the gender pay gap, which are not available for the UK outside of the policy-induced reported statistics. With ASHE's 1% sample, only for the very largest employers can we estimate a pay gap with any degree of accuracy. In my cleaned data sample, for most employers with fewer than 250 employees, there exist only one or two observations.[13]

[11] FAME, 2020, Bureau van Dijk Electronic Publishing, accessed through Wharton Research Data Services

[12] Compustat, 2020, Standard and Poor's, accessed through Wharton Research Data Services

[13] The increasing availability of HMRC (the UK tax authority) data for research suggests that in future years, Specification (1) could become feasible.

Instead, the following worker-level regression is estimated:

$$\begin{aligned} log(wage_{ijt}) = \beta_0 + \beta_1 \text{Over250}_{jt} + \beta_2 \text{Post}_t + \beta_3 \text{Female}_i \\ +\beta_4 \text{Over250}_{jt} * \text{Post}_t + \beta_5 \text{Over250}_{jt} * \text{Female}_i + \beta_6 \text{Post}_t * \text{Female}_i \\ +\gamma \text{Over250}_{jt} * \text{Post}_t * \text{Female}_i + \delta X_{ijt} + \nu_{ij} + \epsilon_{ijt} \end{aligned} \tag{2}$$

Here the outcome variable is the log hourly wage of worker i at employer j in period t. X_{ijt} is a set of worker-level controls and ν_{ij} is a worker X firm fixed effect.[14] This is a triple-difference specification, in line with that of Bennedsen et al. (2019).[15] The coefficient of interest is γ, which gives the relative changes in the wages of female workers at larger firms before and after the policy was introduced. The inclusion of worker X firm fixed effects means that any effect identified is within employment spell. The identification of γ will then stem entirely from changes to the wages of workers at the same firm before and after the policy introduction.

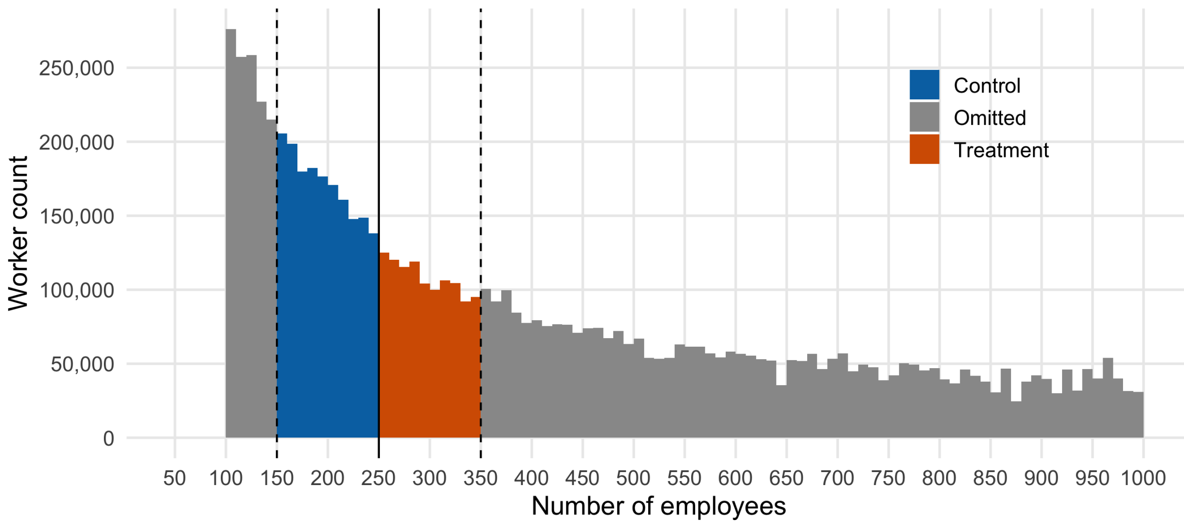
Under what conditions will a regression estimate of γ provide an unbiased and consistent estimate of the effect of the policy on the gender pay gap? Before questions of econometric identification, a first simple but important consideration is how parameter γ relates to the pay gap response within employers. If most employers had a pay gap in favor of women and were induced to narrow their pay gap as a result of the policy, $\gamma < 0$ would be expected. In practice, the reports shown in Figure 2 demonstrate that the majority of employers have pay gaps in favor of men. This means that if the policy induces a narrowing of the pay gap within employers, we would see $\gamma > 0$. This corresponds to an increase in the wages of women relative to men on aggregate.

This specification is averaging over employers with widely heterogeneous pay gaps, who could be affected by the policy's introduction differently. Employers with low or zero pay gaps have no incentive to react. The actions of employers with pay gaps in favor of women could plausibly partly offset those with pay gaps in favor of men. There is little evidence that employers with pay gaps in favor of women have been subject to any attention in the media, so to the extent that media plays a key role in driving any effects, or reflects broader public sentiment, this seems unlikely. The prospect of full firm-level data in the future would help address concerns such as this.

[14] The inclusion of fixed effects ν_{ij} means that Female_i is not identified when estimating this equation, but it is left in here for expositional clarity.

[15] Though they remain commonplace, difference-in-difference methods have been subject to significant critiques (Bertrand et al. (2004), Athey and Imbens (2006)). While the methodology applied here is subject to these criticisms, I will attempt to alleviate some of these concerns by applying a now-standard battery of robustness tests.

Figure 3: Treatment and control group selection



Notes: Histogram of firm size (at worker level) with treatment and control group indicated. Workers at firms with under 100 employees or over 1,000 employees omitted from figure. Bin width = 50. Source: Amadeus

4.2 Common trends

As in any difference-in-difference strategy, a first key identifying assumption is common trends. In this setting, the required assumption is that in the absence of the policy introduction, the gap between male and female wages would have followed the same trend in larger and smaller employers.

The common trends assumption is more plausible if the treatment and control group share similar characteristics. Employer size is correlated with other aspects of employers, most importantly with how internal pay setting operates.[16] Rather than including all employers above and all employers below the threshold, a bandwidth is chosen around the threshold. Employers outside of this bandwidth are excluded, such that very large employers will not be compared to very small employers. The chosen baseline bandwidth is 200 employees, meaning that employers with 150-249 employees constitute a control group and employers with 250-349 employees a treatment group. Robustness will be shown around different choices of the bandwidth. The selection of the treatment and control group is shown graphically in Figure 3, along with the distribution of firm size at the worker level. This shows that the bandwidth is found in a dense part of the firm-size distribution, with many workers and firms falling into the treatment and control groups.

As 250 employees is used as the threshold identifying "large" firms in the UK, there are changes in other reporting requirements around this threshold. For example, firms above the threshold are mandated to submit more detailed accounts than firms below the threshold. To the

[16] Larger employers are more likely to have formal HR departments, with structured pay bands and less flexibility over pay.

best of my knowledge there were not any other reforms around this period which differentially affected firms based on this employment threshold.

Table 1 shows means and standard errors for the those at smaller employers (column (1)), larger employers (column (2)) and the combined sample (column (3)) for the latest pre-treatment year available among those in the main analysis sample. Comparing (1) and (2), we see that despite our narrowing of the sample to those in a bandwidth, there are some differences between the two groups. The treatment group, those working at larger firms, tend to have marginally higher earnings. The difference, for weekly and annual earnings, is significant at the 5% level. This is consistent with the substantial literature on the positive wage returns to firm size. Other than number of employees, which differs by construction, there are no remaining statistically significant differences between the two groups.

Column (4) of Table 1 shows the statistics for the full ASHE population.[17] Comparing means to those of the estimation sample in column (3), here we can see that our estimation sample is more male, which primarily results from the full-time worker restriction, also reflected in the hours variable. The estimation sample is older, higher earning, more experienced, works in higher-skilled occupations and at smaller firms than the average UK worker. While the local nature of my empirical strategy is beneficial for identification, it comes at a cost of generality. If treatment effects are heterogeneous, the local effects identified using the estimation sample may differ to the aggregate effect on the full sample.

I also estimate the following generalized difference-in-difference (event study) design:

$$\begin{aligned} log(wage_{ijt}) = \beta_0 + \beta_1 \text{Over250}_{jt} + \sum_{t=0}^{t=T_{max}} \beta_{2,t}\phi_t + \beta_3 \text{Female}_i \\ + \sum_{t=0}^{t=T_{max}} \beta_{4,t}\phi_t * \text{Over250}_{jt} + \sum_{t=0}^{t=T_{max}} \beta_{5,t}\phi_t * \text{Female}_i + \beta_6 \text{Over250}_{jt} * \text{Female}_i \\ + \sum_{t=0}^{t=T_{max}} \gamma_t\phi_t * \text{Over250}_{jt} * \text{Female}_i + \delta X_{ijt} + \nu_{ij} + \epsilon_{ijt} \end{aligned} \tag{3}$$

Here ϕ_t represent time fixed effects. In this specification, parameters β_2, β_4, β_5, and γ are now vectors, with an estimate for each time period. This allows the inspection of pre-trends, which indicates the plausibility of the common trends assumption. Flat pre-trends are indicated by $\gamma_t = 0$ for all periods before the policy introduction.

4.3 Timing and anticipation

For all main specifications, Post_t is set to equal 1 from 2018 onwards. The first reporting deadline was in 2018 and referred to pay as of April 2017. The 250-employee cutoff was built into the

[17] Firm-level summary statistics are contained in Table 9 in Appendix B.

Table 1: Summary statistics by employer size

	(1) Control	(2) Treatment	(3) Full	(4) Population
Male	0.60	0.62	0.61	0.48
	(0.007)	(0.008)	(0.005)	(0.001)
Age	38.32	38.42	38.36	36.72
	(0.134)	(0.167)	(0.105)	(0.022)
Hourly wage	15.74	16.05	15.86	13.73
	(0.119)	(0.158)	(0.095)	(0.018)
Weekly earnings	605.49	621.20	611.49	476.63
	(4.280)	(5.791)	(3.449)	(0.978)
Annual earnings	30,362	31,558	30,819	23,229
	(280)	(360)	(221)	(71)
Hours	39.43	39.52	39.47	32.79
	(0.088)	(0.108)	(0.068)	(0.026)
Tenure	6.46	6.74	6.56	5.65
	(0.091)	(0.120)	(0.072)	(0.015)
Skill (1-6)	2.88	2.87	2.88	2.50
	(0.024)	(0.030)	(0.019)	(0.004)
Employees in firm	196.07	295.39	234.01	13750.83
	(0.385)	(0.482)	(0.594)	(78.518)
N	5,495	3,397	8,892	232,149

Notes: Statistics based on latest available pre-treatment year for each worker. Mean and standard errors shown, with number of observations in final row. Wages and earnings in 2010 GBP. Column (1) contains control group sample, column (2) contains treatment group sample and column (3) contains the full estimation sample (columns (1) and (2) combined). Column (4) contains the full 18-55 ASHE population, conditional on presence in any of pre-treatment year 2012-2017. Source: ASHE

legislation several years before the policy was introduced, so one might argue that forward-looking firms adjusted their pay gaps earlier than 2018. Using the event study estimates, we will be able to inspect treatment dynamics fully to see how plausible this timing assumption is. If affected firms were induced to reduce their pay gaps in anticipation of the policy introduction, we would expect to see pre-trends and the estimated treatment effect would be biased towards zero.

More concerning for identification is that some firms with particularly high pay gaps may adjust their firm size in anticipation of the reporting requirement. To counteract this concern, specifications with a "donut hole", a missing mass around the 250 employee threshold, are estimated. These are designed to exclude any firms which could have manipulated their employment to land on one side of the threshold. This is illustrated in Figure 12 in Appendix B. I will also inspect the employer distribution for bunching around the threshold directly.

4.4 Spillovers

Another identifying assumption is the Stable Unit Treatment Value (SUTVA) assumption. In this setting, the assumption requires that the policy cannot have also affected workers at employers with fewer that 250 employees. There are three reasons to question this assumption.

Firstly, there are potential general equilibrium effects in the labor market. If wages for women at larger employers are increased as a result of the policy, smaller firms who share a common labor market will experience pressure to match these increases. In equilibrium, it would then be expected that women at smaller employers would also see relative wage increases. Given the substantial evidence demonstrating the presence of labor market frictions, it is unlikely that full adjustment would occur immediately, so we would expect there to be at most small effects in the short run. This would bias the estimated effect of the policy towards zero.

Secondly, one may be concerned that the policy not only affected those employers who were required to report but also led to a greater general awareness of pay gaps as an issue. Figure 9 in Appendix B shows the number of UK print and online news articles mentioning the term “gender pay gap" each week. There is a clear spike around the date at which the first deadline for reports in April 2018. This growing awareness of gender pay gaps could have affected all employers irrespective of size, either by changing norms in pay setting or due to pressure from workers. Again, this would bias the estimated effect of the policy in the above strategies towards zero.

Finally, smaller employers could anticipate that the reporting requirements would at some point in the future cover themselves also. This could be due to employers expecting to grow into the existing threshold. This concern is similar to the that of employers adjusting their size to land one side of the threshold. This would result in a bias towards zero and can be mitigated by the donut hole strategy discussed above and by inspecting the employer size distribution for bunching around the threshold. Relatedly, employers could expect the legislation to be extended to smaller employers at some point in the future.[18] This too would bias the estimated effect of the policy towards zero.

5 Results

In this section I present my main results, alongside a battery of robustness tests.

5.1 Main results

Table 2 shows estimates from the triple difference-in-difference strategy set out in Specification (2). Years 2012-2017 constitute the pre period and 2018-2019 the post period. Standard errors are clustered at the firm level.[19] The sample is as described in Section 3, constrained such

[18]This expansion of coverage to smaller employers occurred in the Austrian implementation of the policy.
[19]In this section, I will refer to ‘employers’ as ‘firms’.

that the estimation of worker X firm fixed effects is possible. This means that any single-year worker-firm employment spells are excluded from analysis.

Table 2: Triple difference-in-difference specifications

	(1)	(2)	(3)	(4)	(5)
>250	-0.0046	0.0097	0.000077	0.011***	0.011***
	(0.015)	(0.0089)	(0.0066)	(0.0042)	(0.0035)
Post	0.045***	0	0	0	0
	(0.011)	(.)	(.)	(.)	(.)
Female	-0.086***	-0.066***	-0.070***	0	0
	(0.012)	(0.0079)	(0.0096)	(.)	(.)
>250*Post	0.0044	-0.0095	-0.0099	-0.014**	-0.014***
	(0.018)	(0.011)	(0.0075)	(0.0056)	(0.0047)
Post*Female	0.012	0.000083	-0.0067	-0.000056	-0.000029
	(0.014)	(0.010)	(0.0083)	(0.0056)	(0.0047)
>250*Female	0.026	0.0068	-0.0047	-0.012*	-0.012**
	(0.018)	(0.012)	(0.012)	(0.0071)	(0.0059)
>250*Female*Post	0.0078	0.022	0.025*	0.015*	0.016**
	(0.025)	(0.016)	(0.013)	(0.0084)	(0.0070)
N	42,059	42,059	42,059	42,059	42,059
N workers	12,885	12,885	12,885	12,885	12,885
N firms	6,707	6,707	6,707	6,707	6,707
Controls		✓	✓	✓	✓
Firm FE			✓	✓	
Worker FE				✓	
Worker X Firm FE					✓

Notes: * p<0.1, ** p<0.05, *** p<0.01. Dependent variable is real log hourly earnings. Standard errors clustered by firm. Control variables are age, tenure, industry, skill, public sector and region, interacted with year fixed effects. Sample constrained to be identical to that of column (5) for all specifications. Source: ASHE

The final row contains the estimated effect of the policy on female wages relative to male wages. In column (1), there are no controls or fixed effects included. The estimated treatment effect is positive, small in magnitude and statistically insignificantly different to zero. The coefficient on 'Female' is negative and statistically significant, showing a pay gap in favor of men. The -0.086 estimate corresponds to an approximate 8.6% pay gap, lower that would be expected based on previous UK research. This can likely be explained by the sample selection. These are full-time workers with strong labor market attachment.

In column (2), control variables are added. These are age, tenure, industry, skill, public sector and region, interacted with year fixed effects. In column (3) of Table 2, firm fixed effects are included. Firm wage premiums are an important component of wages and hence absorb variation in the outcome variable, allowing for greater precision. Their inclusion results in an estimate of 0.025, statistically significant at the 10% level. In column (4), both firm and worker fixed effects are included, and in my baseline specification column (5) these are combined into worker X firm fixed effects. The estimated effect is 0.016, which is statistically significant at the 5% level. Again using the approximate interpretation of log differences as percentages, the interpretation is that the introduction of the policy led to a 1.6 percentage point increase in female hourly wages relative to those of men. As described in the previous section, the variation driving this estimate is within each workers' employer spell at a particular firm, meaning that this cannot be driven by affected employers hiring new highly paid women into their organization, or firing low paid women.

This constitutes a large increase when compared to the estimates on 'Female' in previous columns. Taking the estimate of -0.086 from column (1), this corresponds to a 19% decline in the gender pay gap. Compared to other reforms designed to reduce gender pay inequality, these estimates suggest the policy to be particularly effective. For example, Kleven et al. (2020) provide an exhaustive analysis of numerous expansions of parental leave and child care subsidies in Austria, finding that they had virtually no impact on gender convergence.

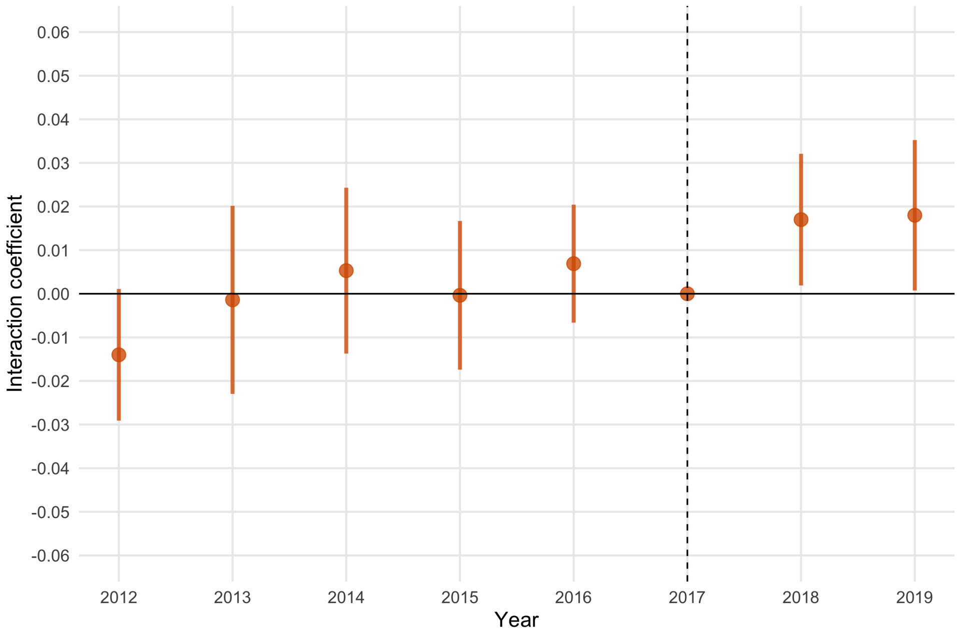
An inspection of the $>$250*Post coefficient demonstrates that the effect is driven primarily by a reduction in the pay of male workers. This is shown more clearly in the separate specifications in Table 10 in Appendix B. This is consistent with men experiencing nominal wage freezes due to the introduction of the policy. This also suggests that on average, wages fell as a result of the policy.

Figure 4 shows the year-by-year triple interaction estimates from the worker X firm fixed effects version of general difference-in-difference specification (3). The full set of estimates for this specification is given in Table 11 in Appendix B. Interactions for the year 2017 are normalized to equal zero. The figure demonstrates flat pre-trends. In none of the pre-intervention years is the treatment interaction coefficient statistically significant. In 2018 the coefficient becomes positive and statistically significantly different to zero, and in 2019 the coefficient is similar in magnitude and is also statistically significant. This figure lends support to the common trends assumption. While precision is low, the flat effect in both 2018 and 2019 is consistent with an immediate and persistent effect of the policy.

The event study results suggest the effect to be present in April 2018. Most pay gap reports become available close to this date and media attention coincided with the deadline of April 2018. This suggests that wages were adjusted in anticipation of the reports being made public, rather than in response to their publication.

In Table 3 I investigate the policy's impact on alternative outcomes. The ASHE data allow the investigation of one potential route through which employers could increase hourly wages without changing their total costs, the adjustment of hours. In the table, the main specification (column (5) in Table 2) is replicated in the first column. In column (2), the outcome variable is

Figure 4: Treatment effect over time



Notes: Dependent variable is real log hourly earnings. 95% confidence intervals shown based on standard errors clustered by firm. Control variables are age, tenure, industry, skill, public sector and region, interacted with year fixed effects. Figure based on estimates shown in column (5) of Table 11 in Appendix B. Source: ASHE

hours worked per week. I do not detect any change in hours. Correspondingly, when the outcome variable is weekly earnings rather than hourly wages (column (3)), the effect is almost identical to the main specification with hourly wages as the outcome.

In column (4) I use the annual earnings measure available in ASHE. The coefficient is of a similar magnitude to the hourly wage and weekly earnings coefficients in columns (1) and (3), but it is not statistically significant. The standard errors in this column are large, reflecting additional noise in the annual earnings variable. This is partly driven by the fact that when employees change firm, their annual earnings will refer only to the period of the year for which they were working for the firm. As annual earnings is the more noisy measure, most ASHE analysis uses either weekly or hourly earnings, consistent with the approach in this paper.

One possibility is that the change in wages may reflect men at affected firms becoming less likely to be promoted as a result of the policy. To investigate this, in column (5) I use as an outcome a variable which equals one if the worker is in a management occupation, drawn from the SOC code information in ASHE. The coefficient, while positive, is once again statistically insignificant. Finally, in column (6) I inspect whether there is any evidence of gender differences in the probability of leaving a firm resulting from the policy. Here, the outcome variable equals one if an individual is registered with a new firm in the period. The estimate is once again not statistically significantly different to zero.

Table 3: Alternative outcomes

	(1) Hourly wages	(2) Hours	(3) Weekly earn	(4) Annual earn	(5) Management	(6) Leaves firm
>250	0.0110***	0.298**	0.0172***	0.0205	0.00268	-0.00895
	(0.00318)	(0.143)	(0.00439)	(0.0128)	(0.00412)	(0.0100)
>250*Post	-0.0143***	-0.0770	-0.0146***	-0.00637	-0.00616	-0.00339
	(0.00374)	(0.168)	(0.00538)	(0.0139)	(0.00534)	(0.0120)
Post*Female	-0.0000287	-0.111	-0.00101	0.000633	-0.00615	0.00331
	(0.00396)	(0.165)	(0.00547)	(0.0148)	(0.00616)	(0.0127)
>250*Female	-0.0122**	-0.0489	-0.0139**	-0.0175	-0.00771	-0.00123
	(0.00506)	(0.248)	(0.00664)	(0.0200)	(0.00646)	(0.0166)
>250*Female*Post	0.0155**	0.124	0.0159**	0.0130	0.0130	0.00699
	(0.00610)	(0.240)	(0.00799)	(0.0221)	(0.00845)	(0.0193)
N	42,059	42,059	42,059	41,885	42,059	42,059
N workers	12,885	12,885	12,885	12,826	12,885	12885
N firms	6,707	6,707	6,707	6,683	6,707	6,707

Notes: * $p<0.1$, ** $p<0.05$, *** $p<0.01$. Dependent variable is log hourly wages in column (1), hours worked in column (2), log weekly earnings in column (3), log annual earnings in column (4), a management indicator in column (5) and a firm departure dummy in column (6). Standard errors clustered by firm. Control variables are age, tenure, industry, skill, public sector and region, interacted with year fixed effects. All specifications include worker X firm fixed effects. Annual earnings missing for a small number of observations, leading to smaller sample. Source: ASHE

In Appendix I I estimate the effect of the policy on stock market returns, finding no effect.

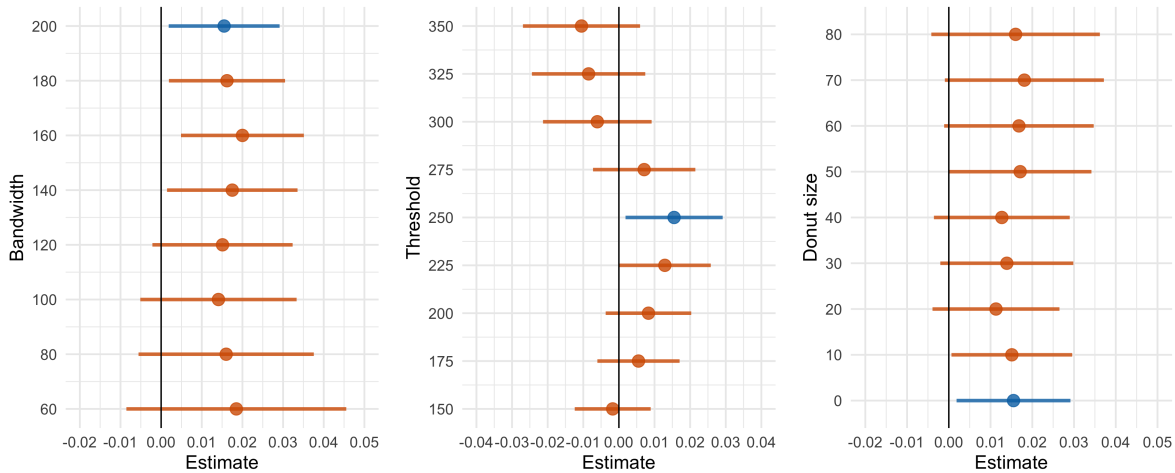
5.2 Robustness

5.2.1 Bandwidths, thresholds and donuts

In this section I inspect the robustness of my main estimated treatment effect, the triple interaction coefficient in column (5) of Table 2. I first test how stable the coefficient is to varying the bandwidth around the threshold. In Figure 5(a), I show the estimated coefficient for a variety of bandwidths. The baseline bandwidth is 200. As the bandwidth gets smaller, the estimates become less precise, but the identifying assumptions become more plausible as the treatment and control workers become more similar. The coefficient is stable for all bandwidths. The estimated coefficient for the smallest bandwidth (60) is 0.019, close to the main baseline estimate of 0.016. The full estimates are shown in Table 12 in Appendix B.

Secondly, I vary the threshold of 250 employees, applying the main specification to a variety of 'psuedo thresholds'. In the spirit of Fisher's randomization inference, the estimated effect is less credible if an effect is found across psuedo thresholds. Figure 5(b) shows that the only significant effect is found at the true threshold of 250 employees, lending credibility to the result. The profile of estimates is consistent with the 'bleeding' of treatment effects, with thresholds

Figure 5: Main robustness tests



(a) Bandwidths (b) Thresholds (c) Donuts

Notes: All panels show main estimated treatment effect and 95% confidence interval for a variety of specifications. All specifications include controls and worker X firm fixed effects. Panel (a) shows estimates for a variety of bandwidths, where a bandwidth of 200 corresponds to the main estimate. Panel (b) shows the estimates for the true employer size threshold of 250 and a set of 'psuedo thresholds'. Panel (c) shows estimates for "donut hole" specifications, where the main specification corresponds to a donut hole of size 0. See Tables 12, 13 and 14 in Appendix B for full estimates. Source: ASHE

close to 250 approaching positive and significant results. The full estimates are shown in Table 13 in Appendix B.

The final panel of Figure 5 contains coefficient estimates where specifications are estimated on data which excludes a mass of workers around the (true) threshold. A "donut hole" of size 10 corresponds to omitting from analysis all individuals at employers with between 245 and 254 employees. As the donut hole becomes larger, the coefficient is estimated less precisely as more observations are excluded, however the figure shows the estimates to be stable for a range of donut sizes. The estimates range from 0.011 to 0.018, close to the baseline estimate of 0.016. This demonstrates that my core results are robust to the exclusion of firms close to the threshold. These estimates are presented in Table 14 in Appendix B.

5.2.2 Other robustness tests

In Table 15 of Appendix B I include three additional robustness tests. In the first (column (1)) I restrict all controls and most importantly employer size to equal their values in the pre-treatment period. This limits concerns of endogenous responses both in the control variables and also in the employer size variable, which determines treatment and control group allocation. The estimated triple interaction coefficient is highly statistically significant and equal to 0.017. This is indistinguishable to the estimated effect in the main specification.

In column (2) of Table 15 I apply a "regression discontinuity-difference in difference" design, which is identical to the standard specification except that I interact a continuous measure of employer size with year fixed effects. The estimated treatment effect is 0.011, close in magnitude to the main estimate, but is statistically insignificant to zero. The inclusion of a continuous measure of firm size reduces the precision with which we can detect a discontinuity around the threshold.

In column (3) of Table 15 I restrict my sample to a balanced panel, including only workers who appear in every year of my estimation period. This reduces the sample size to 2,073 workers, a sixth of the full sample, so the treatment effect is estimated with substantially lower precision. However, the point estimate (0.014) is close to that in the main specification (0.016). This suggests that the effect is not driven by workers moving in and out of the sample.

5.2.3 Bunching around the employment threshold

One concern is that employers respond to the pay gap legislation by changing their firm size.[20] Consider an employer with a particularly high pay gap and 250 employees. If the employer is concerned that having their high pay gap revealed could be damaging, they may find it less costly to lay off a single employee than to adjust their pay gap. In this event, we would expect to see some evidence of bunching just below the 250-employee threshold.

Figure 13 in Appendix B tests for bunching in 2018 data on UK firms. A counterfactual distribution around the threshold is estimated by using a range of data ('CF calc range'), with an inner area excluded ('Excluded bins'). A visual comparison of the counterfactual distribution and the plotted histogram makes clear that there is no excess bunching around the cutoff. In Figure 14 in Appendix B I show the firm size distribution over time. There is no evidence of changes in the distribution of firms around the 250 employee cutoff around the reporting policy introduction in 2018. As adjusting firm size is somewhat costly and difficult in the short run, this will be an interesting outcome to measure in future years.

5.3 Heterogeneity

A disadvantage of the local nature of the empirical strategy applied here is that the small sample size provides relatively little precision. This makes a thorough investigation of heterogeneity challenging. In Table 4 I show estimated interaction effects of the main treatment and five different characteristics. These are age, skill (1-6 scale), years of tenure, whether a worker's wage is covered by a collective agreement and share of women in the industry (2 digit SIC code). In all but one case, the interaction coefficient is statistically insignificantly different to zero. In column (5), we see some evidence that the effect is greater in industries where there are fewer women, though the effect is only marginally significant and the interpretation of this relationship

[20] An alternative interpretation is that firm-size adjustment is an outcome of interest in itself.

is unclear.[21]

Table 4: Heterogeneity

	(1)	(2)	(3)	(4)	(5)
>250*Female*Post	0.0536*	0.0194	0.0152	0.0186**	0.0408**
	(0.0322)	(0.0126)	(0.00960)	(0.00807)	(0.0185)
*Age	-0.000955				
	(0.000745)				
*Skill		-0.00140			
		(0.00396)			
*Tenure			-0.00000529		
			(0.00105)		
*Agreement				-0.00673	
				(0.0141)	
*Fem share					-0.0606*
					(0.0363)
N	42,059	42,059	42,059	42,059	42,059
N workers	12,885	12,885	12,885	12,885	12,885
N firms	6,707	6,707	6,707	6,707	6,707

Notes: * $p<0.1$, ** $p<0.05$, *** $p<0.01$. Dependent variable is real log hourly earnings. Standard errors clustered by firm. Controls and worker X firm fixed effects included in every specification. Control variables are age, tenure, industry, skill, public sector and region, interacted with year fixed effects. Sample constrained to be identical to that of column (5) of Table 2 for all specifications. Source: ASHE

6 Mechanism: Worker preferences over gender pay gaps

In this section, I use newly-gathered survey data to test whether worker preferences over gender pay gaps can explain the effect identified in the main results. As outlined more formally in the conceptual framework in Appendix F, if workers hold a preference against working at a high pay gap employer and the policy increases information on which employers have high pay gaps, employers will have an incentive to narrow their pay gaps. The survey provides qualitative evidence both on how the reporting policy influences the informational environment and on the preferences of workers. Full details on the survey design are given in Appendix H.

6.1 Worker information

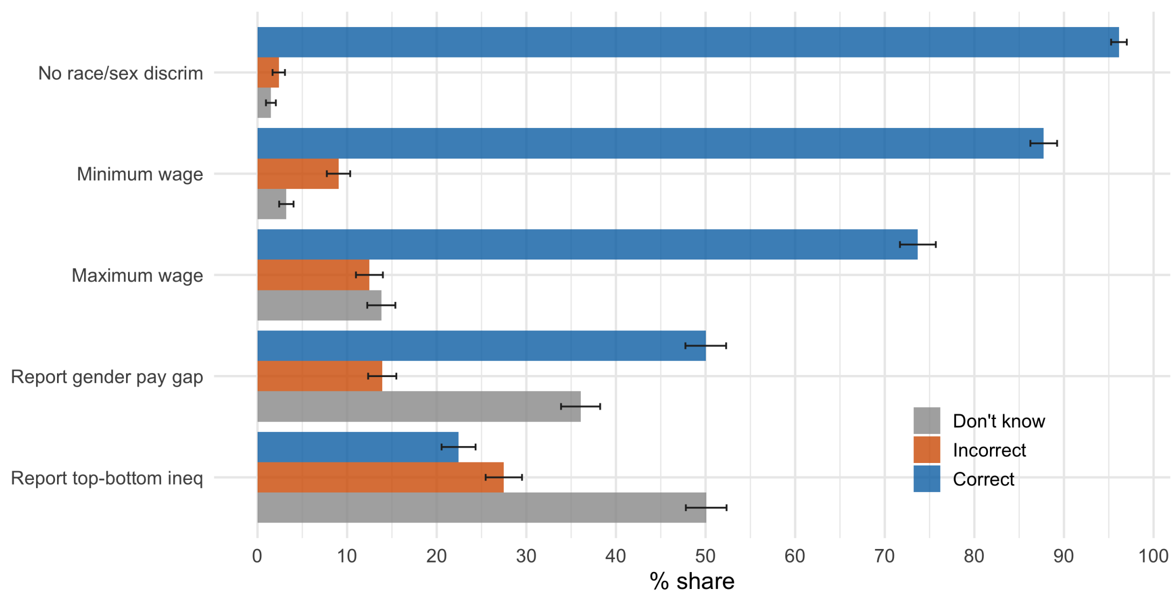
The first issue addressed by the survey data is whether or not workers were informed of the policy. If workers are not aware that pay gap information exists, it is implausible that they

[21] It would be appealing to investigate heterogeneity by the declared pay gap. Unfortunately, the matching of external company-level data to ASHE is not sufficiently high quality to do so.

would react. Figure 6 below shows responses to questions on whether five labor market policies exist. Three of these are real UK policies, for example minimum wage legislation, and two are fictional, for example maximum wage legislation. At this point in the survey, the gender pay gap legislation had not been made salient, with respondents being asked a broad set of questions on working life and views. Responses suggest a strong degree of comprehension among respondents, with a high proportion correctly identifying the best-known policies, such as anti-discrimination legislation and the minimum wage.

In this indirect approach, we can see that 50% of respondents correctly stated that the gender pay gap reporting policy exists. Of the remainder, 36% responded that they were unsure and 14% stated incorrectly that the policy did not exist. Later in the survey respondents are directly asked whether they were aware of the policy before taking the survey, to which 40% responded yes, broadly consistent with the result here.

Figure 6: Share identifying real or fictional policies correctly



Notes: Proportion of respondents giving each type of response when asked whether each of five potential policies exists. Full text of each policy given in Appendix H. 95% confidence intervals shown. Source: Own survey

In Table 16 in Appendix B, I inspect how this varies across types of workers. Awareness is higher among more educated and higher income workers. These patterns are robust to the inclusion of an overall 'Index of labor market policy awareness' variable, which is the average number of correct responses to the other four policies asked about in the indirect question.

Participants were also asked several direct questions on their engagement with the pay gap information. 25% of respondents at covered employers reported that their employer had communicated their pay gap to them,[22] 11% had looked up their employer's pay gap online, and

[22] The pay gap legislation states that employers must communicate the information to their employees directly.

71% had seen a news story about the gender pay gap reports. Most individuals interacted with the information through the media. In Table 5 I show how each of these varies by worker characteristics. While there is little heterogeneity in who has seen the reports in the news, when it comes to looking up their employer's pay gap, men are 7 percentage points less likely to have done so. This is a substantial difference relative to the mean of 11% and is suggestive of strong heterogeneity in workers' interest in their employer's pay gap.

Table 5: Heterogeneity in engagement with pay gap information

	(1) Emp com	(2) Emp lookup	(3) Seen in news
Male	-4.644	-7.313***	-1.058
	(3.741)	(2.468)	(4.150)
Age	-0.337**	-0.250**	-0.296
	(0.169)	(0.108)	(0.184)
Degree	0.301	6.714***	6.896*
	(3.722)	(2.549)	(3.923)
Income	3.676**	2.387**	2.087
	(1.593)	(1.064)	(1.878)
Hours	0.521**	0.0275	0.248
	(0.203)	(0.146)	(0.282)
Mean dep var	25.15	10.77	70.58
N	954	954	954

Notes: * p<0.1, ** p<0.05, *** p<0.01. Outcome variable in (1) is indicator for whether employer communicated pay gap, in (2) is indicator for whether respondent looked up employer's pay gap, (3) is whether respondent saw news reports on policy. Robust standard errors shown in parentheses. Omitted controls are employment status, industry and region. Sample restricted to those at employers covered by policy. Source: Own survey

A final valuable insight from the information section of the survey is the estimated pay gap at participants' employers.[23] On average, participants estimate their employer to have a pay gap of 6.7% towards men. The most-recent complete gender pay gap reports give the average within-firm pay gap as 11.9%. Combining these statistics, on average, participants believe the pay gap to be between a half and two-thirds of the reported value. The difference between mean perceived and actual pay gaps is highly statistically significant. A histogram of estimated pay gaps is given in Appendix B Figure 15. A little over half (56%) of respondents estimate there to be no gap. This demonstrates that individuals on average hold biased beliefs on the magnitude of pay gaps, suggesting there to be scope for informational interventions to change workers' beliefs.

In Table 17 in Appendix B I show substantial gender heterogeneity in the estimated gap. Men estimate the gap to be 2.7 percentage points smaller, which is 40% of the mean estimated gap of 6.7%. Men are also 8.4 percentage points more likely to state there to be no gap.

[23]Participants are given clear instructions on what is meant be the median pay gap.

6.2 Worker preferences

On aggregate, I find that 64% of respondents agree with the statement "it is unfair that earnings differ between men and women". As shown in Figure 16 in Appendix B, compared to other dimensions of earnings inequality, respondents are likely to consider gender pay inequality to be unfair. Table 18 in Appendix B shows men to be 14 percentage points less likely to agree with the aggregate gender pay gap being unfair. The picture is similar when respondents are asked whether employers should seek to reduce their gender pay gap. 62% agree with this on aggregate, and men are 17 percentage points less likely to agree.

Figure 7: Ranking of job attributes

Notes: A rank of 1 means attribute is most important to respondents when choosing a job. Ordered by average rank in pooled sample. Source: Own survey

While attitudes towards aggregate gaps and opinions on employers' normative obligations provide valuable context, more relevant for understanding the estimated effect on wages is whether these gaps affect workers' decisions. To assess this, I use two complementary strategies. In the first, I ask participants to rank eight job attributes in order of each attribute's importance. These include conventional job attributes such as salary and commute time, but also include less familiar attributes such as environmental impact and salary relative to colleagues. This question was positioned before the focus on the gender pay gap policy was made salient to participants.

Each attribute's average rank by gender are shown in Figure 7. Salary is the highest-ranked attribute, with 57% of respondents assigning it their top rank and an average rank of 1.85. This is followed by commute time and hours flexibility. Gender pay gap is on average the lowest-ranked attribute, with 25% of respondents ranking it last.[24] However, the ranking differs by gender with women putting the gender pay gap above both perks and an employer's environmental impact.

Table 6 shows the results of a linear probability model in which the outcome variable is (-) the rank of the gender pay gap attribute. This shows that women rank the gender pay gap 0.73 positions higher than men. There are also strong profiles in age, with younger female workers more likely to value employers with lower gender pay gaps. This is consistent with concerns over career trajectories. A large gender pay gap at an employer in many cases signals a lack of representation of women among the highest earners. As discussed in Goldin (2014), for many this concern is justified.

Interestingly, the age profile for men runs in the opposite direction, and more educated men assign the gender pay gap a lower rank. In Table 19 in Appendix B I show the same patterns hold in an ordered logistic regression model.

Table 6: Heterogeneity in ranking of gender pay gap

	(1) Full	(2) Male	(3) Female
Male	-0.729***	0	0
	(0.107)	(.)	(.)
Age	-0.000635	0.0139**	-0.0122**
	(0.00468)	(0.00706)	(0.00599)
Degree	-0.0819	-0.280**	0.142
	(0.0976)	(0.134)	(0.134)
Income	-0.0305	-0.0806	0.0174
	(0.0417)	(0.0553)	(0.0664)
Hours	0.00741	0.00530	-0.00329
	(0.00573)	(0.00882)	(0.00705)
>250 emp	-0.184	-0.218	-0.151
	(0.114)	(0.176)	(0.149)
Mean dep var	-6.132	-6.526	-5.750
N	1840	619	1221

Notes: * p<0.1, ** p<0.05, *** p<0.01. Outcome variable is (minus) rank position of gender pay gap in job attributes question, where a positive number corresponds to a higher rank. Column (1) is estimated on full sample, column (2) on male participants and column (3) on female participants. Linear regression model in which omitted controls are employment status, industry and region. Source: Own survey

An advantage of the previous approach is that it does not draw particular focus to the gender pay gap, as participants are presented with multiple attributes. A disadvantage is that we are limited to drawing conclusions on relative comparisons across employer attributes rather than

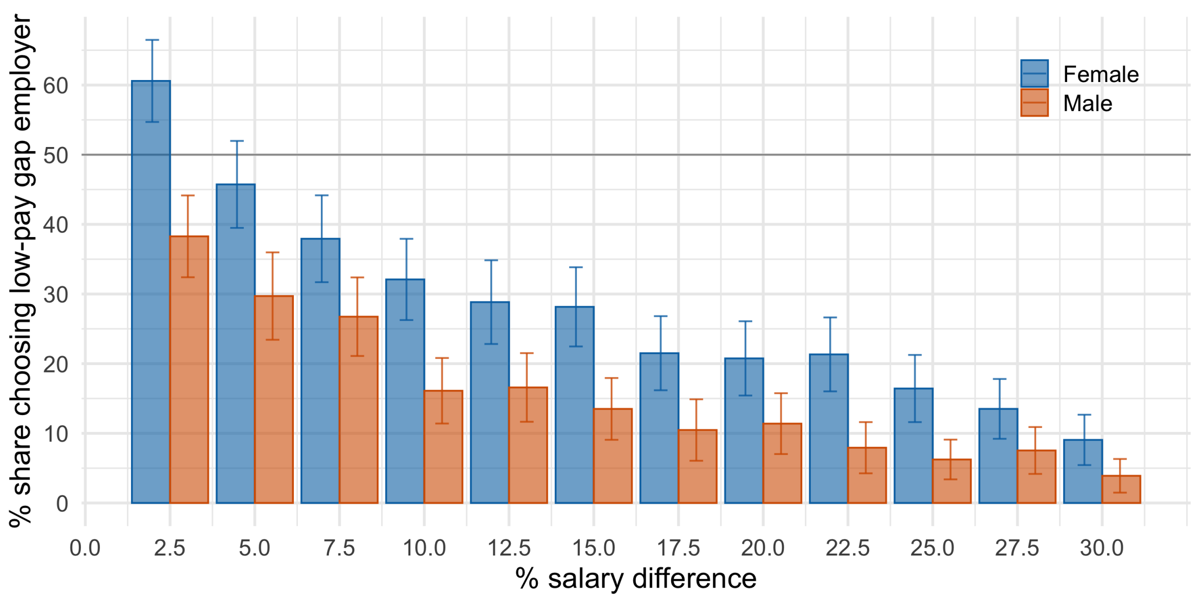
[24]This position is conditional on both own absolute salary and also salary compared to other workers. This means that the gender pay gap information valued here is a narrower signal than it would be in practice.

absolute magnitudes. Given this, I implement a hypothetical choice experiment.

Hypothetical choice experiments have enabled researchers to calculate worker valuations for a variety of job attributes (Mas and Pallais (2017), Datta (2019)). In my application, participants are asked to choose between Job A and Job B, which differ in salary. Participants are told *"The employer offering job A was recently reported as having the highest gender pay gap in their industry. This means that the gap between average male and female wages was higher at this employer than at other employers."*[25] The employer with the greater gender pay gap is associated with a higher salary. The salary difference between the two is varied from 2.5% to 30%, in increments of 2.5 percentage points. Each respondent is asked to choose between the two jobs for three random differences in salary.

The share choosing the lower pay gap employer by gender is plotted in Figure 8. As the salary difference grows, fewer individuals are prepared to sacrifice salary in order to work at the lower pay gap employer. Among women, 61% of respondents would accept a 2.5% lower salary to join the lower pay gap employer. When the salary difference is 15%, 28% of women would choose the low pay gap employer. Clear from the figure is that for any salary difference, women are significantly more likely to choose the low pay gap employer. Inspecting the point at which the 50% share line is crossed, these shares imply a median valuation among men of below 2.5%, and among women of 2.5-5%.

Figure 8: Valuations of employer gender pay gaps



Notes: Share choosing low gender pay gap employer as function of salary difference between low and high pay-gap employer. Source: Own survey.

[25] Contracted hours, location and industry are all drawn from actual responses earlier in the survey. Salaries are set to be close to the self-reported participant salary at the start of the survey. This is to ensure the job choice is as close as possible to real-world decisions.

To obtain a more precise estimate of worker valuations of lower pay gap firms, in Table 7 I show estimated valuations drawn from a logistic regression of employer choice on salary difference. Let $\hat{\alpha}$ and $\hat{\beta}$ represent the estimated intercept and slope coefficient from a logistic regression in which the outcome variable is choosing the low pay gap employer and the sole predictor is the % salary difference associated with that employer. Assuming that valuations for each option are distributed according to the extreme value distribution, the mean and median valuation can be estimated as $-\hat{\alpha}/\hat{\beta}$. Consistent with the above discussion, the estimated aggregate value is estimated as 0.78% of salary and is insignificantly different to zero. For men, the value is negative and insignificantly different to zero, but for women the value is 4.91% of salary.

This represents a large-magnitude valuation for low gender pay gap employers among women. While choices here are hypothetical and do not necessarily entirely reflect real-world decisions, this is consistent with worker responses driving wage adjustment by firms.

Table 7: Coefficients and implied valuations of low gender pay gap employers

	(1) Full	(2) Male	(3) Female
Intercept	0.067	-0.249	0.398***
	(0.087)	(0.158)	(0.104)
Salary diff	-0.086***	-0.102***	-0.081***
	(0.006)	(0.014)	(0.007)
Value	0.780	-2.435	4.909***
	(0.971)	(1.806)	(1.005)

Notes: * p<0.1, ** p<0.05, *** p<0.01. Drawn from logistic regression of employer choice on salary difference as described in main text. Value represents estimated % salary cut individuals would be prepared to take to avoid high gender pay gap employer. Standard errors clustered by participant. Source: Own survey

6.3 Interpretation of information and preferences over pay gaps

There are many potential drivers of the gap in the average earnings of men and the average earnings of women in a particular organization. This leaves ample room for interpretation of pay gap statistics. In this section, I explore how workers interpret the pay gap information, and ask whether this matters substantively for their response to pay gap information.

As discussed in Starr (2014), using qualitative evidence from open-ended questions can help recover a full picture of factors and processes underlying an observed empirical pattern. To gauge worker interpretation of pay gap information, I ask each participant to provide a free text response to “Why do you think the gender pay gap might differ between Employer A and Employer B?", where A and B are employers with different pay gaps in the participant’s industry. The set of the most commonly-occurring words is given in Figure 17 in Appendix B. These are then hand-classified into a set of non-mutually-exclusive classes. The process underlying this is detailed in Appendix H. In this section I work with the six most common classes, which are “Occupations", “Seniority", “Skills", “Hours", “Family" and “Values". 79% of observations provide responses which fall into at least one of these classes. These six responses classes are all factors underlying the pay gap which have been investigated in the literature discussed in Section A.

The results in Table 21 demonstrate substantial heterogeneity in the interpretation of pay gap differences. Male workers are more likely to interpret differences as being due to occupations. Female workers are more likely to attribute them to differences in seniority or family-related differences, such as childcare responsibilities. Workers with more education are 12 percentage points more likely to report seniority as a factor, relative to a mean of 33%.

In Table 8 I show results from a linear regression of an indicator for choosing the low pay gap employer in the hypothetical choice experiment on a set of predictors. These include dummy variables indicating which whether or not an individual lies in each of the six classes. In the first column, I omit the interpretation class variables. We see that men are 12 percentage points less likely to choose the low pay gap employer. There is also a strong age profile, with younger workers more likely to choose the low pay gap employer. This are aligned with the heterogeneity patterns on aggregate views on gender inequality shown in Table in Appendix B.

In column (2) of Table 8 I include each interpretation class as a separate dummy variable. These coefficients show that valuations for the low pay gap employer are lower among those who interpret pay gap differences to be due to gender differences in seniority, skills, hours of work or family commitments. On aggregate then, if workers recognize pay gap differences to be due to several key observable differences between men and women in the workforce, they exhibit less of a preference for low pay gap employers. It is interesting that the inclusion of these variables does not eliminate the patterns shown in column (1) across worker and job characteristics. Interpretation differences as measured here do not explain for example, the difference between male and female responses to pay gap information.

In columns (3) and (5), I present the same regression results separately by gender. These show that for men, skill and hours interpretations lead to lower valuations. For women, ‘family’ interpretations lead to lower valuation and ‘values’ interpretations to higher valuation. Here, ‘values’ corresponds to any interpretation of the differences as reflecting the ethical standpoint of managers or the core values of the organization.

One concern with this approach is that 21% of responses do not fall into these groups and could be driving these patterns. In columns (4) and (6) I replicate columns (3) and (5) but restrict only to those who fall one into the six major interpretation classes listed. The results are qualitatively the same, though with the smaller sample size some of the coefficients become insignificant.

Taken together, this suggests that interpretation of pay gap information matters for worker responses. That factors such as hours matter for worker responses has strong implications for policy design. A narrower, conditional form of pay reporting would limit the scope of interpretation

Table 8: Heterogeneity in probability of choosing low pay gap employer

	(1) Full	(2) Full	(3) Male	(4) Male	(5) Female	(6) Female
Characteristics						
Male	-11.69***	-11.66***	0	0	0	0
	(1.963)	(1.932)	(.)	(.)	(.)	(.)
Age	-0.277***	-0.275***	-0.0372	-0.0735	-0.479***	-0.445***
	(0.0837)	(0.0833)	(0.115)	(0.114)	(0.115)	(0.132)
Degree	-0.783	-0.590	1.036	3.894	-2.367	-2.668
	(1.907)	(1.892)	(2.756)	(3.048)	(2.541)	(2.849)
Income	-1.189	-1.036	-1.766*	-1.679	-1.085	-0.235
	(0.825)	(0.823)	(1.060)	(1.224)	(1.159)	(1.238)
Hours	-0.200*	-0.190*	-0.00252	-0.0887	-0.332**	-0.292*
	(0.104)	(0.104)	(0.148)	(0.156)	(0.141)	(0.153)
>250 emp	-0.874	-0.894	1.718	1.399	-2.623	-1.189
	(2.021)	(2.010)	(2.763)	(2.968)	(2.828)	(3.048)
Interpretation classes						
Occupations		0.0514	-2.670	1.414	0.718	3.372
		(2.012)	(2.731)	(2.837)	(3.026)	(3.168)
Seniority		-3.260*	-3.724	0.561	-3.575	0.756
		(1.918)	(2.493)	(2.724)	(2.833)	(3.294)
Skills		-7.753***	-10.30***	-5.152	-4.116	-0.607
		(2.583)	(3.502)	(3.394)	(3.648)	(4.067)
Hours		-7.976**	-16.13***	-10.83***	0.208	1.967
		(3.181)	(3.225)	(3.570)	(4.995)	(5.077)
Family		-5.642*	0.443	1.709	-9.049*	-8.223*
		(3.271)	(4.154)	(3.917)	(4.859)	(4.904)
Values		2.903	-2.012	4.314	6.611**	11.41***
		(2.328)	(3.335)	(3.528)	(3.148)	(3.726)
Mean dep var	23.71	23.71	16.33	14.13	30.84	30.06
N	5,520	5,520	1,857	1,524	3,663	2,847
In 6 classes				✓		✓

Notes: * p<0.1, ** p<0.05, *** p<0.01. Linear probability model in which outcome variable is indicator for choosing low pay gap employer, scaled by 100 for interpretation. Robust standard errors clustered at individual level shown in parentheses. Omitted controls are salary difference, employment status, region and industry. Columns (4) and (6) omit all individuals with “why pay gap" responses outside of top 6. Results split by gender where indicated. Source: Own survey.

and would likely have different effects to the current policy.

7 Discussion

The key result of this paper is that a gender pay gap reporting policy in the UK led to an increase in female wages relative to those of men. The effect was within workers' employment spells at affected organizations. The effect is then not driven by a reshuffling of highly-paid women into affected employers, but rather was due to changes in individual workers' wages who remained at affected firms before and after the policy's introduction.

Survey evidence is consistent with worker concerns over employer gender pay gaps as a partial mechanism driving wage adjustment. There are significant gender differences in the response to pay gap information. Female workers are less likely to choose to work at employers with high pay gaps, and they draw different conclusions from pay gap reports to men.

There are a number of caveats to the results shown in the current paper. Firstly, the sample used for analysis here is constrained to full-time workers who have a strong attachment to the labor market. Much of the aggregate gender pay gap is driven by differences between part-time and full-time work (Manning and Petrongolo, 2008), so this sample misses a key part of the aggregate gender pay gap. Given this, an investigation into part-time workers would be valuable in future work. Additionally, future work should explore more directly worker flows between employers. The current datasets available in the UK are not well-suited to this type of analysis. One promising potential avenue is to use job search data to identify if employers revealed as having high pay gaps experienced greater outward, and lower inward, job search.

The topic would also benefit from further theoretical innovations, integrating existing ideas in personnel and labor economics to better capture how pay gap reports affects wage setting within the firm. A promising direction may be to draw more closely on the literature on regulation, as in Laffont and Tirole (1993). In doing so, researchers would be more able to inform policymakers on optimal policy design, rather than estimating reduced form effects of existing policies.

In practice, labor is differentiated within gender, firms can outsource workers, and payment can be shifted towards in-kind benefits. There exists anecdotal evidence of non-pay adjustments being used to alter pay gaps without changing overall remuneration. In future work it would be valuable to explore additional datasets that would allow the investigation of substitution towards in-kind benefits and changes in the workforce at affected firms. An alternative form of shifting is to move wage bills to different points in the year. For example, over the period on which the gender wage gap is assessed, workers of a certain gender could be temporarily paid more.

One important component of an assessment of the aggregate welfare effect of the policy is whether these changes in the wage structure have any effect on firm productivity. If pay gap reporting induces changes in wages such that individual wages better-track each worker's productivity, the policy could result in improvements in productivity. On the other hand, given that the declared pay gaps do not account for differences across workers, firms wanting to reduce their pay gap may find it optimal to enhance the difference between workers' wages and their productivity. This would distort the allocation of labor within the firm and lead to lower productivity.

Recall that on aggregate, I find that the policy reduced wages at affected employers. At least in standard models, this is consistent with firms experiencing lower productivity. In the future, theoretical and empirical contributions in this direction would be valuable for policymakers to understand the normative implications of the policy.

References

S. Andersen, J. Marx, K. M. Nielsen, and L. Vesterlund. Gender Differences in Negotiation: Evidence from Real Estate Transactions. Working Paper 27318, National Bureau of Economic Research, June 2020. URL `http://www.nber.org/papers/w27318`. Series: Working Paper Series.

S. Athey and G. W. Imbens. Identification and Inference in Nonlinear Difference-in-Differences Models. *Econometrica*, 74(2):431–497, 2006. ISSN 0012-9682. URL `http://www.jstor.org/stable/3598807`. Publisher: [Wiley, Econometric Society].

I. Ayres and P. Siegelman. Race and Gender Discrimination in Bargaining for a New Car. *The American Economic Review*, 85(3):304–321, 1995. ISSN 0002-8282. URL `http://www.jstor.org/stable/2118176`. Publisher: American Economic Association.

M. Baker, Y. Halberstam, K. Kroft, A. Mas, and D. Messacar. Pay Transparency and the Gender Gap. Technical Report w25834, National Bureau of Economic Research, Cambridge, MA, May 2019. URL `http://www.nber.org/papers/w25834.pdf`.

S. Beggs, S. Cardell, and J. Hausman. Assessing the potential demand for electric cars. *Journal of Econometrics*, 17(1):1–19, Sept. 1981. ISSN 03044076. doi: 10.1016/0304-4076(81)90056-7. URL `https://linkinghub.elsevier.com/retrieve/pii/0304407681900567`.

B. Bell and S. Machin. Minimum Wages and Firm Value. *Journal of Labor Economics*, 36 (1):159–195, Jan. 2018. ISSN 0734-306X, 1537-5307. doi: 10.1086/693870. URL `https://www.journals.uchicago.edu/doi/10.1086/693870`.

M. Bennedsen, E. Simintzi, M. Tsoutsoura, and D. Wolfenzon. Do Firms Respond to Gender Pay Gap Transparency? Technical Report w25435, National Bureau of Economic Research, Cambridge, MA, Jan. 2019. URL `http://www.nber.org/papers/w25435.pdf`.

M. Bertrand, E. Duflo, and S. Mullainathan. How Much Should We Trust Differences-In-Differences Estimates?*. *The Quarterly Journal of Economics*, 119(1):249–275, Feb. 2004. ISSN 0033-5533. doi: 10.1162/003355304772839588. URL `https://doi.org/10.1162/003355304772839588`.

M. Bertrand, C. Goldin, and L. F. Katz. Dynamics of the Gender Gap for Young Professionals in the Financial and Corporate Sectors. *American Economic Journal: Applied Economics*, 2 (3):228–255, July 2010. ISSN 1945-7782, 1945-7790. doi: 10.1257/app.2.3.228. URL `http://pubs.aeaweb.org/doi/10.1257/app.2.3.228`.

D. Bird. Methodology for the 2004 Annual Survey of Hours and Earnings. page 8, 2004.

E. Breza, S. Kaur, and Y. Shamdasani. The Morale Effects of Pay Inequality*. *The Quarterly Journal of Economics*, 133(2):611–663, May 2018. ISSN 0033-5533, 1531-4650. doi: 10.1093/qje/qjx041. URL https://academic.oup.com/qje/article/133/2/611/4430649.

S. G. Bronars and D. R. Deere. Union Representation Elelctions and Firm Profitability. *Industrial Relations: A Journal of Economy and Society*, 29(1):15–37, 1990. ISSN 1468-232X. doi: 10.1111/j.1468-232X.1990.tb00739.x. URL http://onlinelibrary.wiley.com/doi/abs/10.1111/j.1468-232X.1990.tb00739.x. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1468-232X.1990.tb00739.x.

D. Card, A. Mas, E. Moretti, and E. Saez. Inequality at Work: The Effect of Peer Salaries on Job Satisfaction. *American Economic Review*, 102(6):2981–3003, Oct. 2012. ISSN 0002-8282. doi: 10.1257/aer.102.6.2981. URL http://pubs.aeaweb.org/doi/10.1257/aer.102.6.2981.

D. Card, A. R. Cardoso, and P. Kline. Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women. *The Quarterly Journal of Economics*, 131(2):633–686, May 2016. ISSN 0033-5533, 1531-4650. doi: 10.1093/qje/qjv038. URL https://academic.oup.com/qje/article-lookup/doi/10.1093/qje/qjv038.

M. Castillo, R. Petrie, M. Torero, and L. Vesterlund. Gender differences in bargaining outcomes: A field experiment on discrimination. *Journal of Public Economics*, 99:35–48, Mar. 2013. ISSN 0047-2727. doi: 10.1016/j.jpubeco.2012.12.006. URL http://www.sciencedirect.com/science/article/pii/S0047272713000042.

A. K. Chatterji and M. W. Toffel. How firms respond to being rated. *Strategic Management Journal*, pages n/a–n/a, 2010. ISSN 01432095, 10970266. doi: 10.1002/smj.840. URL http://doi.wiley.com/10.1002/smj.840.

Z. Cullen and R. Perez-Truglia. How Much Does Your Boss Make? Harvard Business School Working Paper(19-013):107, 2018.

Z. B. Cullen and B. Pakzad-Hurson. Equilibrium Effects of Pay Transparency in a Simple Labor Market: Extended Abstract. *Working Paper*, pages 193–193, 2019. doi: 10.1145/3328526.3329645. URL http://dl.acm.org/citation.cfm?doid=3328526.3329645.

Z. B. Cullen and R. Perez-Truglia. Privacy Norms and the Diffusion of Information. *Harvard Business School Working Paper*, (69), Nov. 2019.

N. Datta. Willing to Pay for Security: A Discrete Choice Experiment to Analyse Labour Supply Preferences. *CEP Discussion Paper*, (No 1632), 2019. URL http://cep.lse.ac.uk/pubs/download/dp1632.pdf.

E. Duchini, S. Simion, and A. Turrell. Pay Transparency and Cracks in the Glass Ceiling. *arXiv:2006.16099 [econ, q-fin]*, June 2020. URL http://arxiv.org/abs/2006.16099. arXiv: 2006.16099.

W. Espeland and M. Sauder. Rankings and Reactivity: How Public Measures Recreate Social Worlds. *American Journal of Sociology*, 113(1):1–40, July 2007. ISSN 0002-9602, 1537-5390. doi: 10.1086/517897. URL https://www.journals.uchicago.edu/doi/10.1086/517897.

U. Gneezy, M. Niederle, and A. Rustichini. Performance in Competitive Environments: Gender Differences. *The Quarterly Journal of Economics*, 118(3):1049–1074, Aug. 2003. ISSN 0033-5533, 1531-4650. doi: 10.1162/00335530360698496. URL https://academic.oup.com/qje/article-lookup/doi/10.1162/00335530360698496.

C. Goldin. A Grand Gender Convergence: Its Last Chapter. *American Economic Review*, 104(4):1091–1119, Apr. 2014. ISSN 0002-8282. doi: 10.1257/aer.104.4.1091. URL http://pubs.aeaweb.org/doi/10.1257/aer.104.4.1091.

C. Goldin, L. F. Katz, and I. Kuziemko. The Homecoming of American College Women: The Reversal of the College Gender Gap. *Journal of Economic Perspectives*, 20(4):133–156, Dec. 2006. ISSN 0895-3309. doi: 10.1257/jep.20.4.133. URL http://www.aeaweb.org/articles?id=10.1257/jep.20.4.133.

C. Goldin, S. P. Kerr, C. Olivetti, and E. Barth. The Expanding Gender Earnings Gap: Evidence from the LEHD-2000 Census. *American Economic Review*, 107(5):110–114, May 2017. ISSN 0002-8282. doi: 10.1257/aer.p20171065. URL http://pubs.aeaweb.org/doi/10.1257/aer.p20171065.

A. Gulyas, S. Seitz, and S. Sinha. Does Pay Transparency Affect the Gender Wage Gap? Evidence from Austria. page 42, 2020.

F. Guvenen, G. Kaplan, and J. Song. How Risky Are Recessions for Top Earners? *American Economic Review*, 104(5):148–153, May 2014. ISSN 0002-8282. doi: 10.1257/aer.104.5.148. URL http://pubs.aeaweb.org/doi/10.1257/aer.104.5.148.

B. Hansen and D. McNichols. Information and the Persistence of the Gender Wage Gap: Early Evidence from California's Salary History Ban. Working Paper 27054, National Bureau of Economic Research, Apr. 2020. URL http://www.nber.org/papers/w27054. Series: Working Paper Series.

D. Hedblom, B. Hickman, and J. List. Toward an Understanding of Corporate Social Responsibility: Theory and Field Experimental Evidence. Technical Report w26222, National Bureau of Economic Research, Cambridge, MA, Sept. 2019. URL http://www.nber.org/papers/w26222.pdf.

T. Independent. The worst gender pay gap offenders in each sector have been revealed. *The Independent*, Apr. 2018. URL https://www.independent.co.uk/news/business/news/gender-pay-gap-worst-offenders-in-each-sector-revealed-as-reporting-deadline-passes-a8290566.html.

H. Kleven, C. Landais, and J. E. Søgaard. Children and Gender Inequality: Evidence from Denmark. *American Economic Journal: Applied Economics*, 11(4):181–209, Oct. 2019. ISSN 1945-7782, 1945-7790. doi: 10.1257/app.20180010. URL https://pubs.aeaweb.org/doi/10.1257/app.20180010.

H. Kleven, C. Landais, J. Posch, A. Steinhauer, and J. Zweimüller. Do Family Policies Reduce Gender Inequality? Evidence from 60 Years of Policy Experimentation. page 90, 2020.

H. J. Kleven and M. Waseem. Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence from Pakistan. *The Quarterly Journal of Economics*, 128 (2):669–723, May 2013. ISSN 0033-5533. doi: 10.1093/qje/qjt004. URL http://academic.oup.com/qje/article/128/2/669/1943151. Publisher: Oxford Academic.

J.-J. Laffont and J. Tirole. *A Theory of Incentives in Procurement and Regulation.* MIT Press, 1993. ISBN 978-0-262-12174-3. Google-Books-ID: 4iH4Z2wbNqAC.

D. S. Lee and A. Mas. Long-Run Impacts of Unions on Firms: New Evidence from Financial Markets, 1961–1999 *. *The Quarterly Journal of Economics*, 127(1):333–378, Feb. 2012. ISSN 0033-5533, 1531-4650. doi: 10.1093/qje/qjr058. URL https://academic.oup.com/qje/article-lookup/doi/10.1093/qje/qjr058.

A. Leibbrandt and J. A. List. Do Women Avoid Salary Negotiations? Evidence from a Large-Scale Natural Field Experiment. *Management Science*, 61(9):2016–2024, Sept. 2015. ISSN 0025-1909, 1526-5501. doi: 10.1287/mnsc.2014.1994. URL http://pubsonline.informs.org/doi/10.1287/mnsc.2014.1994.

M. Luca. Reviews, Reputation, and Revenue: The Case of Yelp.Com. *SSRN Electronic Journal*, 2011. ISSN 1556-5068. doi: 10.2139/ssrn.1928601. URL http://www.ssrn.com/abstract=1928601.

A. C. MacKinlay. Event Studies in Economics and Finance. *Journal of Economic Literature*, 35 (1):13–39, 1997. ISSN 0022-0515. URL http://www.jstor.org/stable/2729691. Publisher: American Economic Association.

A. Manning and B. Petrongolo. The Part-Time Pay Penalty for Women in Britain*. *The Economic Journal*, 118(526):F28–F51, 2008. ISSN 1468-0297. doi: 10.1111/j.1468-0297.2007.02115.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1468-0297.2007.02115.x.

N. Marriott. Nigel Marriott's Blog, 2020. URL https://marriott-stats.com/nigels-blog/.

A. Mas. Does Disclosure affect CEO Pay Setting? Evidence from the Passage of the 1934 Securities and Exchange Act. *Working Paper*, Mar. 2016.

A. Mas. Does Transparency Lead to Pay Compression? *Journal of Political Economy*, 125(5): 1683–1721, Oct. 2017. ISSN 0022-3808. doi: 10.1086/693137. URL http://www.journals.uchicago.edu/doi/abs/10.1086/693137. Publisher: The University of Chicago Press.

A. Mas and A. Pallais. Valuing Alternative Work Arrangements. *American Economic Review*, 107(12):3722–3759, Dec. 2017. ISSN 0002-8282. doi: 10.1257/aer.20161500. URL http://pubs.aeaweb.org/doi/10.1257/aer.20161500.

M. Niederle and L. Vesterlund. Do Women Shy Away From Competition? Do Men Compete Too Much? *The Quarterly Journal of Economics*, 122(3):1067–1101, Aug. 2007. ISSN 0033-5533. doi: 10.1162/qjec.122.3.1067. URL http://academic.oup.com/qje/article/122/3/1067/1879500. Publisher: Oxford Academic.

Parliament. Oral evidence: The effectiveness of gender pay gap reporting, HC 2240, May 2019.

B. Petrongolo. The gender gap in employment and wages. *Nature Human Behaviour*, 3(4):316–318, Apr. 2019. ISSN 2397-3374. doi: 10.1038/s41562-019-0558-x. URL http://www.nature.com/articles/s41562-019-0558-x.

Pew. The narrowing, but persistent, gender gap in pay, 2018. URL https://www.pewresearch.org/fact-tank/2019/03/22/gender-pay-gap-facts/.

R. S. Ruback and M. B. Zimmerman. Unionization and Profitability: Evidence from the Capital Market. *Journal of Political Economy*, 92(6):1134–1157, Dec. 1984. ISSN 0022-3808, 1537-534X. doi: 10.1086/261278. URL https://www.journals.uchicago.edu/doi/10.1086/261278.

M. A. Starr. Qualitative and Mixed-Methods Research in Economics: Surprising Growth, Promising Future. *Journal of Economic Surveys*, 28(2):238–264, 2014. ISSN 1467-6419. doi: https://doi.org/10.1111/joes.12004. URL http://onlinelibrary.wiley.com/doi/abs/10.1111/joes.12004. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/joes.12004.

A Extended literature review

A.1 Effects of pay transparency

This project relates to a wide recent literature in Labor Economics on the impacts of pay transparency on a variety of outcomes. The earliest empirical paper in this literature is Card et al. (2012). This paper leverages the fact that salaries for public sector workers in the state of California are public. By experimentally varying the salience of this information to employees of The University of California at Berkeley, the authors demonstrate that the salary information leads to lower worker satisfaction. Lower earners become dissatisfied and are more likely to search for other jobs. This result is replicated in Breza et al. (2018) in an Indian manufacturing plant setting. This latter paper also finds that when workers perceive pay differences to be driven by differences in productivity across workers, there is no effect of pay inequality on worker morale and output.

Using the same California public sector disclosure policy, Mas (2017) shows that the highest earners in municipal jobs experience lower wages following public disclosure of wages, due to public aversion to exorbitant salaries. Particularly of interest to the current paper is the finding that exposure to media had a significant effect in restraining top wages. In contrast to these results, Mas (2016) uses data from the Great Depression to demonstrate that disclosure of executive compensation led to an increase in average CEO pay relative to that of other executives.

Using a field experiment, Cullen and Perez-Truglia (2018) show that when workers perceive their peers to be higher paid than themselves, they exert lower effort, leading to lower output and retention. However, when managers are perceived to have higher salary the effect is the opposite. This is consistent with the above results of Breza et al. (2018), that the reasons underlying pay gaps are important when it comes to worker reactions to new information. Particularly relevant to the current study is the observation that individuals may be tolerant of gender pay gaps if the pay gap is primarily driven by vertical differences, rather than horizontal differences. This paper and the later work in Cullen and Perez-Truglia (2019) show that in general, workers are not well-informed on the salary of their colleagues, but there are no gender differences in how accurate beliefs are.

Cullen and Pakzad-Hurson (2019) extend the above papers by running an experiment in an online marketplace. Through a dynamic wage bargaining model, they show that transparency lowers wages and raises profits. Most relevant for the current paper, they demonstrate that full transparency benefits women by reducing men's advantage in communicating wages.

A recent paper which focuses particularly on the effect of transparency on the gender pay gap is Baker et al. (2019). This paper examines the effect of public sector salary disclosure laws on university faculty salaries in Canada. They present robust evidence that transparency laws reduced the gender pay gap, driven primarily by institutions where the faculty are unionized.

Another related research area covers 'salary history' questions, in which employers ask about previously earned salaries. It has been suggested that this activity perpetuates past gender discrimination. Recent evidence on US city and state policy shifts suggests that banning salary history questions increases female earnings relative to male earnings (Hansen and McNichols, 2020).

While these papers demonstrate a link between transparency and pay gaps it is difficult to directly extrapolate their results to pay gap disclosure policies, since these papers tend to investigate cases in which all individual-level data is released. This is a different type of transparency to that induced by the policy in this paper. Many of the empirical settings are also often restricted to public sector workers, where pay is substantially more rigid than in the private sector.

With these caveats in mind, a general theme that emerges is that workers are poorly informed of colleagues' wages, and that transparency can have substantial effects on wages when workers perceive there to be an element of unfairness. Worker responses to course measures of gender pay gaps will likely then depend on whether they reflect an injustice or unfairness from workers' perspectives.

A.2 Causes of the gender pay gap

While this paper does not seek to shed new light on the causes of the gender pay gap, when considering the potential impact of this policy it is important to consider the factors underlying pay gaps. This helps build intuition over potential mechanisms and increases understanding of the information that workers and consumers might draw from gender pay gap reports. It also clarifies the factors omitted from the aggregated gender pay gap statistics.

There exists an expansive literature on the proximate and root causes of gender pay gaps, much of which is reviewed succinctly in Petrongolo (2019). As alluded to previously, controlling for education actually widens the gender pay gap in most countries, as women now obtain more education than men (Goldin et al., 2006). Occupational segregation remains an important factor but not overwhelmingly so, with Goldin (2014) showing that absorbing the effect of detailed occupations decreases the gap by no more than one-third.

An increasing focus has been put on the gender earnings gap over the life cycle. Using US administrative data, Goldin et al. (2017) find that the gender earnings gap expands by 33.7 log points from ages 26 to 39 among college graduates. This is due in equal parts to differential earnings trajectories within employer and to differential mobility between establishments. The widening of the gap is substantially larger for college graduates relative to high school graduates, a result which we will return to later in the paper.

Gaps in hours worked have been shown to be important, with UK evidence given in Manning and Petrongolo (2008). Part-time work tends to be associated with an hourly wage penalty and more women are found in part-time positions. Goldin (2014) extends this finding, arguing that the extent to which hours and job continuity matter for wages is the dominant cause of the gender pay gap. In occupations where pay is strongly nonlinear in hours, meaning that hours are complementary, the gender pay gap is significantly higher. In these occupations, the asymmetric childcare responsibilities observed around family formation (Kleven et al., 2019) are particularly costly and long lasting in terms of wage growth. One influential study demonstrated this to be the case for MBA graduates of top US business schools (Bertrand et al., 2010), with the male earnings advantage reaching almost 60 log points ten years after MBA completion.

Another relevant strand of literature covers differences in bargaining power. Card et al. (2016) show that women receive lower firm-specific pay premiums than men, which the authors attribute to differences in bargaining power. Related to this, Leibbrandt and List (2015) demonstrate that women are less likely to engage in salary negotiations, but that when it is explicit that wages are negotiable, differences disappear. This builds on a substantial experimental literature on gender gaps in performance in competitive environments (Gneezy et al., 2003) and preferences for such environments (Niederle and Vesterlund, 2007). There also exists an active literature on bargaining in consumer markets, such as over taxi fares (Castillo et al., 2013), automobiles (Ayres and Siegelman, 1995) and real estate transactions (Andersen et al., 2020).

A.3 Gender pay gap disclosure policies

There exists a pair of papers investigating impacts of targeted gender pay gap disclosure policies on workers and firms. Bennedsen et al. (2019) find that a pay gap disclosure policy in Denmark successfully reduced the pay gap within affected firms, increasing female earnings by 2 percentage points relative to those of men. The effect was driven primarily by lower wages for men. They also find that the policy had a negative effect on firm productivity. Gulyas et al. (2020) study a similar law in Austria, finding that the policy led to an increase in the retention rate of workers.

The policies investigated in these studies are substantively different to the UK policy. In both Denmark and Austria, pay gaps were not made publicly available, but rather were just available to employees within the firm. Pay gap reports were also made specific to occupations, in contrast to the cruder unconditional nature of the UK's reporting policy.

A.4 Firm ratings and Corporate Social Responsibility

There finally exists a literature on the broad theme of organizations' responses to being publicly rated. The literature is best reviewed in Chatterji and Toffel (2010). Espeland and Sauder (2007) provide a theoretical contribution, emphasizing how the act of being observed or measured can induce organizational change independent of any external incentives. This is an intriguing perspective, but one which is difficult to include in a standard model of the firm.

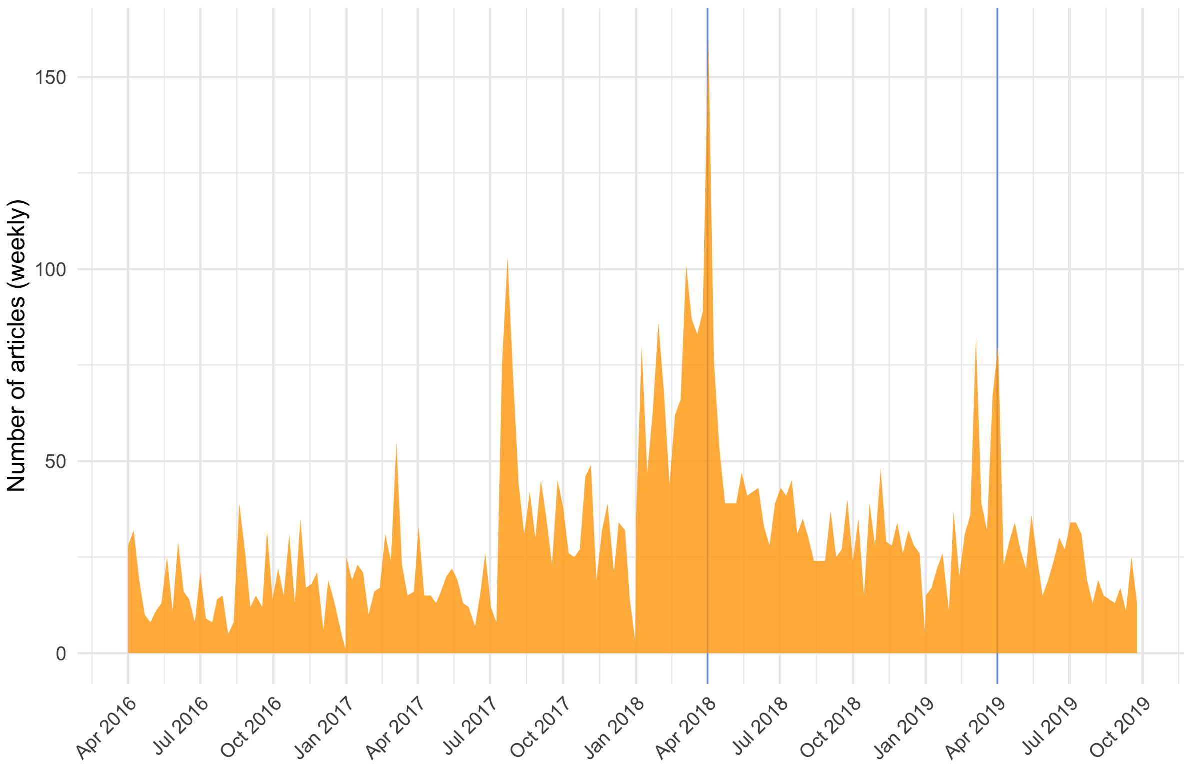
There are a number of papers on the impact of Corporate Social Responsibility (CSR) metrics, which gender pay gap statistics can be considered a form of. A recent related paper in this field is Hedblom et al. (2019). They combine theory and a field experiment to focus on the labor supply-side implication of CSR ratings. They find that when a firm advertises one particular form of CSR (working with charities), they see an uptick in applications from prospective employees.

These new employees are more productive than the firm would be able to recruit otherwise. The study does not allow wages to vary, but this provides further evidence that workers care about non-wage aspects of their firm, which is important when considering the labor supply effects of the policy in this paper.

The evolution of online review systems has stimulated a number of papers on how consumers respond to firm ratings. Luca (2011) shows that consumers respond strongly to restaurant reviews on Yelp.com. The extent to which this extends to gender pay gap disclosure naturally depends on whether consumers place value on, or hold firms responsible for, their gender pay gaps. That the reports of prominent consumer-facing companies have received strong media attention is suggestive that consumers hold some interest in the pay gaps of firms.

B Further tables and figures

Figure 9: Weekly news reporting on the gender pay gap



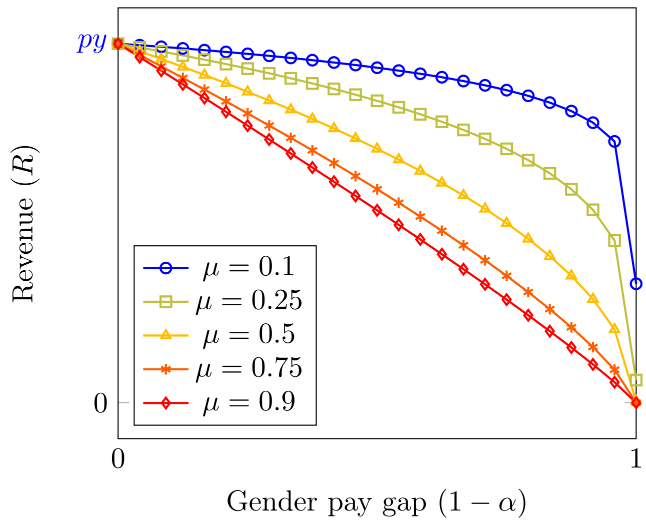
Notes: Number of UK print and online news articles including the term ‘gender pay gap’ each week. Weeks containing pay gap reporting deadlines indicated with blue vertical lines. Drawn from 483 UK national and local news sources. (accessed via AWN Newsbank March 31st 2020).

Table 9: Firm summary statistics by employer size

Variable	Control	Treatment	Overall
Operating profit	2454 (203)	3885 (466)	2894 (2)
Working capital	5389 (442)	11040 (1003)	7078 (4)
Pre-tax profits	2209 (220)	4461 (677)	2903 (2)
Net tangible assets	18603 (570)	29466 (1348)	21845 (5)
Turnover	36688 (676)	54339 (1305)	42124 (6)
Number of employees	191.6 (0.3)	293.7 (0.5)	222.2 (0.0)

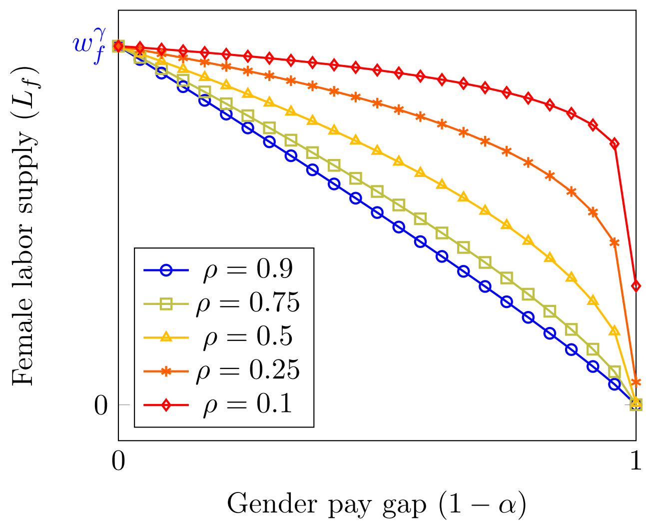
Notes: Year 2017. Standard errors in parentheses. Contains full population of firms. Source: FAME

Figure 10: Consumer preferences, revenue and pay gaps



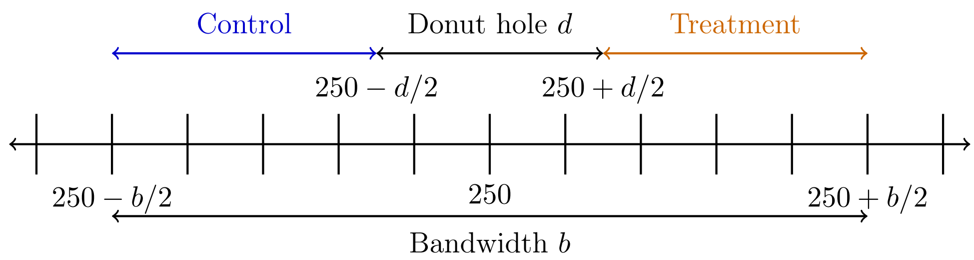
Notes: This figure shows how revenue is affected by the gender pay gap. As μ increases, consumers become more sensitive to the firm's gender pay gap.

Figure 11: Female labor supply and pay gaps



Notes: This figure shows how female labor supply responds to the gender pay gap. As ρ increases, workers become more sensitive to the firm's gender pay gap and reduce their labor supply.

Figure 12: Treatment and control group selection with “donut hole"



Notes: Illustration of treatment and control group selection around threshold of 250 employees. Number line represents number of employees. Bandwidth b refers to maximum number of employees in treatment group minus minimum number of employees in control group. Donut hole d omitted from analysis.

Table 10: Triple difference-in-difference specifications by gender

	Male	Female	All
>250	0.011***	-0.0030	0.011***
	(0.0035)	(0.0050)	(0.0032)
>250*Post	-0.014***	0.00085	-0.014***
	(0.0047)	(0.0053)	(0.0037)
Post*Female			-0.000029
			(0.0040)
>250*Female			-0.012**
			(0.0051)
>250*Female*Post			0.016**
			(0.0061)
N	25,617	16,442	42,059
N workers	7,689	5,196	12,885
N firms	4,777	3,672	6,707

Notes: * p<0.1, ** p<0.05, *** p<0.01. Dependent variable is real log hourly earnings. Standard errors clustered by firm. Controls and worker X firm fixed effects included in every specification. Control variables are age, tenure, industry, skill, public sector and region, interacted with year fixed effects. Source: ASHE

Table 11: Generalized difference-in-difference estimates

	(1)	(2)	(3)	(4)	(5)
>250*Female	0.014	0.0023	-0.017	-0.014	-0.014*
	(0.025)	(0.016)	(0.015)	(0.0093)	(0.0077)
>250*Female*2013	0.027	0.011	0.017	-0.0014	-0.0014
	(0.036)	(0.023)	(0.019)	(0.013)	(0.011)
>250*Female*2014	0.0099	0.018	0.025	0.0054	0.0053
	(0.033)	(0.021)	(0.017)	(0.012)	(0.0097)
>250*Female*2015	-0.0062	-0.013	0.010	-0.00027	-0.00036
	(0.030)	(0.019)	(0.015)	(0.010)	(0.0087)
>250*Female*2016	0.031	0.0094	0.012	0.0070	0.0069
	(0.020)	(0.014)	(0.011)	(0.0083)	(0.0069)
>250*Female*2017	0	0	0	0	0
	(.)	(.)	(.)	(.)	(.)
>250*Female*2018	0.023	0.026	0.036**	0.017*	0.017**
	(0.027)	(0.018)	(0.014)	(0.0093)	(0.0077)
>250*Female*2019	0.015	0.028	0.039**	0.019	0.018**
	(0.032)	(0.022)	(0.018)	(0.011)	(0.0094)
N	42,059	42,059	42,059	42,059	42,059
N workers	12,885	12,885	12,885	12,885	12,885
N firms	6,707	6,707	6,707	6,707	6,707
Controls		✓	✓	✓	✓
Firm FE			✓	✓	
Worker FE				✓	
Worker X Firm					✓

Notes: * p<0.1, ** p<0.05, *** p<0.01. Dependent variable is real log hourly earnings. Standard errors clustered by firm. Controls and worker X firm fixed effects included in every specification. Control variables are age, tenure, industry, skill, public sector and region, interacted with year fixed effects. Sample constrained to be identical to that of column (5) for all specifications. Source: ASHE

Table 12: Alternative bandwidths

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
>250*Female*Post	0.0185	0.0160	0.0141	0.0151*	0.0175**	0.0200***	0.0162**	0.0155**
	(0.0138)	(0.0110)	(0.00981)	(0.00880)	(0.00819)	(0.00771)	(0.00729)	(0.00696)
Bandwidth	60	80	100	120	140	160	180	200
N	10030	14176	18458	22955	27510	32278	37104	42059
N workers	3539	4829	6104	7397	8723	10108	11449	12885
N firms	1941	2601	3240	3897	4566	5253	5985	6707

Notes: * p<0.1, ** p<0.05, *** p<0.01. Coefficients from main specification where bandwidth containing estimation sample allowed to vary. The baseline estimate corresponds to a bandwidth of 200 (column (8)). All specifications include controls and worker X firm fixed effects. Source: ASHE

Table 13: Alternative thresholds

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
>thresh*Fem*Post	-0.00174	0.00550	0.00830	0.0129*	0.0155**	0.00710	-0.00606	-0.00851	-0.0105
	(0.00544)	(0.00589)	(0.00613)	(0.00658)	(0.00696)	(0.00733)	(0.00778)	(0.00812)	(0.00839)
Threshold	150	175	200	225	250	275	300	325	350
N	80332	64706	54383	47964	42059	37481	33400	30035	27378
N workers	23646	19230	16314	14558	12885	11543	10386	9417	8625
N firms	16325	12180	9597	8008	6707	5695	4859	4185	3653

Notes: * $p<0.1$, ** $p<0.05$, *** $p<0.01$. Coefficients from main specification where number of employees threshold identifying affected firms is allowed to vary. The baseline estimate corresponds to a threshold of 250 (column (5)). All specifications include controls and worker X firm fixed effects. Source: ASHE

Table 14: "Donut hole" specifications

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
>250*Female*Post	0.0155**	0.0151**	0.0113	0.0139*	0.0127	0.0171**	0.0168*	0.0181*	0.0160
	(0.00696)	(0.00738)	(0.00776)	(0.00813)	(0.00830)	(0.00871)	(0.00914)	(0.00972)	(0.0103)
Donut size	0	10	20	30	40	50	60	70	80
N	42059	39516	37346	34894	32480	30198	28061	25778	23583
N workers	12885	12319	11828	11201	10575	9939	9321	8662	8020
N firms	6707	6515	6332	6050	5780	5488	5183	4856	4526

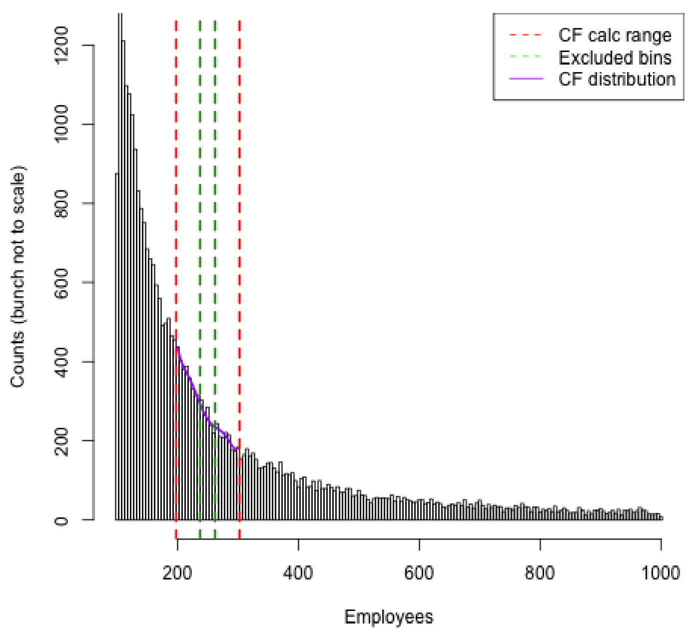
Notes: * $p<0.1$, ** $p<0.05$, *** $p<0.01$. Coefficients from main specification where "donut hole" of individuals around employer size threshold omitted from analysis. The baseline estimate corresponds to a donut hole size of 0 (column (1)). All specifications include controls and worker X firm fixed effects. Source: ASHE

Table 15: Additional robustness tests

	(1) Pre controls	(2) RD-DID	(3) Balanced
>250	0	0	0.026***
	(.)	(.)	(0.0057)
>250*Post	-0.014***	-0.0023	-0.016**
	(0.0052)	(0.0060)	(0.0066)
Post*Female	-0.00043	0.00087	0.0076
	(0.0049)	(0.0048)	(0.0077)
>250*Female	0	0	-0.012
	(.)	(.)	(0.0097)
>250*Female*Post	0.017**	0.011	0.014
	(0.0075)	(0.0073)	(0.011)
N	37,820	42,059	9,777
N workers	10,172	12,885	2,073
N firms	4,608	6,707	1,641

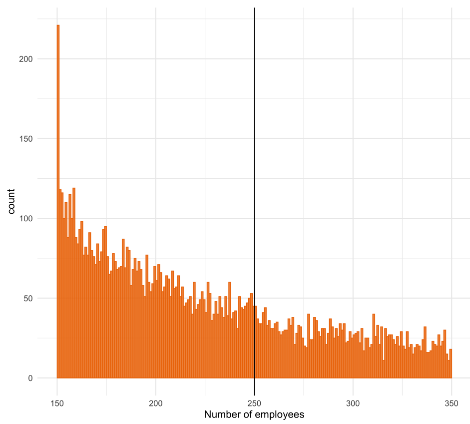
Notes: * $p<0.1$, ** $p<0.05$, *** $p<0.01$. Dependent variable is real log hourly earnings. Standard errors clustered by firm. Controls and worker X firm fixed effects included in every specification. Control variables are age, tenure, industry, skill, public sector and region, interacted with year fixed effects. Maximum possible sample constrained to be identical to that of column (5) of Table 2 for all specifications. See main text for description. Source: ASHE

Figure 13: Test for bunching

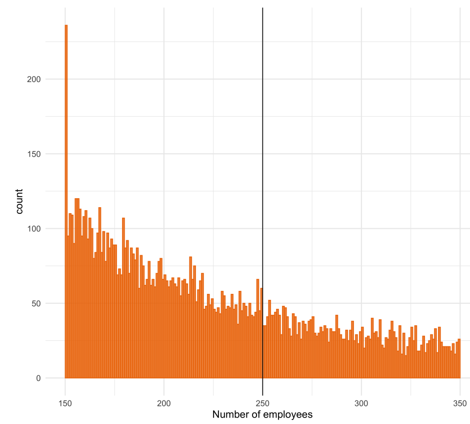


Notes: Visual test based on (Kleven and Waseem, 2013). Data from year 2018. Histogram is true data. Purple line is estimated counterfactual distribution around 250 employee cutoff, using data outside of area between dashed green lines. Source: FAME

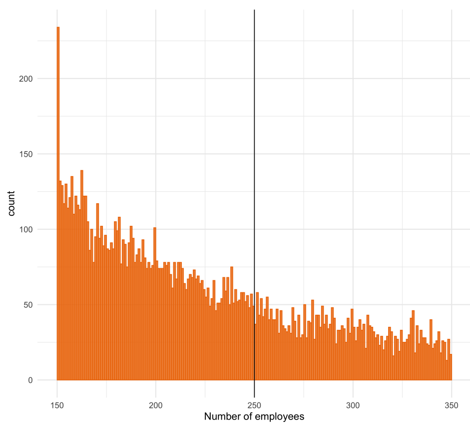
Figure 14: Histograms of firm size over time



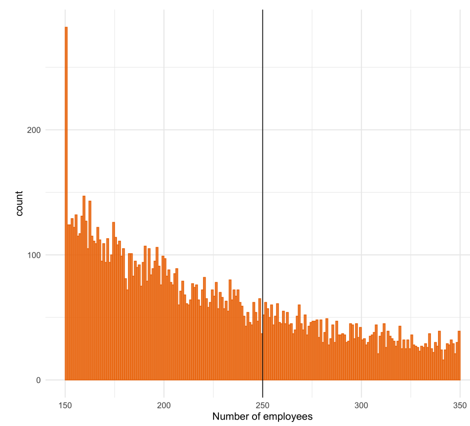
(a) 2014



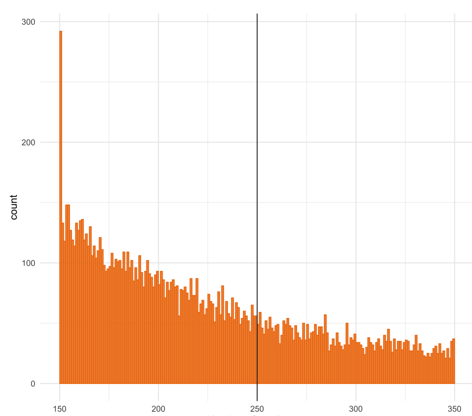
(b) 2015



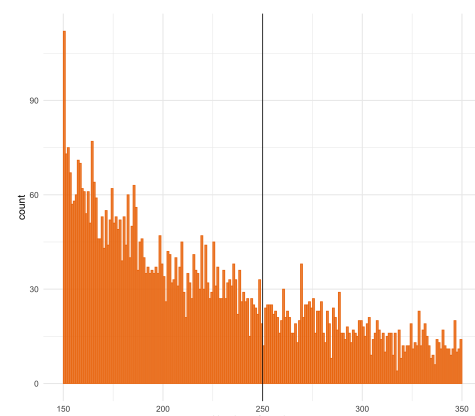
(c) 2016



(d) 2017



(e) 2018



(f) 2019

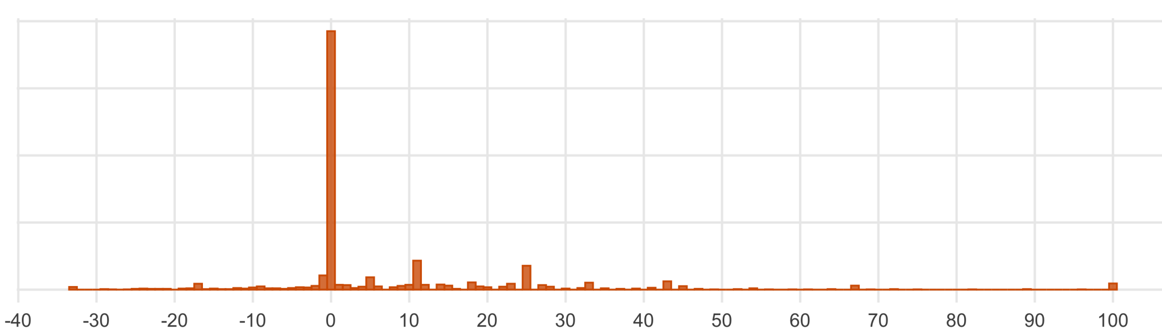
Notes: Data for 2019 is noisier as employment information for many firms not yet updated. Source: FAME

Table 16: Heterogeneity in awareness of policy

	(1) Direct	(2) Indirect	(3) Direct	(4) Indirect
Male	-3.397	-3.481	-3.662	-3.714
	(3.087)	(3.209)	(3.103)	(3.227)
Age	-0.0503	-0.0755	-0.0571	-0.0815
	(0.128)	(0.138)	(0.129)	(0.139)
Degree	14.46***	14.78***	14.16***	14.52***
	(2.971)	(3.032)	(2.984)	(3.052)
Income	2.837**	2.468*	2.670**	2.321*
	(1.265)	(1.313)	(1.268)	(1.311)
Hours	0.188	0.00873	0.201	0.0196
	(0.163)	(0.172)	(0.162)	(0.170)
>250 emp	5.922*	3.826	5.817*	3.734
	(3.160)	(3.366)	(3.167)	(3.372)
ILA			12.40*	10.87
			(7.390)	(7.770)
Mean dep var	39.63	50.03	39.63	50.03
N	1840	1840	1840	1840

Notes: * p<0.1, ** p<0.05, *** p<0.01. Outcome variable is whether demonstrated awareness of gender pay gap reporting policy in direct question (columns (1) and (3)) and indirect question (columns (2) and (4)). Outcome variable multiplied by 100 for ease of interpretation. Robust standard errors shown in parentheses. Model is linear probability model in which omitted controls are industry, employment status and region. Columns (3) and (4) control for “Index of Labor Market Policy Awareness" (ILA), which is average number of correct responses to other four hypothetical policy questions. Source: Own survey

Figure 15: Estimated own-employer pay gaps



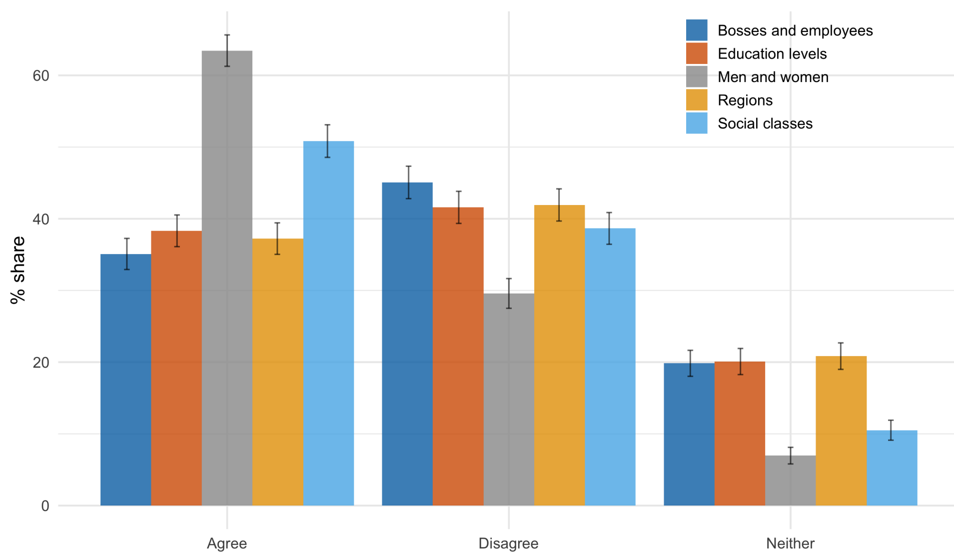
Notes: Estimated % pay gap in favor of male workers at own employer. Source: Own survey

Table 17: Heterogeneity in estimated own-employer pay gaps

	(1) Gap size	(2) No gap
Male	-2.668**	8.364***
	(1.330)	(3.227)
Age	0.00811	0.123
	(0.0556)	(0.139)
Degree	1.477	-3.988
	(0.929)	(3.021)
Income	0.339	-3.349**
	(0.429)	(1.318)
Hours	0.0997	0.0479
	(0.0863)	(0.178)
>250 emp	-0.488	-1.605
	(1.292)	(3.425)
Mean dep var	6.673	56.09
N	1840	1840

Notes: * p<0.1, ** p<0.05, *** p<0.01. Outcome variable in (1) is the estimated own-employer pay gap in favor of men and in (2) is indicator for an estimated gap of zero. Robust standard errors shown in parentheses. Omitted controls are employment status, industry and region. Source: Own survey

Figure 16: Concerns over dimensions of earnings inequality



Notes: Participants asked whether they agree with statement "it is unfair that earnings differ between...". Source: Own survey

Table 18: Heterogeneity in concerns over gender pay inequality

	(1)	(2)
Male	-13.88***	-16.66***
	(3.190)	(3.170)
Age	-0.256*	-0.241*
	(0.138)	(0.137)
Degree	3.351	4.192
	(2.977)	(2.940)
Income	-0.959	-2.367*
	(1.333)	(1.272)
Hours	0.192	0.105
	(0.173)	(0.168)
>250 emp	8.127**	7.171**
	(3.341)	(3.286)
Mean dep var	63.45	62.12
N	1840	1840

Notes: * p<0.1, ** p<0.05, *** p<0.01. Outcome variable in (1) is indicator for whether respondent finds it unfair that earnings differ between men and women, and in (2) is indicator for whether participant agrees that employers should close their gender gaps. Robust standard errors shown in parentheses. Model is linear probability model in which omitted controls are employment status, industry and region. Source: Own survey

Table 19: Heterogeneity in ranking of gender pay gap (Rank ordered logistic regression)

	(1)	(2)
Commute	-0.875***	-0.623***
	(0.0716)	(0.0785)
Env impact	-2.843***	-2.765***
	(0.0888)	(0.100)
GPG	-2.895***	-2.485***
	(0.0901)	(0.0961)
Hours flex	-1.009***	-0.712***
	(0.0770)	(0.0894)
Perks	-2.518***	-2.548***
	(0.0828)	(0.0886)
Rel salary	-2.507***	-2.412***
	(0.0836)	(0.0918)
WFH	-2.404***	-2.320***
	(0.0927)	(0.105)
Commute × Male		-0.547***
		(0.153)
Env impact × Male		-0.234
		(0.189)
GPG × Male		-0.897***
		(0.192)
Hours flex × Male		-0.629***
		(0.164)
Perks × Male		0.0124
		(0.178)
Rel salary × Male		-0.257
		(0.179)
WFH × Male		-0.232
		(0.196)
N	14720	14720

Notes: * p<0.1, ** p<0.05, *** p<0.01. Model is rank ordered logistic regression (Beggs et al. (1981)). Standard errors clustered at participant level. Source: Own survey

Figure 17: Common reasons given for gender pay gap differences across employers

Notes: Most common words in responses to "Why do you think the gender pay gap might differ between Employer A and Employer B?", where A and B are employers with different pay gaps in the respondent's industry. Word size reflects frequency. Stemming performed and stop words removed. Source: Own survey

C Legal Background

The Equality Act 2010 Section 78 allowed the UK government to request pay gap information from large firms and public sector organizations. The 6 points from the act are replicated below:[26]

1. *Regulations may require employers to publish information relating to the pay of employees for the purpose of showing whether, by reference to factors of such description as is prescribed, there are differences in the pay of male and female employees.*
2. *This section does not apply to—*
   - *(a) an employer who has fewer than 250 employees;*
   - *(b) a person specified in Schedule 19;*
   - *(c) a government department or part of the armed forces not specified in that Schedule.*
3. *The regulations may prescribe—*
   - *(a) descriptions of employer;*
   - *(b) descriptions of employee*
   - *(c) how to calculate the number of employees that an employer has;*
   - *(d) descriptions of information;*
   - *(e) the time at which information is to be published;*
   - *(f) the form and manner in which it is to be published.*
4. *Regulations under subsection (3)(e) may not require an employer, after the first publication of information, to publish information more frequently than at intervals of 12 months.*
5. *The regulations may make provision for a failure to comply with the regulations—*
   - *(a) to be an offence punishable on summary conviction by a fine not exceeding level 5 on the standard scale;*
   - *(b) to be enforced, otherwise than as an offence, by such means as are prescribed.*
6. *The reference to a failure to comply with the regulations includes a reference to a failure by a person acting on behalf of an employer.*

The legislation asserts that the UK government would be able to introduce gender pay gap reporting requirements on employers with over 250 employees. However, the precise policies were sidelined until a consultation in 2015, which led to a draft text "Equality Act 2010 (Gender Pay

[26]Source: `http://www.legislation.gov.uk/ukpga/2010/15/section/78`, accessed 16th April 2020 under Open Government Licence.

Gap Information) Regulations 2016" being released for further consultation on 12th February 2016.[27] After consultation, on the 6th of April 2017 a new draft statutory instrument name "The Equality Act 2010 (Gender Pay Gap Information) Regulations 2017" came into force.[28] The regulations set out in detail exactly what must be submitted be employers, including formulae demonstrating how to calculate each of the required statistics.

The regulations assert that the following statistics must be reported both to the government and to directly to employees:

- The difference in mean pay between male and female employees (expressed as a percentage of mean male pay
- The difference in median pay between male and female employees (expressed as a percentage of median male pay
- The difference in bonus pay between male and female employees (expressed as a percentage of mean male pay, as well as the proportion of male and female relevant employees who received bonus pay
- The numbers of male and female employees employed by the relevant employer on the relevant date in quartile pay bands

The regulation contains crucial details on how to construct the hourly pay variable. Firstly, it is specified that the "pay period" is typically the period in respect of which the employer pays the employee basic pay, which in the UK is typically fortnightly or monthly. The hourly rate of pay is then the total pay during that pay period, divided by a multiplier which converts pay to weekly pay, then further divided by weekly hours worked.

Where bonuses are paid out, these are scaled according to the period for which the bonus is related. For example, if an annual bonus is paid out during the pay period, this is divided by approximately 52 to convert to a weekly rate. Hours are either 'typical' working hours, where workers are on a fixed hours contract, or average hours worked over the previous 12 weeks for those not on a fixed hours contract. The definition of hourly wage here is designed to be consistent with that found in the ASHE dataset.

For bonus pay statistics, the relevant period is the 12 months ending with the snapshot date. This implies that Christmas bonuses for example, common in the UK, will be included in the bonus pay statistics but not the hourly pay statistics.

The regulations also offer the following statement on enforcement:

[27] Source: `https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/504398/GPG_consultation_v8.pdf`, accessed 1st April 2020 under Open Government Licence.

[28] Source: `https://www.legislation.gov.uk/ukdsi/2017/9780111152010`, accessed 17th April 2020 under Open Government Licence.

> *Failure to comply with an obligation imposed by these Regulations constitutes an 'unlawful act' within the meaning of section 34 of the Equality Act 2006 (c. 3), which empowers the Equality and Human Rights Commission to take enforcement action.*

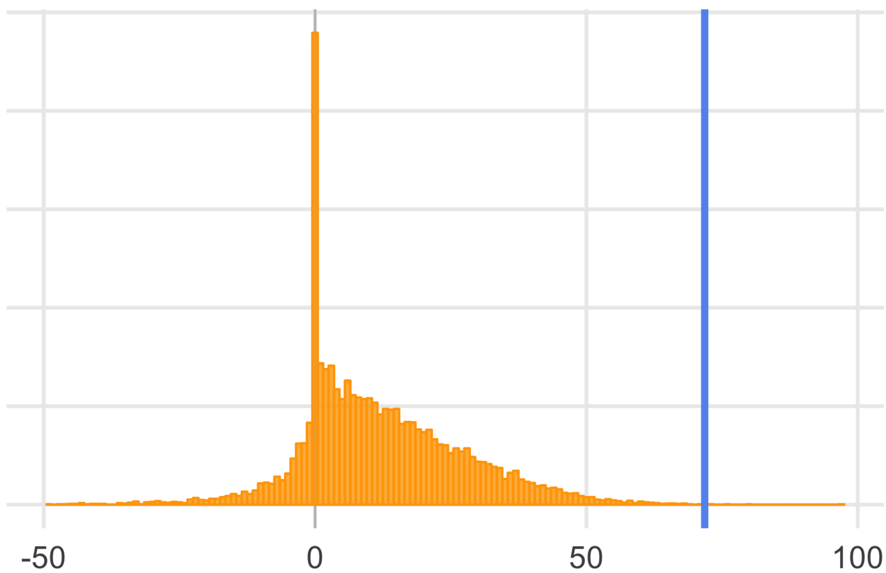
As stated in the main text, there is no evidence that enforcement has moved beyond sending reminder letters to employers who fail to submit a pay gap report. The regulations state also that employers write also make and send a signed statement to confirm that the information provided is accurate. This must be signed by a director, partner or senior official.

D Ryanair case study

In this section, further detail is provided on the experience of a single firm, focusing on the actions they took around the pay gap declaration and the media reaction to their report. The chosen firm is low-cost airline Ryanair, which represents a particularly interesting case.

Figure 18 shows the distribution of median pay gap reports for the first snapshot date of April 2017. Ryanair has a high pay gap in favor of men, with male median wages 71.8% higher than those of median women. This is the 7th highest pay gap across all reporting companies.

Figure 18: Ryanair's gender pay gap



Notes: Median gender pay gaps in hourly pay in 2018 reports. Positive gaps correspond to higher male wages. Gaps less than 50 omitted. Ryanair position highlighted by vertical line. Source: Gender Pay Gap Service

As Ryanair is a well-known consumer facing company, this high pay gap received ample media attention. They appeared in many "worst offender" articles, such as the Independent's "Gender pay gap: worst offenders in each sector revealed as reporting deadline passes" (Independent, 2018). Likely in anticipation of the media response, Ryanair chose to release a narrative statement on their pay gap[29]. In this, they defend their high pay gap, firstly by pointing out the extent of occupational segregation in their industry:

> *Like all airlines, our gender pay in the UK is materially affected by the relatively low number of female pilots in the airline industry*

[29] `https://investor.ryanair.com/wp-content/uploads/2019/04/RYANAIR-GENDER-PAY-GAP-2018.pdf`, accessed 16th April 2020

Indeed, in their report they state that they have 689 male pilots relative to 15 female pilots, and 234 male cabin crew relative to 464 female cabin crew. As pilots are well-paid, this drives their quartile statistics, with men occupying 99.6% of the top-earning quartile. Secondly, Ryanair points out that many of their middle-earning positions are not in the UK:

> *In Ryanair's case our management and administration are based largely in Ireland, so almost all of our UK based colleagues are pilots or cabin crew*

Finally, Ryanair makes clear that within positions, wages are set by collective agreement:

> *All of our UK pilots and cabin crew are covered by negotiated collective agreements, under which our female pilots and cabin crew are paid the same basic salary and the same variable pay rates as their male colleagues.*

The Ryanair case is interesting as it is an extreme example of the gap between high pay gap statistics, notions of unequal pay for equal work and broader concepts of discrimination. The historically unionized nature of the airline industry means that there is little discretion on pay within position, meaning that the company's high pay gap can be clearly attributed to the differentiation of jobs between genders. As the skills and training of pilots and cabin crew differ substantially, it is implausible to expect much movement between these two groups. Furthermore, a significant share of pilots received their training in the military, and this group are predominantly male. This in part drives strong gender differences in labor supply of pilots to the firm.

E Gender pay gap reporting policies outside of the UK

The United States has experienced a fierce debate around gender pay gap reporting. Under the Obama administration, the Office of Management and Budget (OMB) approved a proposal from the US Equal Employment Opportunity Commission (EEOC) to collect pay gap data from employers. The policy met strong criticism from business groups, with the US Chamber of Commerce arguing that the law would be a huge burden for business and would be ineffective at narrowing the pay gap.[30] The Obama-era decision was later reversed by the Trump administration and since then has been contested in the courts. In 2019 in the midst of much uncertainty, the EEOC did collect pay gap statistics from firms, but as of April 2020 it is unclear whether this pay data will be used and if it will be collected again in the future.

With progress stalling at the national level, several states have attempted to implement their own version of the law. The state of California was almost successful in this with SB 1284. As with the federal legislation, this was fiercely opposed by business groups, most notably the California Chamber of Commerce, who labelled the legislation a "job killer". As was the case at the national level, the campaign was effective and the legislation was shelved.

Proponents of pay gap transparency policies have been more successful in Europe, with pay gap reporting requirements now found in nine European countries.[31] There is ample heterogeneity in the details of how these policies are administered. Unlike the UK legislation, in the majority of cases we do not see the public release of pay gap data. Instead, only employees within an organization are given the right to request the information. In some countries, statistics are based not on aggregates but rather on pay gaps within "jobs", typically defined by a detailed occupation code. The detail and openness of reports are constrained by concerns over privacy. This in part explains the lack of public access to pay gap reports in several settings. Outside of Europe there are few examples of gender pay gap transparency legislation, with notable exceptions in Australia, India and Japan.

[30] See for example the letter from the Chamber of Commerce to OMB director here: `https://www.uschamber.com/sites/default/files/eeo-1_coalition_request_for_review.pdf`

[31] These are Austria, Belgium, Denmark, Finland, France, Germany, Norway, Sweden and the UK.

F Conceptual framework

In this section I introduce a simple partial equilibrium model which is used to clarify the mechanisms through which gender pay gap reporting can affect workers' wages. The goal of this exercise is not to derive parameters to be estimated empirically, but rather to establish how pay gap reporting can be embedded into a standard model of the firm. The model also motivates the direction of the survey analyzed in Section 6, in which I investigate one mechanism included in the model.

F.1 Simple model of the firm

Firms produce output y using two imperfectly-substitutable inputs, male labor L_m and female labor L_f. The separation of labor by gender in the production function is designed to capture the fact that men and women tend to hold different types of roles within firms. Firms produce output y with constant returns to scale Cobb-Douglas production technology:

$$y = AL_m^{\beta} L_f^{(1-\beta)} \tag{4}$$

with $0 < \beta < 1$ and TFP given by A, both of which are exogenous. Other inputs such as capital are excluded from the model for simplicity.

I assume that $\beta > \frac{1}{2}$. This is motivated by the fact that empirically, men tend to be found in higher-paid occupations and are more likely to hold senior positions, which suggests higher observed productivity. It is not motivated by any assertion that within-occupation or seniority, male workers are more productive.

Firms are price takers in the product market. The firm is free to scale their output as they see fit, without influencing prices. Their revenue R is determined by exogenous price p and endogenous output y follows:

$$R = py \tag{5}$$

The firm has monopsony power in the labor market and hence is a wage setter, facing upward-sloping labor supply curves as follows:

$$L_f = w_f^{\gamma} \tag{6}$$

$$L_m = w_m^{\gamma} \tag{7}$$

Here, $w_m \geq 0$ and $w_f \geq 0$ are wages for male and female behavior respectively, and γ is the (compensated and uncompensated) labor supply elasticity. We assume $\gamma > 0$ and $\gamma < 1$, meaning that workers have a labor supply elasticity of between 0 and 1. Income effects are zero, which can be motivated by a quasilinear utility function for workers over leisure and consumption. The elasticity is assumed to be constant across the two types of labor.[32]

Solving the firm's optimization problem gives the following simple relationship for male and female wages as functions of exogenous parameters:

$$\left(\frac{w_f}{w_m}\right)^{(1+\gamma)} = \left(\frac{1-\beta}{\beta}\right) \tag{8}$$

The ratio between male and female wages is determined by their relative productivities and the labor supply elasticity γ.

Finally, we introduce the following α parameter, which is the ratio of female to male wages:

$$\alpha = \frac{w_f}{w_m} \tag{9}$$

This parameter is useful as (1-α) is the typical definition of the gender pay gap. If wages for women are 10% lower than wages for men, $\alpha = 0.9$. As $w_f \geq 0$ we have that $\alpha \geq 0$.

In this simple model, we see that:

$$\alpha = \left(\frac{1-\beta}{\beta}\right)^{\frac{1}{1+\gamma}} \tag{10}$$

The assumption $\beta > \frac{1}{2}$ implies that $\alpha < 1$ and $(1-\alpha) > 0$, so there is a gender pay gap in favor of men. Note that to keep the model tractable and to focus on the coming integration of pay gap preferences on the consumer demand and labor supply side, we model the gender pay gap as being entirely due to productivity differences. This is abstracting from many factors which have been shown play a part in explaining the gender pay gap in the literature examined in Section A.

F.2 Gender pay gaps on the demand side

In this section we introduce consumer preferences over pay gaps by including a parameter which allows the pay gap to directly influence firm revenue. To simplify exposition, we assume that realized pay gaps will remain weakly in favor of men ($\alpha \leq 1$). Consumer adjustment must not be so strong that firms pay women more than they do men.

Revenue is modified as follows:

$$R = \alpha^{\mu} p y \tag{11}$$

with $0 \leq \mu \leq 1$.

[32]Empirically, it is typically found that female workers have higher elasticities than men. In future work it would be interesting to explore how pay gap reporting interacts with differences in labor supply curves between men and women. However, to focus on the primary mechanisms in the model we do not explore that possibility here.

Term α is defined as previously and is decreasing in the pay gap. Term μ reflects sensitivity to the gender pay gap among consumers.[33] When $\mu = 0$, revenue is unaffected by the gender pay gap. As μ increases, consumers exhibit stronger negative preferences for a gender pay gap. Parameter μ can therefore be thought of as a simple way of modeling consumers' exhibited preferences for high-pay gap firms. This is shown graphically in Figure 10 in Appendix B. For all $\mu < 1$, revenue falls exponentially in the pay gap, meaning that larger pay gaps are strongly disfavored relative to small pay gaps. If $\mu = 1$, revenue falls linearly in the pay gap for a given price and output. As will be shown below, the introduction of this parameter will affect the wage-setting behavior of the firm.

One interpretation for the inclusion of α in revenue R is that the effective price that firms can charge is now lower if they have a non-zero pay gap. For an individual firm, this is a reasonable assumption. If consumers dislike pay gaps, having a pay gap means that the price charged has to fall to maintain the same level of demand.

F.3 Gender pay gaps on the labor supply side

How might pay gaps affect workers' labor supply? One approach would be to attempt to model how pay gap information affects worker bargaining power. Most directly, if a declared pay gap indicates unequal pay for equal work, the threat of legal action may induce firms to offer pay increases.[34] However, in this setting and particularly given the large size of the employers who have to report, the pay gap reports do not provide substantial information on how well paid an individual worker is relative to their colleagues in comparable roles. Given this, it is unlikely that any direct claim of unequal pay for equal work could be made. Recall that the pay gap reports provide no information on absolute levels of pay, only relative, so it is also difficult to argue that workers become more informed on outside options as a result of the pay gap reports.

Rather than changes to bargaining power, the mechanism highlighted here is a direct worker preference for low pay gap employers. This can be motivated by the evidence presented in previous literature in Section A demonstrating that workers care directly about fairness, and we will explore this further in Section 6.

The labor supply curve for women is modified as follows:

$$L_f = \alpha^{\rho} w_f^{\gamma} \tag{12}$$

with $0 \leq \rho \leq 1$.

Here, ρ functions analogously to μ on the consumer side. For a given wage w_f and pay gap α, higher values of ρ increase the labor supply of female workers. This is plotted in Figure 11 in

[33] To be clear that μ does not necessarily directly reflect consumer preferences but rather a combination of information and preferences, when referring to this parameter the terms "exhibited preference", "sensitivity" or "adjustment" are used. This point will be returned to later in the section when we consider the policy introduction.

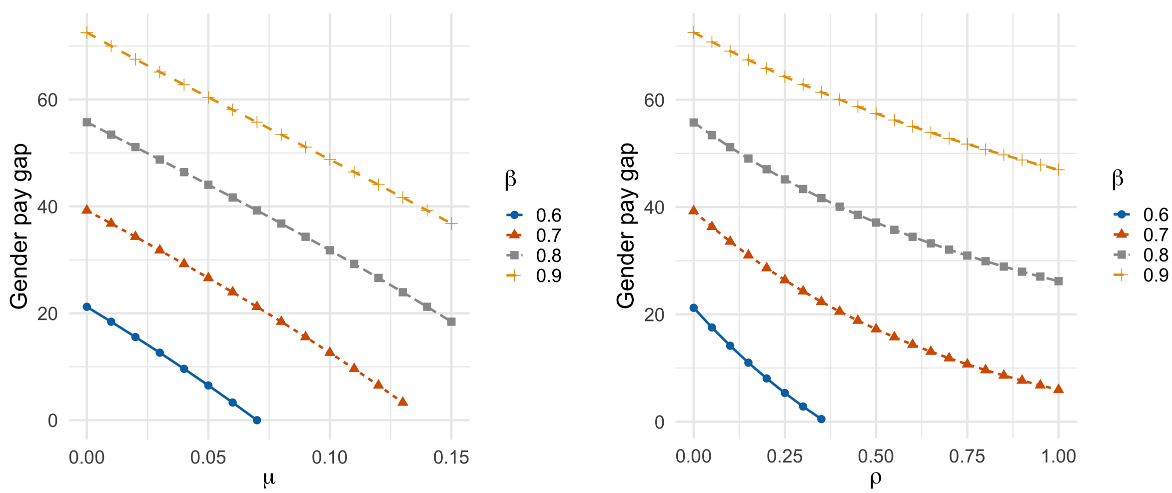
[34] The most famous case of this is Lilly Ledbetter, who famously successfully took her employer to court after learning through an anonymous note that she was underpaid relative to male colleagues at Goodyear.

Appendix B. Again, it is useful to consider ρ as a combination of information and preferences.

F.4 The impact of pay gap reporting

We now consider how pay gap reporting could affect the gender pay gap, once the basic model has been augmented with consumer and labor supply preferences above. In Figure 19 I show the pay gap emerging from firm optimization for a number of parameters. In panel (a), consumer exhibited preferences over pay gaps are introduced and in panel (b) worker exhibited preferences over pay gaps are introduced. The negative slopes shown in each panel of the figure demonstrate that increased sensitivity from consumers ($\uparrow \mu$) or workers ($\uparrow \rho$) leads to falls in the gender pay gap. Firms respond either in order to maintain revenue (panel (a)) or to continue to attract workers (panel (b)).

Figure 19: Pay gaps under consumer and labor supply pressure



(a) Consumers

(b) labor supply

Notes: Optimal gender pay gap for different values of μ, ρ and β. γ fixed at 0.7 and p at 1 throughout.

Let us formalize the decomposition of μ and ρ into an information term and a preference term:

$$\mu = \delta \nu_\mu \tag{13}$$

where δ represents how informed consumers are of pay gaps and ν_μ represents preferences. The same parameterization can be given for worker exhibited preferences:

$$\rho = \delta \nu_\rho \tag{14}$$

The effect of moving to pay gap reporting can be thought of as a shift from $\delta = 0$ to $\delta > 0$.

Consumers and workers become more informed as a result of the policy. Whether or not this will impact wages depends on the ν preference parameters above. If $\nu_\mu > 0$, the introduction of reporting represents a shift from $\mu = 0$ to $\mu > 0$ in Figure 19, and hence a reduction in the gender pay gap. If $\nu_\rho > 0$ on the worker side the same is true. The introduction of pay gap reporting results in a right-ward shift along the curves in each panel of Figure 19 and a lower gender pay gap.

In Section 6, I ask targeted survey questions to qualitatively capture (though not formally estimate) both the information parameter and preference parameter. Embedded in ν_μ and ν_ρ are consumers' and workers' interpretations of the pay gap information, which could affect their preferences over pay gaps. The survey in Section 6 provides evidence on this.

G Further details on ASHE dataset

Formerly known as the "New Earnings Survey", the Annual Survey of Hours and Earnings has been running since 1970 and is now administered by the Office for National Statistics (ONS). Workers enter the sample frame by having a particular pair of digits at the end of their National Insurance Number (NIN), the UK equivalent of a Social Security Number. Surveyors then identify the employer(s) of these individuals by Her Majesty's Revenue and Customs (HMRC) Pay As You Earn (PAYE) system, which is the UK government's tax withholding system used for the vast majority of employees. This takes place each January.

In most cases, survey forms are then sent to employers requesting information on the worker(s) in the sampling frame who are identified as working for that employer in the PAYE records. Employers then complete the forms using information from payroll records. For larger employers, much of the process is automated, with surveyors accessing payroll records and extracting information directly. The survey covers both the public and private sectors, but as it is administered via employers it excludes the self-employed, who constituted 15% of UK employment as of 2019.

The survey delivers useable information on 140,000 to 180,000 employees each year. The variables available include detailed wage and hours information for a snapshot week at the start of April, along with a limited measure of annual wages. Wages in ASHE are not top-coded. The hours data refer to the hours the employee worked that week, according to the employer.[35] The worker's three-digit occupation code is also available, along with worker age and sex. Employer identifiers allow the construction of within-employer tenure variables. Unless a worker migrates or becomes unemployed/self-employed, workers are followed throughout their entire working lives.

Employer data is linked to employees via employer identifiers ("entrefs"). Other Office for National Statistics data on employers is matched to provide some basic employer-level statistics.[36] This crucially includes the number of employees at the organization, which will be used in the empirical strategy outlined in Section 4 below. Employer information also includes other basic characteristics such as industry and region.

It is possible to link external data to ASHE. Using information on 3-digit occupation codes, I integrate skill data from the UK Government's Immigration rules,[37] which assigns one of six skill levels to each occupation. These range from lower-skilled to PhD level and reflect the average levels of skills required for an occupation.

I am also able to link a subset of employers to the gender pay gap reports themselves. One

[35]The hours variable is more accurate for hourly-paid workers. As is typical in administrative employer-reported hours data, there is significant bunching at a handful of discrete thresholds representing standard contracts. For some workers, these will not reflect the true hours worked over the week.

[36]The "entref" id variable refers to the 'enterprise', which is close to the typical definition of a 'firm' and is the standard grouping of organizations used by the ONS. Formally, an enterprise is the smallest combination of legal units (based on VAT or PAYE records) that is an organizational unit producing goods or services, benefiting from a degree of autonomy in decision-making, especially for the allocation of current resources. Enterprises carry out one or more activities at one or more locations and may be sole legal units.

[37]Immigration Rules Appendix J: codes of practice for skilled work, UK Home Office, Published 25 February 2016. Used under Open Government Licence.

potential issue for this project is that the ASHE firm size variable does not align precisely with that which is used to determine coverage of pay gap reporting requirements. The level at which firms report may also not be exactly the level at which they are identified by an "entref" code in ASHE. By matching reports to the ASHE data, I have verified that the two are in practice well-aligned. While there may be a small degree of misclassification, there is a discontinuous jump in the probability of reporting at 250 employees in ASHE.

While ASHE is thought to be exceptionally high quality, there exist several data deficiencies. Firstly, a set of workers and employers are excluded from the sampling frame. Employees that work at businesses outside of the interdepartmental business register (IDBR) are excluded as they cannot be identified. These are typically very small businesses who fall below the Value Added Tax (VAT) threshold. Given that my estimation sample will exclude small firms as discussed in the section below, this will not influence results. Relatedly, employees who earn below the National Insurance Lower Earnings Limit (LEL) are not subject to PAYE tax withholding so will not be in the PAYE system. The current LEL is set to £120, or approximately $150 per week. This means that a number of the very lowest-paid workers are omitted from the survey. A minimum-wage earner working 15 hours a week in the UK makes £130, so the majority of missing workers will be those working fewer than 15 hours a week on the minimum wage. This is a small group, but will disproportionately be made up of women, who are more likely to work part time. The Office for National Statistics (Bird, 2004) estimates that biases caused by omitting these missing firms and employees are likely to be small, though it is difficult to precisely quantify their extent. In my analysis I will restrict my sample to full-time workers, in part due to this concern.

Secondly, there is both worker-level and employer-level non-response in ASHE data. At the worker level, the cause of this is primarily that the employee has moved employers between the point at which employers are identified in the tax data and the point at which the survey is sent out. It can also be that employers have retained previous workers in their PAYE records, despite the fact that they no longer work there. This means that job movers are under-sampled in ASHE. The second type of non-response is at the firm level. Businesses are legally obliged to comply, though it is not clear how well this is enforced in practice. It is possible to identify in the data a small number of cases which are strongly suggestive of complete firm non-response.

A final broader concern is that when an individual is missing, it is not possible to definitively identify why this is. This makes the dataset inappropriate for studying movements into or out of employment. Missing individuals could be unemployed, have left the labor market, be self-employed or be missing due to the reasons described above.

Throughout my analysis, the sample is restricted to full-time workers aged 18-55. Consistent with other academic work using this data, observations are dropped if there is a loss of pay due to absence in the reference period. To be consistent with pay gap reports, our main outcome is hourly wage, deflated by the CPI. Observations are at the job rather than worker level, however the restriction to full-time workers means that in practice, very few workers in our sample hold multiple jobs.

H Further details on survey

The survey was piloted throughout August 2020 and administered in September 2020. Of the 2,000 survey takers, 160 were dropped in the estimation sample. These included 36 individuals with data deficiencies, and 124 individuals who either completed the survey too quickly to yield usable responses or failed an attention test.

Eligibility was determined based on records help by Prolific Academic, who were also able to provide basic demographics on participants to complement survey questions. The survey took respondents 7 minutes to complete on average and was hosted by Qualtrics. Here I go through the five main sections of the survey before discussing weighting and representativeness in Section H.6.

H.1 Background demographics, attitudes towards inequality and policy awareness

The survey started by asking for background information on demographics and work, before moving into more detailed questions on attitudes towards inequality and on policy awareness. This included the following question on policy awareness.

Respondents were asked "Under UK law, which of the following statements are correct? (ignore any temporary policy shifts due to Covid-19)"

- Employers have to pay employees under a certain maximum hourly wage
- Employers have to pay employees over a certain minimum hourly wage
- Employers have to pay employees with different educational qualifications the same hourly wage
- Large employers have to publicly report the difference between their lowest-paid worker's salary and their highest-paid worker's salary
- Employers are not allowed to discriminate on the basis of race or sex
- Large employers have to publicly report the difference between how much they pay men and how much they pay women

Three options were given, 'True', 'False' or 'Don't know'.

H.2 Hypothetical job choices

Participants were then asked to choose between two hypothetical jobs, in which one of the jobs was labelled as follows:

The employer offering job A was recently reported as having the highest gender pay gap in the industry. This means that the gap between average male and female wages was higher at this employer than at other employers.

H.3 Estimated own-employer pay gaps

Participants were next asked to estimate the gender pay gap at their primary employer. They were first given the following information on the median:

Consider lining up all male workers in the UK in order of their hourly wage. The median male worker is the worker in the middle of this line. The median female worker is defined in the same way.

The following prompt was given:

For every £1 earned per hour by the median male worker at your primary employer, the median female worker at your primary employer earns (in £):

Participants were then presented with a slider between 0.5 and 1.5.

H.4 GPG policy questions

In this section, participants were asked if:

- they were aware of the reporting policy before taking this survey
- their employer had communicated the pay gap to them
- they had looked up the pay gap of their employer or any employer
- they had seen news stories about the pay gap reports

H.5 Questions about particular pay gap reports

Participants were presented with a graphic based on the government's method of communicating pay gaps online. Two (unnamed) employers in the participant's industry were shown. Participants were then asked some basic comprehension questions, and then why they think the gender pay gap might differ between these two employers (details below). They were also asked whether they could name any companies which had been named and shamed in the media for having particularly high pay gaps.

H.6 Survey weighting and representativeness

The survey has been weighted to match sex and education of the 18-65 working population, using statistics from the labor Force Survey (LFS).[38] Exact weighting is used, such that each sex-education-age group cell is assigned a weight to match the LFS proportions.

Table 20 shows summary statistics for the unweighted and weighted samples, and population estimates drawn from the Labour Force Survey for comparison. While representative on gender, education (degree) and age, the survey under-represents foreign-born workers, the self-employed and full-time workers.

Table 20: Summary statistics of survey and LFS comparison

	(1) Unweighted	(2) Weighted	(3) LFS
Male (%)	33.5	49.0	50.9
	(1.10)	(1.10)	(.259)
Degree (%)	64.9	36.7	36.7
	(1.11)	(1.11)	(.25)
Age	34.9	39.6	41.4
	(.242)	(.242)	(.065)
Foreign (%)	10.5	8.10	16.7
	(.720)	(.720)	(.180)
Self-employed (%)	7.30	8.30	13.3
	(.720)	(.610)	(.180)
Hours	32.7	32.6	34.9
	(.247)	(.247)	(.058)
Observations	1,840	1,840	37,215

Notes: All statistics reported are means and standard errors. Column (1) shows statistics for the unweighted (raw) survey, column (2) the weighted sample, and column (3) the Labour Force Survey for comparison. Source: Own survey and UK Labour Force Survey.

H.7 Interpretation of pay gap information

Participants were presented with two pay gap reports and asked "why do you think the pay gap differs between these two employers?". Responses were free text. They were classified into a number of non-mutually exclusive classes. Below I describe the classification of the six most common classes:

- **Occupations**. Differences in the type of job, tasks the workers are expected to perform.

[38]Office for National Statistics, Northern Ireland Statistics and Research Agency. (2020). Quarterly Labour Force Survey, July - September, 2020. [data collection]. UK Data Service. SN: 8720, http://doi.org/10.5255/UKDA-SN-8720-1

- **Seniority**. Men being in more senior positions within an organization or having more experience.

- **Skills**. Differences in abilities, education or qualifications by gender.

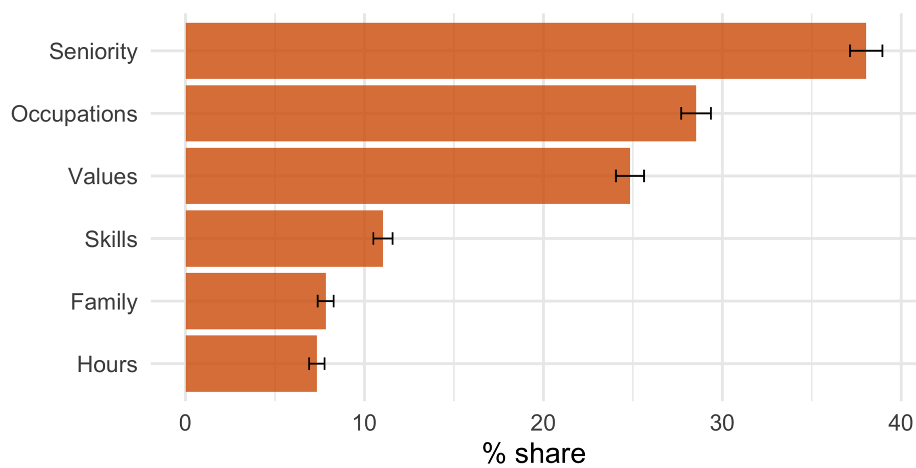
- **Hours**. Differences in the number of hours worked, stating that women are more likely to work part time.

- **Family**. Women having more childcare commitments, family responsibilities or taking time away from work for childbirth

- **Values**. Pay gaps reflecting different ethical viewpoints of employers, or the values of senior leaders.

These cases collectively covered 79% of responses. Of those not covered, a third mistakenly interpreted the information to reflect differences in the share of men and women at each organization, 20% stated that they were unsure, 6% gave no response and the remainder gave unclassifiable answers. The shares of the six most common responses are given here in Figure 20.

Figure 20: Frequency of interpretation classes



Notes: Classes based on responses to “why do you think the pay gap differs between these two employers?". Classes are not mutually exclusive. 95% confidence intervals given. Unweighted responses. Source: Own survey

Table 21 shows the predictors of each interpretation class. This demonstrates a number of intuitive patterns, for example that more educated workers are more likely to discuss seniority and family as underlying the gender pay gap differences.

Table 21: Predictors of interpretation class

	(1) Seniority	(2) Occupations	(3) Values	(4) Skills	(5) Family	(6) Hours
Male	-4.971*	10.20***	0.513	3.271	-3.471**	-0.341
	(2.865)	(3.015)	(2.890)	(2.322)	(1.503)	(1.709)
Age	-0.321***	0.153	0.0875	0.120	0.0630	0.0361
	(0.117)	(0.129)	(0.120)	(0.101)	(0.0635)	(0.0666)
Degree	11.62***	4.736*	2.099	-2.187	2.567*	0.153
	(2.894)	(2.871)	(2.600)	(2.014)	(1.455)	(1.683)
Income	1.594	0.894	-3.104***	-1.200	-0.451	1.496**
	(1.219)	(1.264)	(1.041)	(0.910)	(0.620)	(0.740)
Hours	0.0281	-0.0886	0.0581	0.232*	0.0771	-0.119
	(0.169)	(0.169)	(0.146)	(0.137)	(0.102)	(0.101)
>250 emp	-1.866	-3.613	2.830	-2.279	2.192	0.496
	(2.972)	(3.138)	(3.039)	(2.337)	(1.386)	(1.641)
Mean dep var	32.78	30.15	24.22	13.49	6.802	7.272
N	1840	1840	1840	1840	1840	1840

Notes: * $p<0.1$, ** $p<0.05$, *** $p<0.01$. Linear probability model in which outcome variable is indicator for being in prediction class indicated. Robust standard errors clustered at individual level shown in parentheses. Omitted controls are employment status, region and industry. Source: Own survey.

I Effects on stock market returns

The revelation of information on firms' gender pay gaps could plausibly influence firm value if investors anticipate an impact on consumer demand or that pay gap reports signal future regulatory intervention. While not the primary focus of this paper, here I adopt an event-study approach to investigate whether the publishing of gender pay gap statistics affects stock market returns. This approach has a rich history in finance (MacKinlay, 1997) and has also been applied to labor questions in the context of unions (Ruback and Zimmerman (1984), Bronars and Deere (1990), Lee and Mas (2012)) and more recently to minimum wage legislation (Bell and Machin, 2018).

I focus on the first round of reporting, the deadline for which was April 2018. I successfully matched 1,283 (12%) of reporting employers to 408 stocks on the London Stock Exchange, using FAME to match Company Numbers to ISIN codes. Each ISIN code corresponds to a single stock listing. Of these 1,283 total matches, 132 are direct 1:1 matches. Remaining employers are matched to one or more subsidiaries of a listed company. Where multiple reporting employers are matched to a single ISIN, I averaged over all matches.

I retrieved stock price data from Compustat, calculating abnormal returns for stock i at time t as:

$$r_{it}^{A} = r_{it} - \hat{\alpha}_i - \hat{\beta}_i \tilde{r}_t \tag{15}$$

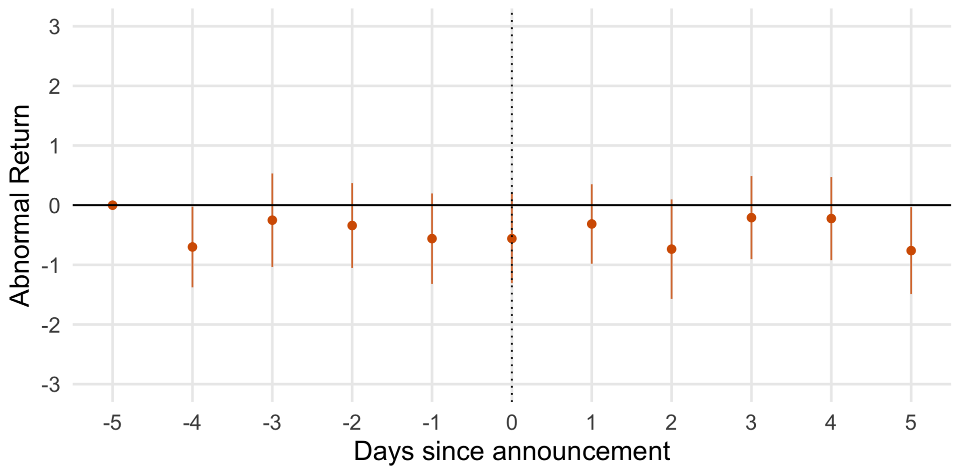
where r_{it} is the stock's raw return, $\tilde{r}_t$ is the FTSE all-share return and $\hat{\alpha}_i$ and $\hat{\beta}_i$ are the intercept and slope estimates from a regression of r_{it} on $\tilde{r}_t$ using all trading days in 2017 and 2018. I then estimate an event study specification, where I regress abnormal returns on trading-day dummies around the date on which pay gap reports are released. Panel (a) of Figure 21 shows the estimated date coefficients for an 11-day window around reports, with the estimate for day -5 normalized to 0. There is no evidence of any movement in returns around the date.

The previous figure pools all companies, irrespective of their pay gap report. In panel (b) of Figure 21 I split the sample into two based on whether they lie in the top or the bottom 50% of the distribution of (median) pay gap declarations. There is no evidence of a change in returns for either group. While with a small number of stocks it is not possible to obtain a great degree of precision, we cannot reject the null of no stock market response to pay gap declarations.[39]

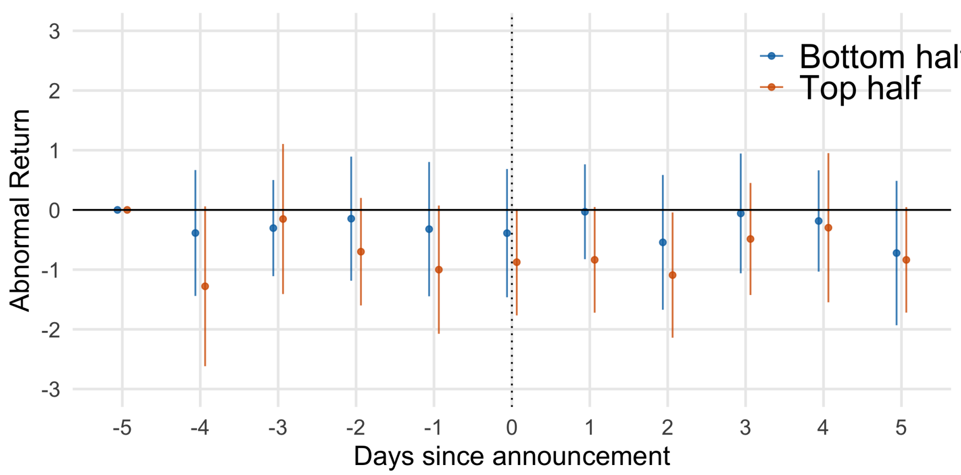
There are a number of reasons why despite finding evidence of firm adjustment the effect on stock market returns is not statistically different to zero. The first is that investors are more attuned to consumer behavior than worker preferences, and the latter factor drives a large share of adjustment. A second is that investors were unaware that the policy would have an effect on wages, unobservable at the precise time of reporting. This is not implausible for an entirely new type of policy.

[39]This echoes the perception of the media. After a thorough search, the Financial Times carries no mention of pay gap reporting having any influence on investors. Given the large number of potentially affected firms, any movements in stock price would have likely attracted media attention.

Figure 21: Abnormal returns around report release date



(a) aggregate



(b) by pay gap report

Notes: Outcome variable is abnormal returns. Day 0 corresponds to date which company released gender pay gap statistics. 95% confidence intervals based on standard errors clustered at the stock level. Sample divided into companies falling into the top half and bottom half of median gender pay gaps. Source: Compustat

CENTRE FOR ECONOMIC PERFORMANCE
Recent Discussion Papers

1749	Andrew E. Clark Conchita D'Ambrosio Simone Ghislandi Anthony Lepinteur Giorgia Menta	Maternal depression and child human capital: a genetic instrumental-variable approach
1748	Lee Elliot Major Andrew Eyles Stephen Machin	Unequal learning and labour market losses in the crisis: consequences for social mobility
1747	Tom Kirchmaier Monica Langella Alan Manning	Commuting for crime
1746	Andrew E. Clark Maria Cotofan Richard Layard	The true returns to the choice of occupation and education
1745	Yuhei Miyauchi Kentaro Nakajima Stephen J. Redding	Consumption access and agglomeration: evidence from smartphone data
1744	Philippe Aghion Antonin Bergeaud John Van Reenen	The impact of regulation on innovation
1743	Christian A.L. Hilber Andreas Mense	Why have house prices risen so much more than rents in superstar cities?
1742	Swati Dhingra Rebecca Freeman Hanwei Huang	The impacts of non-tariff barriers on trade and welfare
1741	Josh De Lyon João Paulo Pessoa	Worker and firm responses to trade shocks: The UK-China case

1740	Carlo Altomonte Domenico Favoino Monica Morlacco Tommaso Sonno	Markups, intangible capital and heterogeneous financial frictions
1739	Chad Bown Paola Conconi Aksel Erbahar Lorenzo Trimarchi	Trade protection along supply chains
1738	Christian Krekel Julia Rechlitz Johannes Rode Alexander Zerrahn	Quantifying the externalities of renewable energy plants using wellbeing data: The case of biogas
1737	Nick Jacob Giordano Mion	The UK's great demand and supply recession
1736	Gabriel M. Ahlfeldt Fabian Bald Duncan Roth Tobias Seidel	Quality of life in a dynamic spatial model
1735	Kilian Huber	Are bigger banks better? Firm-level evidence from Germany
1734	César Ducruet Réka Juhász Dávid Krisztián Nagy Claudia Steinwender	All aboard: the effects of port development
1733	Jonathan Colmer John Voorheis	The grandkids aren't alright: the intergenerational effects of prenatal pollution exposure

The Centre for Economic Performance Publications Unit
Tel: +44 (0)20 7955 7673 Email info@cep.lse.ac.uk
Website: http://cep.lse.ac.uk Twitter: @CEP_LSE