The Gender Gap in Compensation among UK firms

An-Chiao Liu, Gabriel Naylor-Leyland & Jasper Ginn

Introduction

The Gender Pay Gap defines the difference in hourly wage between men and women. This can be measured either by the mean wage, the median wage, or any other average. It is of interest not only to assess the size of the gender gap across a population, but also to uncover the driving factors behind its existence. For instance, are certain high-wage industries dominated by male workers? Does the population wage gap increase in positions of high wage? Do men and women earn comparable salaries for comparable jobs?

This report will focus on the wage gap in the UK in 2017. The data we analyse throughout the report comprise all companies with over 250 employees, and a small amount of smaller companies. As of 2017, it is required by the UK government that all such companies publish information pertaining to the pay of employees for the purpose of showing whether there is indeed a difference in the pay of male and female employees. This dataset is thus well protected against non-response errors, given that failure to comply with the regulations can lead to enforcement from the Equality and Human Rights Commission [1]. In addition, non-sampling errors are also well protected via the use of clear and simple categories, for example, size brackets for the number of employees.

Our aim is to create an efficient sampling design, from which we can make accurate calculations by using a smaller sample. We will compare the precision of a number of different such designs. Their accuracy will be measured against the results obtained from the complete data analysis.

One element of the data which we will refer to throughout this report is the Standard Industrial Classification (SIC). This is a widely used system for classifying companies into certain larger industries. Another important variable in the data is the size of an employer, ("*EmployerSize*"), which categorises each company into seven distinct sizes: size not provided, less than 250, 250 to 499, 500 to 999, 1000 to 4999, 5000 to 19.999, and 20.000 or more. We will discuss in detail whether using information on the gender pay gap within these categories can lead to better sampling designs.

Before conducting our analysis, we first made some modifications to the dataset. Firstly, the information that was not necessary was stored in a separate dataframe, which we did not use. For example, the person responsible for the submission of the data was not relevant, and was thus not included in our main

dataset. Six companies were then reclassified into different SIC divisions. Said observations fell under the divisions *activities of extraterritorial organisations and bodies* and *activities of households as employers*. The first reason for the reclassification was that, with just three companies each, these divisions would be too small to stratify. More importantly, however, information from the websites of said companies led to to their fitting in other SIC divisions. With this modified dataset of 6713 companies we proceeded with our analysis.

Population data analysis

Our analysis found that, across the population (of provided data), the average monthly wage for men was 14.4% higher than that for women. It should be noted that this provides only an initial insight into the extent of the gender gap in the UK, and should be interpreted with caution. For instance, it can often be more relevant to assess the size of gender pay gaps in terms of the median wage, so that the few disproportionately high earners do not skew the results. To improve the accuracy of this 'population' gender gap we should have weighted the population mean based on the size of the companies, so that the larger companies hold more influence on the population statistics. However, with the data provided, it would be difficult to achieve a meaningful accuracy, given that we are not provided with the exact size of the companies.
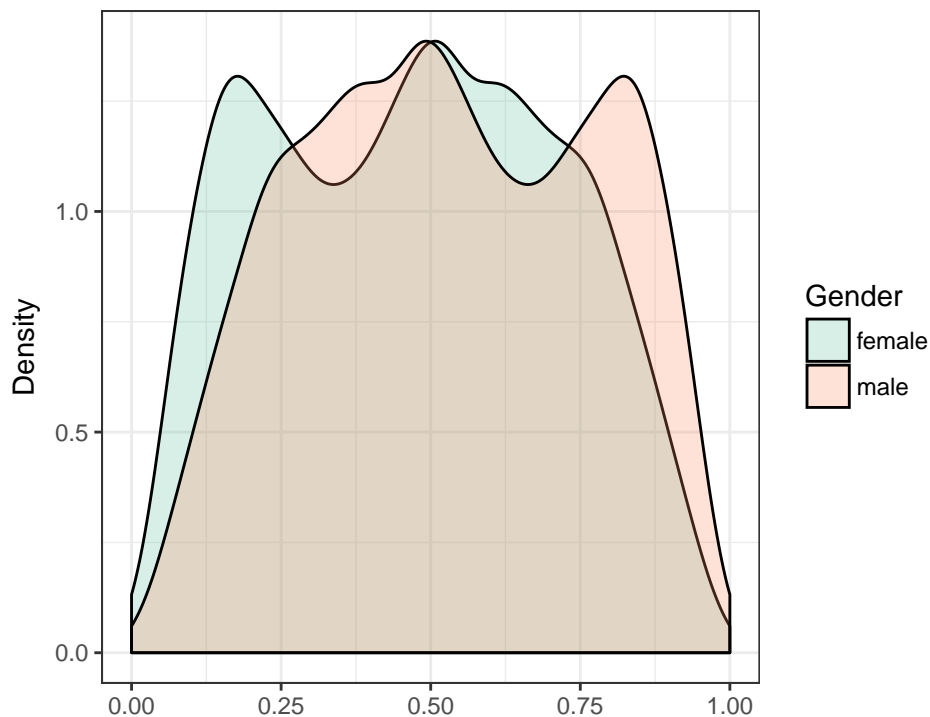
Figure 1: Density of different male-female employee proportions

Figure 1 gives us an early impression of the distribution of male and female workers within companies. We can clearly see that there are many more companies in which there are fewer female employees, with female employment proportion peaking at ~22% and 50%. The graph shows proportions, thus male employment peaks at ~78% and 50%. This is an interesting finding, which would well have implications on the size of the gender pay gap. We now seek to determine whether this holds across all sectors. Figure 2 shows a great deal of variation within sectors, in terms of the proportion of male and female employees. This marks SIC division as a potential variable upon which to stratify when taking a sample.[1]

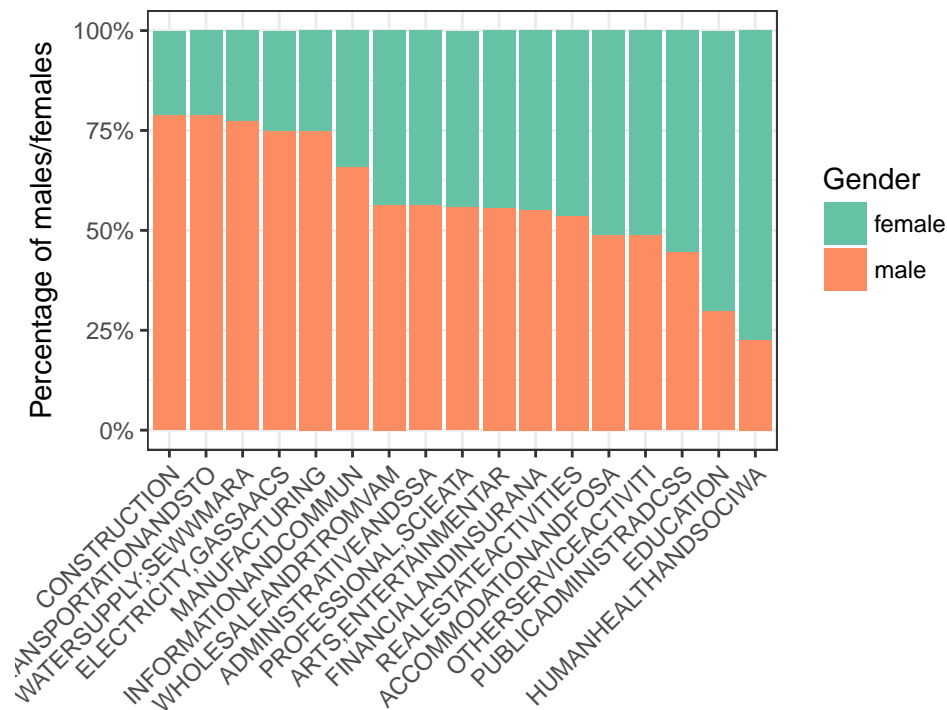[1]See the section on sampling methods below.

Figure 2: Proportion of males against females across the UK and SIC division

A brief analysis into the prevalence of the gender gap, namely in how many companies there exists a disparity in wages based on gender, revealed the following: 749, or 11.15%, of the companies in the dataset exhibited a gender gap (in the average hourly rate) favouring women, whereas 5910 companies, or 88.03%, paid men a higher average wage. (The remaining 0.008% of companies exhibited no gender wage gap). This indicates that many of the industry-wide differences in pay are likely to favour men. These figures are accompanied by the histogram (Left hand side: Figure 3), showing a heavy skew towards a 'positive' pay gap. Recall that a positive pay gap represents one in which males receive a higher wage than females.
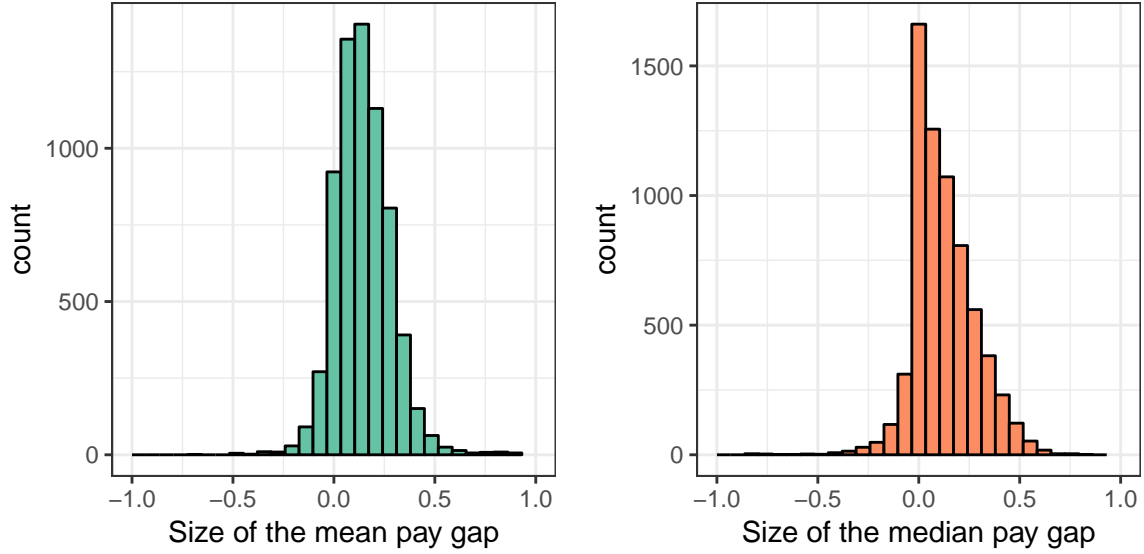
Figure 3: Histograms of the frequency of different sized mean and median pay gaps

The gender pay gap in the UK is also evident in the distribution of bonus payments; defined as any payment additional to an employee's salary. The gender gap for bonus payments stands at 14.1% at the UK population level, again in favour of men. The median value for bonus payments stands at 17.3% after correcting for a single unplausible outlier. As previously, this statistic falls short of a true population average, owing to the lack of appropriate weighting. However, for the purpose of brevity, these will be referred to herein as the true population values. As previously mentioned, comparing the median (hourly) wage within a company can also be an effective way to study the gender pay gap. Therefore, by averaging the median wage across all UK companies, we found that the median male employee earns 12.2% more than his female counterpart.

A glance at the distribution of both positive and negative wage inequalities again proves interesting: 903, or 13.45%, of UK companies pay a higher median wage to women in comparison to men - a significantly smaller figure than the 5258 companies, or 78.32%, who provide a higher median wage to males. (The remaining 552 companies, or 8.22%, provide an equal median wage). Similarly to the analysis of the mean wage, we can see from the histogram (Right hand side: Figure 3) that the distribution is strongly skewed in favour of higher male income.

Using the data on the gender of employees working within each income quartile (herein referred to as lower, lower middle, upper middle, top), we were able to get a sense of the distribution of employees at the company level. For example Figure 4 shows that a higher concentration of women currently work in the lower quartile, while male employees dominate the top and upper middle quartiles. This is likely to be one of the driving factors behind the 14.4% nation-wide pay inequality. However, in order to further this analysis, we would

require data pertaining to the size of the gender pay gap within each income quartile. For example, we know that males dominate the top quartile - and females the lower quartile - but we do not know in which quartile is the pay gap more pronounced.
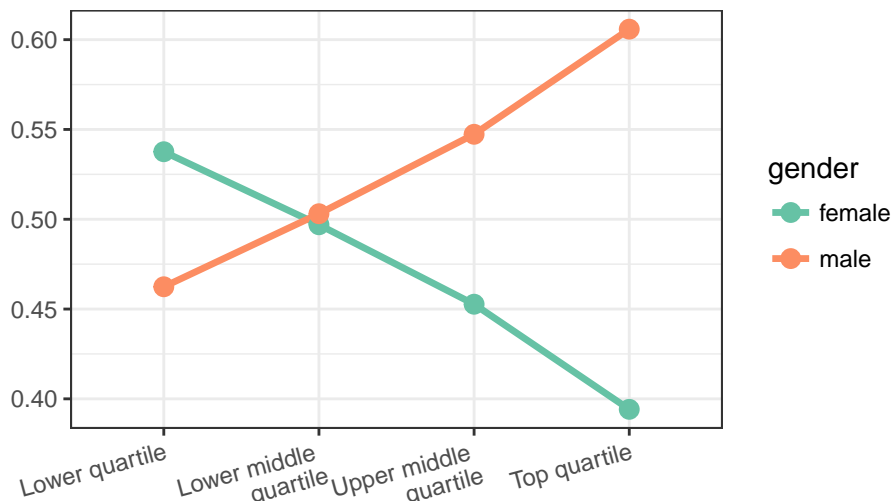

Figure 4: Distribution of males and females within each pay quartile

Sampling methods

The focus of this report largely falls on sampling methods, with the aim of making inferences on the target population using a (smaller) designed sample. The target population for our sampling was the UK working population; the goal of assessing the gender wage gap was to do so in the context of its widespread presence in the entire population. In the current situation, whereby companies with over 250 employees are required to provide relevant data, we worked with a dataset containing 6713 companies. The sample size in mind when conducting the following sample analyses was 1000. Thus, we endeavoured to make calculations related to the gender pay gap on such a sample which accurately reflect the UK population.

Firstly, we opted to generate a sample of 1000 random companies out of the 6713 available, with each company given an equal probability of being sampled (1000 / 6713 = 14.9%). This method is known as a simple random sample (SRS), and requires the assumption that a random sample is representative of the target population. The sample is taken without replacement, meaning that no company can be included more than once in the sample.

Working with this SRS of 1000 companies, we found that the gender pay gap stands at 11.71 % and 13.65 %

when assessing the median and mean wage, respectively. Both of these pay gaps are in favour of men. These values do not differ drastically from the population values of 12.2% and 14.4%. 95% confidence intervals for these percentages show little variation from our original population values, with the population values falling within the 95 % CI for both metrics (Table 1, below).

Table 1: Lower and upper CI for the SRS design

	Lower CI	Upper CI
Mean	12.88%	14.43%
Median	10.84%	12.57%

The standard error for the estimate for the difference in mean hourly wage using the SRS is 0.0040. Using this sampling design, we achieve a margin of error of $0.0040 \cdot 1.96 = 0.78\%$ at the 5% level. This relates to the sampling error - the difference between the sampled statistic and the population statistic is just 0.78%. We will use this SRS as a reference to compare all future sampling designs against, with the aim of creating a sampling design with a small standard error - this will indicate a small sampling error. These comparisons will be done via the design effect

$$deff = \frac{Variance_{design}}{Variance_{SRS}} \tag{1}$$

When the statistic of interest (here, gender wage inequality) is known to correlate in some way with information provided in the sampling frame, it can be useful to stratify. Arranging the data into strata prior to sampling can optimise the sampling design in the sense that the standard error of the 'predictions' can be reduced, thus improving the precision. This will only be the case if we see small variance within the strata, and larger variance between the strata. This will make for a more efficient design for the reason that, if a large portion of the total variance is between the strata, only a small sample from each stratum will be necessary to make accurate predictions.

NB: 3 observations have been removed due to excessive values (+−100% mean pay gap)
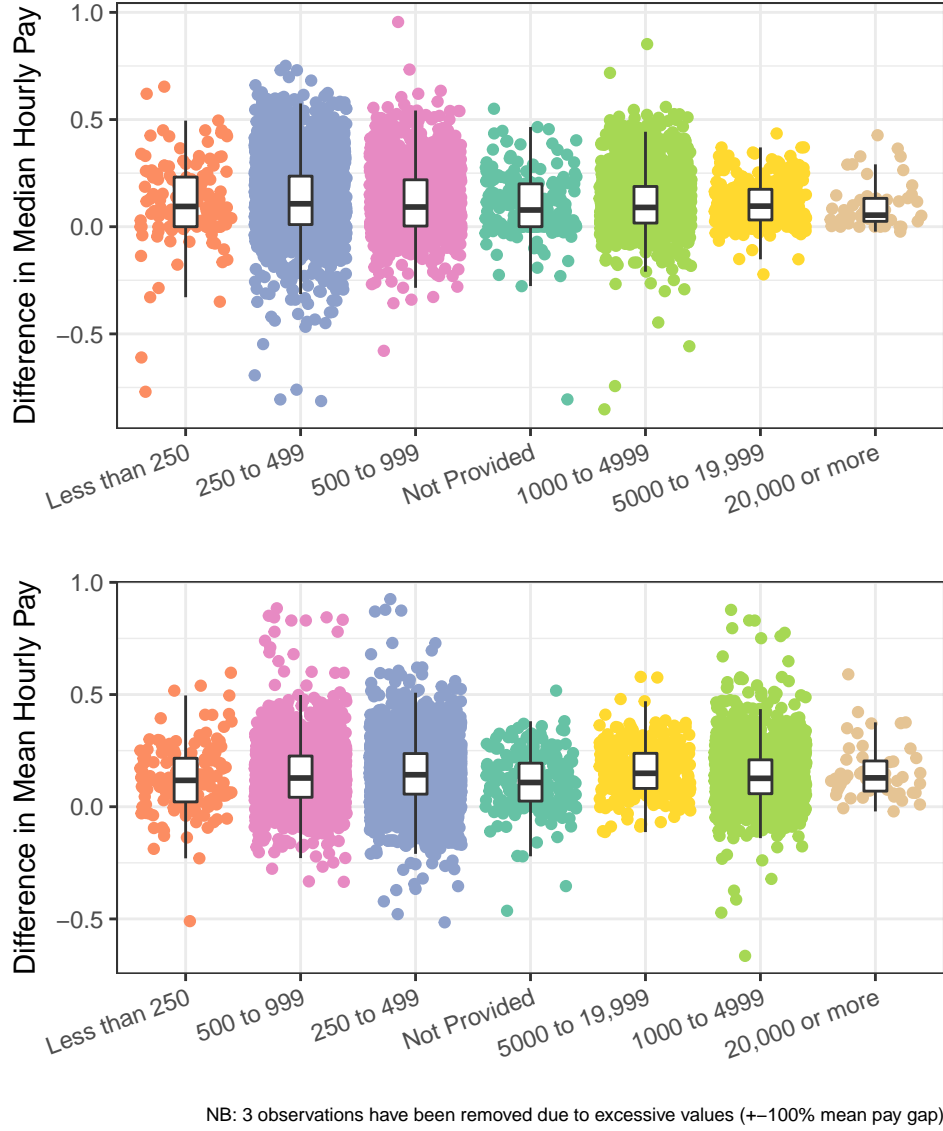
Figure 5: Distribution of median and mean pay gap by employer size (ordered by interquartile range)

Based on rudimentary analysis of the two potential stratification variables SIC and *EmployerSize*, and with the sample of 1000 still the goal, we decided to stratify based on SIC division. Evidence supporting this decision is provided in Figure 5: boxplots of the mean and median gender pay gap for the strata, here *EmployerSize*. From these plots, we do not see a great deal of variation between different company sizes, which suggests that *EmployerSize* may not be a particularly useful variable on which to stratify. Additionally, post-hoc tests following an ANOVA on the different categories of *EmployerSize* revealed very little evidence of between-group differences (only one of the 21 pairwise comparisons showed a significant difference). On the contrary, running an ANOVA on the SIC divisions yielded many significant pairwise differences in the

median pay gap. Figure 6 also shows that there is significant variation in the pay gap within SIC divisions.[2]

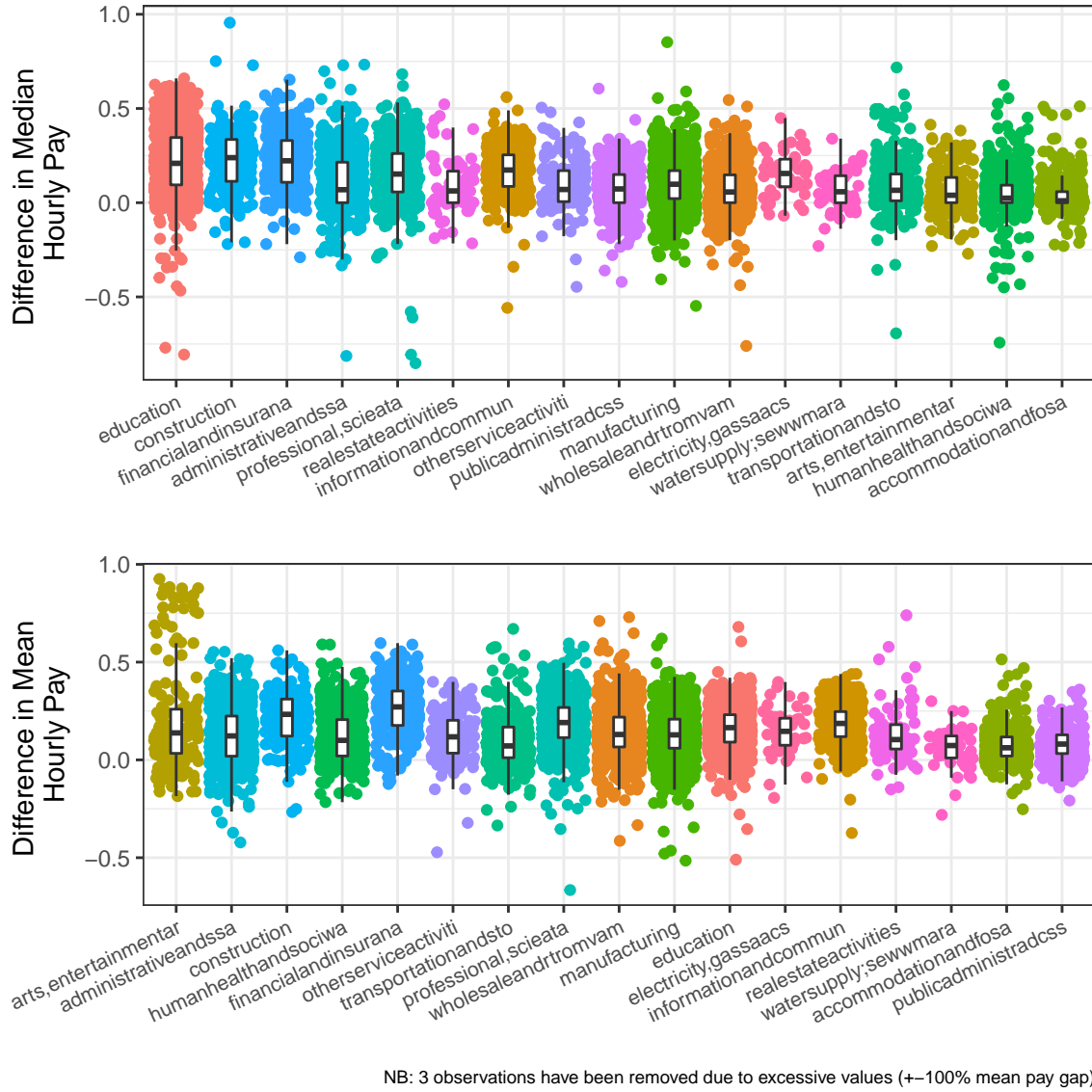NB: 3 observations have been removed due to excessive values (+−100% mean pay gap)

Figure 6: Distribution of median and mean pay gap by SIC division (ordered by interquartile range)

We then used a proportional-to-size (PPS) procedure to draw a stratified sample of (at most) 1000 companies. This technique ensures that the drawn sample well represents the population, given that the proportions of the divisions within the sample match those within the population. For example, the division *education* contains 986 companies, or 14.69% of the population. Within the PPS sample, education will comprise as close to this figure as possible. This method of sampling should ensure lower sampling errors, when

---

[2]We briefly considered geographic location as another potential stratification variable. In particular, we expected London to exhibit a different profile from the rest of the UK. A preliminary analysis for which we extracted the counties of some 4.500 companies did not show any difference between counties and therefore the idea was abandoned.

comparing with an SRS. This is because a PPS stratified sample is expected to better represent the target population (in this case all UK companies with over 250 employees).

Given that is often more useful to use the median to gauge the size of the pay gap, we will now proceed by comparing the predicted difference in median wage of our sampling designs.

The PPS sample tells us that there is a 12.29% difference in the median wage. This is closer to the population value (12.2%) than the SRS (12.4%), and the margin of error here is lower, at 0.8561% (standard error = 0.004368). This gives a sense of how far our predictions of the wage difference differ from the population values. This is calculated by $1.96 \cdot (s.e.)$, where 1.96 is the 95% z-value used to construct confidence intervals. The design effect of 0.8516 means a sample of only 852 would be required to achieve the same precision as our earlier SRS. These figures indicate that we can more efficiently (design effect) and accurately (margin of error) make inferences on the population with this sampling design. All values referred to here can be found in Table 2.

However, it is still possible to optimise the sample, via the choice of stratum sample size. This can be done by using a Neyman Allocation. This method allocates an optimal size for each stratum via the formula

$$n_h, Neyman = \frac{(N_h \cdot V_h)}{\sum_{i=1}^{H}(N_i \cdot V_i)} \cdot n \tag{2}$$

where N_h and V_h are the population stratum sizes and variances respectively and n is the desired sample size. In this way, it is possible to oversample from certain strata - namely those with higher variances - in order to achieve a smaller standard error. According to this sampling design, the difference in median pay stands at 12.1%. This value is about as close to the population value (12.2%) as the PPS sample. The standard error of 0.004115 is also marginally lower in comparison. This means we now have a design effect of 0.8313 and that we need a sample size of 832 to achieve comparable performance with this design when compared to an SRS. Consequently, we consider this design more efficient than the PPS, requiring fewer companies to achieve the same precision. We therefore conclude that, when seeking to maximise precision within a sample of 1000 companies, stratifying based on SIC division and using Neyman optimal allocation is better than taking an SRS or a PPS stratified sample, although it should be noted that all three designs are efficient.

Table 2: Comparison of sampling designs.

PG Metric	Design	Estimate	SE	MOE	Deff
Median	SRS	0.117055	0.0044254	0.0086738	1.0000
Median	PPS	0.122900	0.0043680	0.0085610	0.8516
Median	Neyman	0.121000	0.0041150	0.0080650	0.8313
Mean	SRS	0.136534	0.0039694	0.0077800	1.0000
Mean	PPS	0.142600	0.0039100	0.0076640	0.8997
Mean	Neyman	0.147900	0.0038690	0.0075830	0.8404

Sample size determination based on the coefficient of variation

Another common way in which to sample is to base the sample size around a prerequired level of precision. In this situation, the precision of the sample estimates dictates the sample size, as opposed to a fixed budget or sample size. With this technique in mind, we aim to find the sample size required in order to achieve a coefficient of variation of 0.01. The coefficient of variation (CV) is defined as the standard error of the wage gap, relative to the size of the wage gap - a standardised error term.

Firstly, we calculated the existing CV, taken from the optimised stratified sample, to be 0.026. This is significantly higher than the target of 0.01. We employed two approaches to find the desired sample size. Using manual calculations, we established that the sample size needed to achieve a CV of 0.01 is $n_{pps} = 3797$ and $n_{optim} = 3643$ for the PPS and Neyman stratified samples respectively. Additionally, we implemented a computational procedure in which we systematically increased the sample size, n, until the CV falls at or below the target of 0.01. Because this algorithm is not entirely exact[3], we run the algorithm several times, which leads us to conclude that a required sample size between 3640 and 3770 will suffice. The average sample size needed is nearly the same as that derived using manual calculations for the PPS design at $n_{avg} = 3785$.[4] Therefore, with a level of precision as determined by a CV of 0.01, it would be necessary only to sample roughly 57% of UK companies with over 250 employees, while still obtaining the same information.

---

[3]For example, the algorithm takes a slightly different sample for each iteration. This means that the CV oscillates and at times may increase even though the sample size is increasing and we would expect a *lower* CV

[4]We could not execute the algorithm for the Neyman design because of a lack of observations in certain subgroups. Hence, we opted to computationally determine the CV for the PPS sample only.

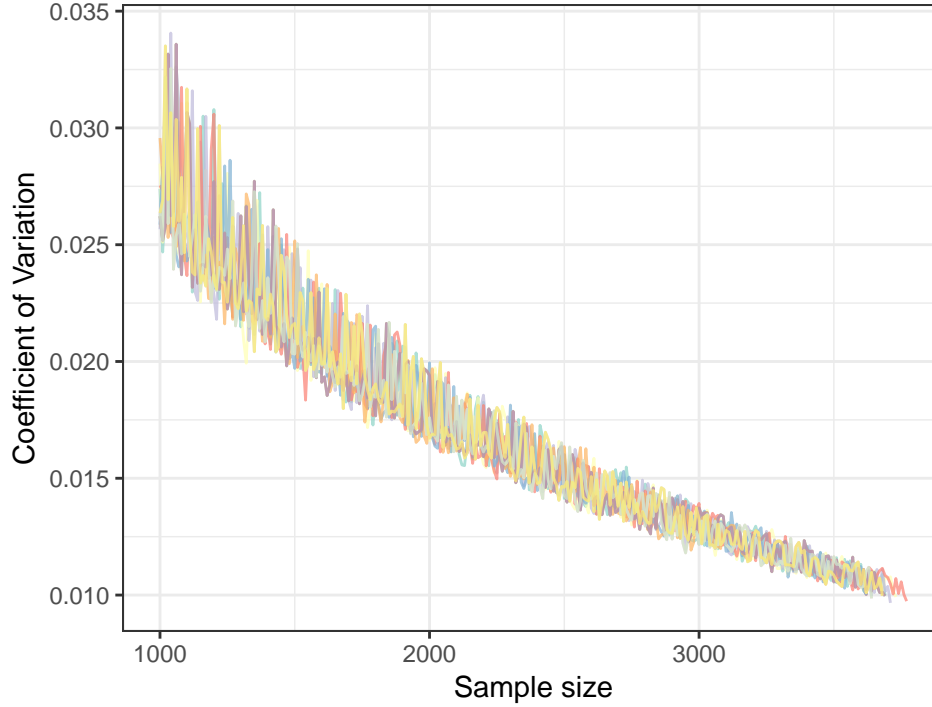NB. each line is one run of the algorithm. The total number of runs is 10.

Figure 7: Results of 10 runs of the algorithm that computes sample size given a CV of 0.01. Each colored line represents an individual run.

It is also of great interest to sample at the level of the individual employee, rather than at the company level. This can be done with a cluster sampling design. This method requires the weighting of the data according to the size of the company. In absence of the exact company size data, we opted to use the midpoints of the size categories as surrogate company sizes. The largest size category was an exception; we used the value of 20,000 to weight these companies. Companies not providing the number of employees were weighted as 0, and thus excluded from this analysis. Using this sampling design, we found that a weighted average for the mean gender pay gap stands at 15.6%, in favour of men, while the median pay gap stands at 10.4%, also in favour of men.

From our analysis we can conclude that, whether measuring the mean or median pay gap, there is a pronouced disparity in wages received by men and women across the UK population. In addition, a sufficient level of information to make inferences on the size of the wage gap can be obtained by taking a markedly smaller sample of companies. Even a simple random sample of 1000 companies was surprisingly successful in terms of standard error, when the average wage gap was required. This owes partly to the fact that, aside from SIC division, there is very little variation among natural strata across geographic areas or employer size. With very little added effort, a stratified sample can be obtained, which further increases the precision of

the estimates. Therefore, when attempting to calculate the population-wide gender pay gap in the future, it would suffice to take a sample of companies. We have also shown that it is possible to achieve a specific level of precision, by drawing a large enough sample.

APPENDIX I: mean/median gender pay gap ("PG") across gender/sic division

Table 3: Gender distribution and mean and median gender gaps
by SIC division

SIC division	Perc. Male	#Companies	Perc. Companies	Avg. Mean PG	Avg. Median PG
CONSTRUCTION	0.79	199	0.03	0.20	0.23
TRANSPORTATION AND STORAGE	0.79	278	0.04	0.10	0.09
WATERSUPPLY;SEWERAGE,WASTEMARA	0.77	46	0.01	0.06	0.06
ELECTRICITY,GAS,STEAMANDAIRCOS	0.75	49	0.01	0.14	0.15
MANUFACTURING	0.75	960	0.14	0.13	0.10
INFORMATION AND COMMUNICATION	0.66	284	0.04	0.18	0.17
WHOLESALEANDRETAILTRADE;ROMVAM	0.56	654	0.10	0.15	0.08
ADMINISTRATIVEANDSUPPORTSERVIA	0.56	708	0.11	0.13	0.11
PROFESSIONAL,SCIENTIFICANDTECA	0.56	489	0.07	0.19	0.16
ARTS,ENTERTAINMENTANDRECREATIO	0.56	190	0.03	0.21	0.07
FINANCIALANDINSURANCEACTIVITIE	0.55	308	0.05	0.26	0.22
REAL ESTATE ACTIVITIES	0.54	78	0.01	0.13	0.09
ACCOMMODATIONANDFOODSERVICEACT	0.49	336	0.05	0.07	0.04
OTHER SERVICE ACTIVITIES	0.49	139	0.02	0.12	0.10
PUBLICADMINISTRATIONANDDEFECSS	0.45	399	0.06	0.08	0.07
EDUCATION	0.30	986	0.15	0.16	0.22
HUMANHEALTHANDSOCIALWORKACTIVI	0.23	610	0.09	0.12	0.05

References

McGuinness, F., and D. Pyper. 2018. "The Gender Paygap." Government of The United Kingdom. www.parliament.uk/commons-library.