GAPSQUARE

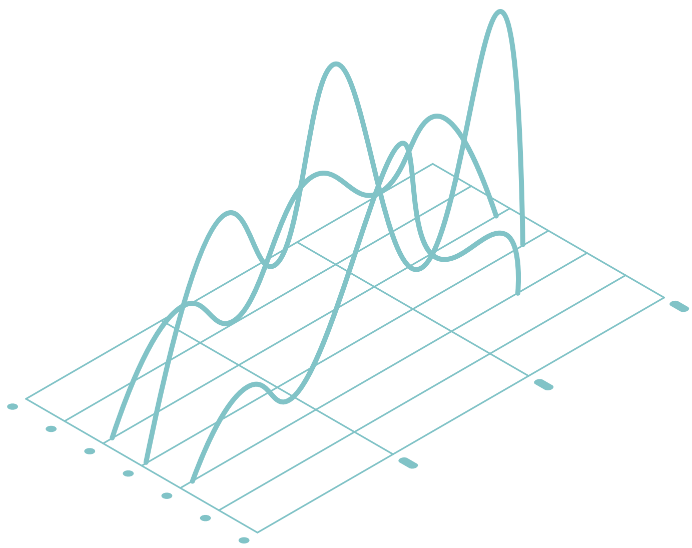
Regression Analysis, Unadjusted and Adjusted Pay Gaps Explained

A Technical Whitepaper

Mar 2021

GAPSQUARE

Contents



The Unadjusted Pay Gap

The unadjusted pay gap is the raw difference between average wages for a reference group and a comparator group. In the case of the gender pay gap (GPG), your reference group will be male and the comparator group will be female.

$$\textit{Unadjusted GPG} = \frac{\textit{average male salary} - \textit{average female salary}}{\textit{average male salary}}$$

This is the number that you see in your headline figures in the UK.

The unadjusted GPG, however, doesn’t take into account pay determining characteristics that can justify the disparities in pay, whether that be for gender, ethnicity, disability, etc. For example, tenure is a pay determining characteristic; the longer someone works in a company or industry, the higher their pay typically is. Protected characteristics like ethnicity or gender are not pay determining characteristics, because employers cannot pay someone more (or less) than others based on their protected characteristics.

Other pay determining characteristics can include:

- Age*
- Experience, i.e. tenure in the company or tenure in the industry
- Education
- Performance
- Job level
- Business unit/department
- Worked hours
- Location

*The older someone is, the longer their tenure typically is. However, experience and age are correlated with each other – therefore we only assume age to be a pay determining characteristic when there is no tenure data.

The Adjusted Pay Gap

The adjusted pay gap, on the other hand, does take into account pay determining characteristics. As a result, it is often much lower than the unadjusted pay gap. Some part of your unadjusted pay gap can be explained by pay determining characteristics, and the remaining pay gap is the adjusted pay gap. Therefore, the adjusted pay gap represents the unexplained portion of your unadjusted pay gap.

Adjusted Pay Gap = Unadjusted Pay Gap - Explained Pay Gap

The adjusted pay gap is important in understanding how much of your unadjusted pay gap could be due to either a lack of data, unobserved factors, or potential discrimination. Some factors behind pay gaps cannot be quantified. For example, different employees may have different skill sets even when working in the same job function, which could lead to higher pay, but this often is not recorded as a quantifiable pay determining characteristic. Additionally, the accumulation of historical gender segregation - where women typically work in lower-paid roles - causes women to earn less than men when they have different job functions. Therefore, although men and women having different job functions explains the unadjusted pay gap, it does not justify the reason behind this occupational segregation. This is not considered direct discrimination from the employer under the Equal Pay Act but is an effect of societal structures.

In the UK, EU and US, we expect a good adjusted pay gap to be <|5|%. Other regions in the world are benchmarked on a case-by-case basis. Despite this, employers should still look to decrease their adjusted pay gap entirely. An adjusted pay gap of <|5|% may not signify discrimination, but it does mean that over a lifetime, women will still earn less than men. Employers have the power to change the structures that have historically held women back, by aiming for an adjusted pay gap of 0%.

To fully tackle the adjusted pay gap, companies must first collect as much pay determining characteristic data as possible. This allows companies to see a more accurate adjusted pay gap. Looking at both types of pay gaps will allow you to identify the causes of pay differences and point you in the right direction to begin correcting problems as necessary.

To tackle the unadjusted pay gap, companies will often focus on changes in recruitment, promotion and retention. For instance, to close the unadjusted gender pay gap, companies can promote or hire more women into the top levels if there is a representation gap.

These methods will not necessarily work for the adjusted pay gap. The adjusted pay gap already uses, for example, differences in job level to explain why the unadjusted pay gap exists. Therefore, increasing representation of women in higher job levels may not close the adjusted gender pay gap. Instead, more work needs to be done on tackling biases and pay transparency in the workplace. Bias often goes unnoticed in the workplace, and quantitative data does not completely reflect how bias manifests itself. A lack of pay transparency also makes it difficult for women, ethnic minorities and disabled people to challenge the structures in place without knowing whether or not they are receiving fair pay.

Regression

Regression analysis is a statistical method whereby the relationship between a dependent variable can be drawn with one or more independent variables. In this case, the dependent variable is the gender (or ethnicity, disability) pay gap, and the independent variables are Pay Determining Characteristics.

While many regression models exist, the Blinder-Oaxaca decomposition is used as it specifically analyses pay gaps. The two Blinder-Oaxaca methods utilised are the threefold and twofold models. As the names suggest, threefold decomposition involves dividing the output into three components while twofold decomposition divides the output into two.[1] The methodology for both models are described below. For the purpose of explanation, males are the reference group and females are the comparator group - although this is changed when analysing ethnicity or disability.

The final output depends on how the data is structured, as we will choose the decomposition results that are most appropriate for the dataset.

The Threefold Decomposition Method

Let W_i be the wage of employee i with a vector of observed characteristics X_i

Using linear regression, the log hourly earnings[2] for male and female employees can be written as follows:

$$ln(W_i) = \beta^m X_i + U_i$$

$$ln(W_i) = \beta^f X_i + U_i$$

1 As of Mar 2021, Gapsquare uses twofold decomposition, moving to threefold in the latter half of 2021. Speak to your account manager about which type of decomposition you are using.

2 Most pay distributions are right-skewed (earnings are distributed towards the lower end of the salary distribution), so using a log scale helps to normalise the distribution.

Where β^m and β^f are vectors of parameters (wage returns to characteristics X_i) for males and females, respectively, and U_i is the error term.
Using the assumption that the error term is 0 on average ($E[U_i] = 0$) then the mean log hourly wages of males and females (represented by the bar on top of the variables) gives:

$$\overline{ln(W_m)} = \beta^m \overline{X^m}$$

$$\overline{ln(W_f)} = \beta^f \overline{X^f}$$

The average wage difference between males and female is then

$$\overline{ln(W_m)} - \overline{ln(W_f)} = \beta^m \overline{X^m} - \beta^f \overline{X^f}$$

This can be rearranged into a *threefold*[3] decomposition:

$$\overline{ln(W_m)} - \overline{ln(W_f)} = \beta^f (\overline{X^m} - \overline{X^f}) + (\beta^m - \beta^f)\overline{X^f} + (\overline{X^m} - \overline{X^f})(\beta^m - \beta^f)$$

Where

1st term: characteristic effect = **explained component.** This effect comes from **differences in characteristics between men and women** e.g. education, age, tenure etc. if this component is 0 then it means that, on average, age, education etc are the same for men and women.

2nd term: coefficient effect = **unexplained component** this effect captures the differences in the **returns to observed characteristics.**

3rd term: interaction effect = correlation component. This term accounts for the fact that between both groups, differences in characteristics and returns (coefficients) can exist at the same time. For example, the 1st term fixes the coefficient to that of women and only looks at differences in characteristics (vice versa for the 2nd term). The 3rd term consolidates the two prior terms for differences in both characteristics and coefficients.

3 https://journals.sagepub.com/doi/pdf/10.1177/1536867X0800800401, https://cran.r-project.org/web/packages/oaxaca/vignettes/oaxaca.pdf

The characteristic effect measures the average difference in females' mean wage if females had males' observed characteristics. The coefficient effect again is expressed from the viewpoint of the comparator group; it measures the expected change in the mean wage of females if they had male returns to a characteristic (male coefficient). The interaction effect means that threefold decomposition is most useful when there are differential effects between components across the two groups. For example, tenure could impact men and women's careers differently; women may have a cut in salary the longer they work if they have had to take maternity leave. In this case, the assumption that tenure increases pay may not be true i.e. women's yearly salaries at a point in their life could actually be decreasing.

Of course, this is not the case in every dataset, and thus twofold decomposition can be used.

The Twofold Decomposition Method

Let W_i be the wage of employee i with a vector of observed characteristics X_i

Using linear regression, the log hourly earnings for male and female employees can be written as:

$$ln(W_i) = \beta^m X_i + U_i$$

$$ln(W_i) = \beta^f X_i + U_i$$

Where β^m and β^f are vectors of parameters (wage returns to characteristics X_i) for males and females, respectively, and U_i is the error term.

Using the assumption that the error term is 0 on average ($E[U_i] = 0$) then the mean log hourly wages of males and females gives:

$$\overline{ln(W_m)} = \beta^m \overline{X^m}$$
$$\overline{ln(W_f)} = \beta^f \overline{X^f}$$

The average wage difference between males and females is then:

$$\overline{ln(W_m)} - \overline{ln(W_f)} = \beta^m \overline{X^m} - \beta^f \overline{X^f}$$

Let β* be the non-discriminatory coefficient vector i.e. the set of regression coefficients that would exist if there was no labour market discrimination. To calculate this, the twofold decomposition model uses an assumption that discrimination is only directed towards the comparator group i.e. women or BAME people. This method will not account for cases of "positive" discrimination.

The average wage gap can then be written as the sum of two components; an explained and unexplained component

$$\overline{ln(W_m)} - \overline{ln(W_f)} = (\overline{X^m} - \overline{X^f})\beta^* + [\overline{X^m}(\beta^m - \beta^*) + \overline{X^f}(\beta^* - \beta^f)]$$

Where

1st term = explained component: This effect comes from differences in characteristics between men and women e.g. education, age, tenure etc. if this component equals 0 then it means that on average age, education etc are the same for men and women.

2nd term = unexplained component: This compromises of the differences in the returns to observed characteristics (if coefficients differ then men and women do not receive the same returns to a characteristic ⟶ possible discrimination) and differences in unobserved characteristics (not in the data and/or discrimination).

There is no interaction term for the twofold model as it comes with the assumption that pay determining characteristics all have the same impact on men's and women's careers. Taking the aforementioned example of tenure, the twofold model assumes that the longer men and women work, the more their pay should be (up to a point i.e. retirement) - therefore if most men have worked in a company longer than women, this can be used to explain the GPG provided that women also see the general trend of their tenure increasing the longer they work.

Both threefold and twofold models can be used to analyse the adjusted pay gap. The Blinder-Oaxaca decomposition analyses and compares the results of both models, although the threefold results are considered default.

Limitations of the Regression Model

Explanation of pay gaps

While regression models will explain *what* contributes to your unadjusted pay gap, it does not explain *why* these factors exist. Regression analysis may tell you that the unadjusted GPG is due to men working at higher job levels than women, but it cannot tell you why women are underrepresented in the upper quartiles. Equally, it may tell you that more men work in business units - such as tech - which generally pay more. However, it does not tell you why it would be difficult for women to break into these roles.

Nuances with statistical inference

The statistical significance of the *entire* result is shown on our app as the Prob F-statistic. The Prob F-statistic is a measure of the variance - the lower it is, the lower the results vary within the groups. If this value is 0.0001, it means that there is a 0.01% chance that the regression results will be zero. This means that the lower the Prob F-statistic is, the higher the validity of our regression model. To conclude that a regression result is statistically significant, the Prob F-statistic must be *less than* 0.05

However, another value we can look at is the P-value of each pay determining characteristic. It represents the probability that test results will be at least as much as the results that are actually observed. When a P-value is less than 0.05, we usually say that there is correlation between a factor and the result. This is interpreted slightly differently in the Blinder-Oaxaca results.

This is because a high P-value can mean two different things:

1) The pay determining characteristics have no correlation with pay

2) The effect of pay determining characteristics are the same for both comparator and reference groups

The first interpretation is straightforward. The second one means that for example, in terms of gender, a pay determining characteristic like job level will affect men and women the same way - the higher someone's job level, the higher they are paid. Therefore, there is no difference in trends between men and women. This does not mean that there are no differences in pay between men and women, thus the unadjusted pay gap still exists. The Blinder-Oaxaca method takes into account all differences in pay determining characteristics between the reference and comparator group - regardless of each individual characteristics' P-value. This is because our aim is to look at explaining the unadjusted pay gap using pay determining characteristics.

To interpret how accurate the Blinder-Oaxaca results are, we recommend looking at the Adjusted R-squared and the Prob F-statistic together.

The R-squared describes how well the dataset fits the regression curve. The Adjusted R-squared shows how well the dataset fits the regression model, taking into account the actual number of data points considered by the model. Therefore, if you add insignificant data points into the regression, the Adjusted R-squared will decrease as more and more data points are unused by the model.

For both R-squared and Adjusted R-squared, we describe 80-100% as high accuracy, 70-80% as fair accuracy, 50-70% as low accuracy. We tend not to use results with R-squared or Adjusted R-squared lower than 50%.

Robustness of the model

Regression cannot be run on all datasets. Two main factors that affect regression analysis are:

1. The size of the dataset
2. The distribution of data points within the dataset

Therefore, while we expect the minimum size of the dataset to be 25 (+/- 10) employees in both the reference and comparator group, we can expect regression to fail on some larger datasets too. Linear regression essentially attempts to fit the data points into a straight line. Even if the dataset is large, it can still fail if the data distribution cannot fit into this line.

Endogeneity issues

The Blinder-Oaxaca model does not completely remove the risk of endogeneity. Endogeneity is when an explanatory variable (i.e. a pay determining characteristic) is correlated with a variable included within the error term. This happens when some variables are omitted from the data; as previously stated, not every factor that impacts a pay gap is necessarily quantifiable or recorded by employers. It can also happen when an outcome is not only a result of the predictor variable used by the model. For example, if you only had education data, the resulting unadjusted pay gap is not necessarily just a result of education. Education itself can be impacted by many other factors such as private versus public schooling and country of origin. In the worst case scenario, a pay determining characteristic can be correlated with the regression outcome itself (the pay gap). A single mother who has earned consistently less than men throughout her whole life is less likely able to afford quality education for her children. This can lead to bias, in which the effect of a pay determining characteristic is *overestimated* without considering endogeneity issues.

Interpretation of the Results

Since regression analysis looks at the contribution of different pay determining characteristics separately towards the unadjusted pay gap, we are then able to see which characteristics explain the unadjusted pay gap most. For example, if tenure contributes 10 percentage points towards an unadjusted gap and education contributes only 3 percentage points, then tenure has the larger impact by contribution. Each contribution is then subtracted from the unadjusted pay gap to give the adjusted pay gap. The adjusted pay gap can therefore be interpreted as the gap that is *unexplained* by these variables. But contribution by variables may be negative too.

A positive pay gap is when a group (such as women or BAME) is paid less than their reference group (such as men or White people). Conversely, a negative pay gap is when the reference group is paid less. For both positive and negative pay gaps, a negative contribution by variable will make your adjusted pay gap increase; subtracting a negative number is the same as adding a positive number.

There are a few ways to interpret negative contributions. The first is that based on the variable(s) with negative contribution, a pay gap should be higher than it is. Taking the gender pay gap as an example, if the only pay determining characteristic in the data is education, and women are paid more than men for the same level of education, then the regression results would tell us that men need to be paid more. However, this is not conclusive, as other factors – such as job level or performance – have not yet been provided in the data. Your actual adjusted pay gap will be influenced by the sum of all variables' contribution, so even if some variable contributions are negative, there could be larger positive values.

A second interpretation depends on the magnitude of the negative value. If a value is fairly close to zero – for example, -0.00001 – then the variable in question is not as significant.

Equally, the result of adjusted pay gap itself can be negative – not just the contributions by variable. If it is within the +/-5% boundary, then it can still be considered a "good" adjusted pay gap. Generally, a negative adjusted pay gap would imply that some women are being paid more than men for *reasons that the data cannot explain*. This does not necessarily mean that "positive discrimination" is occuring; it may mean that we require more data or more granular analysis to further explain the pay gap.

Still have questions?

Get in touch at hello@gapsquare.com