Office for National Statistics

Article

Disability pay gaps in the UK: 2018

Earnings and employment for disabled and non-disabled people in the UK, raw disability pay gaps and factors that affect pay for disabled people.

Contact:
Tom Evans
policy.evidence.analysis@ons.gov.uk
+44 (0) 1633 65 1833

Release date:
2 December 2019

Next release:
To be announced

Table of contents

1 . Main points

- This report presents the first analysis of disability pay gaps in the UK using newly reweighted earnings data from the Annual Population Survey.
- Median pay was consistently higher for non-disabled employees than for disabled employees; in 2018, median pay for non-disabled employees was £12.11 an hour whilst for disabled employees it was £10.63 an hour, resulting in a pay gap of 12.2%.
- The disability pay gap was wider for men than for women.
- In 2018, London had the widest disability pay gap at 15.3% and the narrowest pay gap was in Scotland, at 8.3%.
- Disabled employees with a mental impairment had the largest pay gap at 18.6%, while for those with a physical impairment the pay gap was 9.7% and those with other impairments had the narrowest gap, at 7.4%.
- Around a quarter of the difference in mean pay can be accounted for by factors such as occupation and qualification.

2 . Introduction

Disability is one of the nine protected characteristics under the Equality Act 2010. Previous work by the Office for National Statistics (ONS) has analysed pay gaps for two of the protected characteristics, the pay gap between men and women in the UK [Ardanaz-Badia and Rawlings (2018)] and between different ethnicities in Great Britain [Anderson (2019)].

For the first time, this report presents analysis on the disability pay gap using a new earnings weight on the Annual Population Survey. This allows for more detailed analysis of disability and pay than was previously possible. As this is the first presentation of such analysis by the ONS we would welcome feedback to Policy.Evidence.Analysis@ons.gov.uk on how this analysis could be developed in the future.

3 . Definitions of disability, impairment and pay gap

Throughout this report we will be using the following definitions of disability and impairment.

Disability

There are multiple definitions of disability. The Office for National Statistics (ONS) has previously published the effect that these different definitions of disability have on the prevalence rate.

To define disability in this publication we refer to the Government Statistical Service (GSS) harmonised “core” definition: this identifies “disabled” as a person who has a physical or mental health condition or illness that has lasted or is expected to last 12 months or more, that reduces their ability to carry-out day-to-day activities.

The GSS definition is designed to reflect the definitions that appear in legal terms in the Disability Discrimination Act 1995 (DDA) and the subsequent Equality Act 2010.

The GSS harmonised questions are asked of the respondent in the survey, meaning that disability status is self-reported.

Impairment

An impairment is defined as any physical or mental health conditions or illnesses lasting or expecting to last 12 months or more. Respondents were presented with a list of impairments and then asked to select all that apply and subsequently their "main health problem". The commentary in this report refers to the main health problem and does not explore where disabled people experienced more than one impairment, though this is something we could explore in future analysis.

This report groups impairments into mental, physical and other categories, based on the following classifications:

- mental impairments cover those with "depression, bad nerves or anxiety", "epilepsy", "learning difficulties" or "mental illness or nervous disorder"
- physical impairments cover those with "problems with arms or hands", "problems with legs or feet", "problems with back or neck", "difficulty in seeing", "difficulty in hearing", "speech impediment", "skin conditions or allergies", "chest or breathing problems" "heart, blood pressure, or blood circulation problems", "stomach, liver, kidney or digestion" or "diabetes"
- other impairments relate to those with "progressive illness not included elsewhere (for example, cancer, symptomatic HIV or multiple sclerosis)" or "other health problems or disabilities"

It should be noted that due to small sample sizes the "severe or specific learning difficulties" response option has been grouped alongside wider mental impairments. Though this shouldn’t affect the robustness of our results, we recognise that learning difficulties are typically regarded separately as they often have different causes and effects to mental health conditions. If users feel this response option should be grouped elsewhere then, as noted previously, please provide feedback to Policy.Evidence.Analysis@ons.gov.uk.

Pay gap

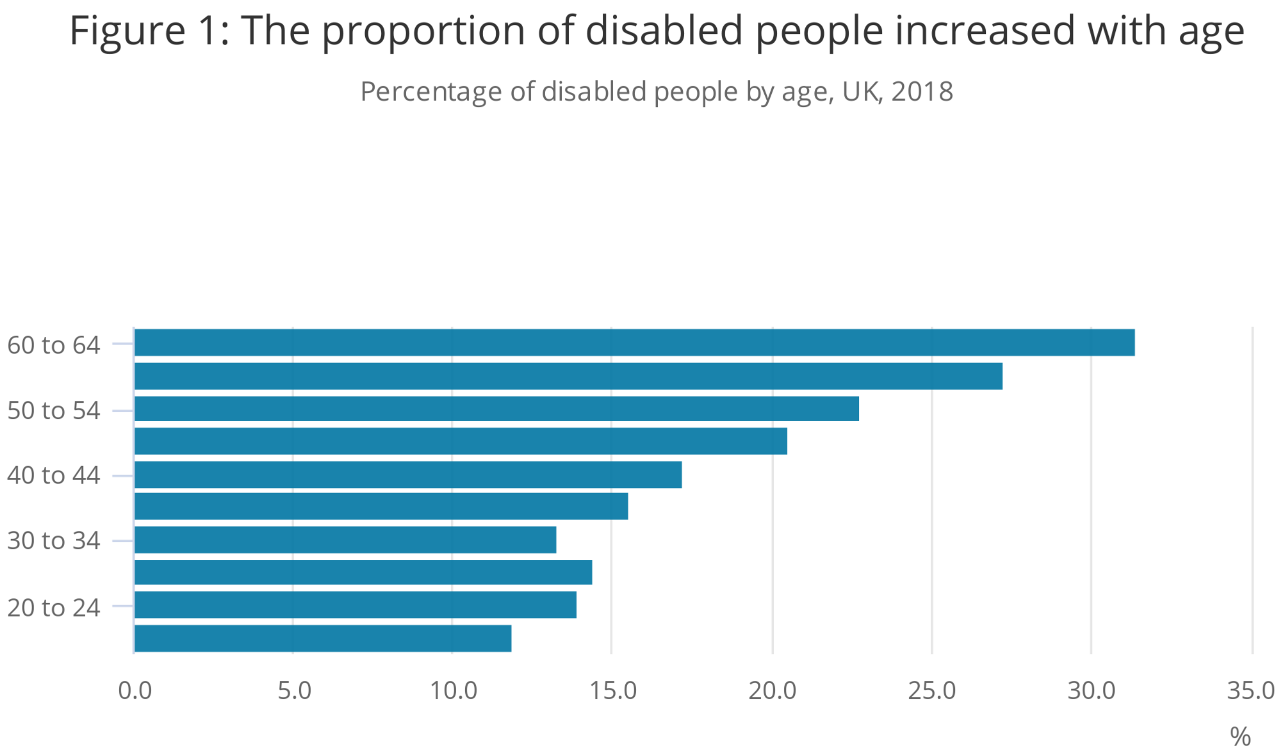
In this study the headline measure for pay gaps is defined as the difference between disabled and non-disabled average (median) hourly pay as a proportion of non-disabled average (median) hourly pay. For example, if the pay gap is 5.0% then disabled people are on average being paid 5.0% less than their non-disabled counterparts, whilst a negative 5.0% pay gap would denote that a disabled person is being paid on average 5% more than their non-disabled counterparts.

4 . Nearly one in five of the UK population aged 16 to 64 years was disabled

Using the Government Statistical Service (GSS) harmonised definition of disability, 18.9% of people aged 16 to 64 years were disabled in 2018. Women were more likely to be disabled than men, at 21.1% and 16.6%, respectively.

Figure 1: The proportion of disabled people increased with age

Percentage of disabled people by age, UK, 2018

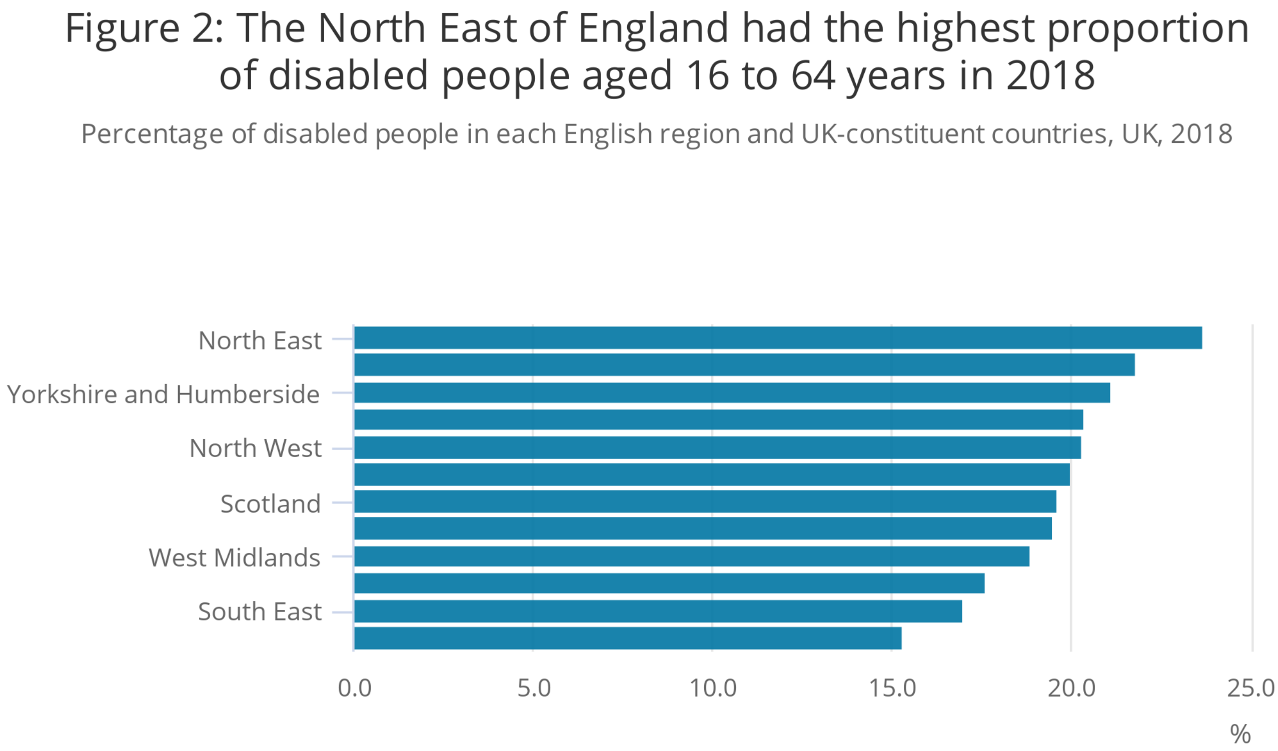


Source: Office for National Statistics – Annual Population Survey

As might be expected, the prevalence of disability increased as people get older (Figure 1). In 2018, people aged 60 to 64 years were almost three times more likely to be disabled, at 31.4%, than people aged 16 to 19 years, of whom 11.9% were disabled.

Figure 2: The North East of England had the highest proportion of disabled people aged 16 to 64 years in 2018

Percentage of disabled people in each English region and UK-constituent countries, UK, 2018



Source: Office for National Statistics – Annual Population Survey

The region with the highest share of disabled people in its 16 to 64 population was the North East at 23.7%. London was the region with the lowest share of disabled people, at 15.3% of its 16 to 64 population.

As age influenced the prevalence of disability, differences in the regional breakdowns could be influenced by the age structure of the regional populations. Figure 3 shows the mean age in each region including breakdowns by disability status. The mean age in London was lower than the other regions, which may partially explain why London had a lower share of disabled people in 2018.

Figure 3: The mean age in London was the lowest out of all parts of the UK

Mean age by region, 2018

[Download the data]

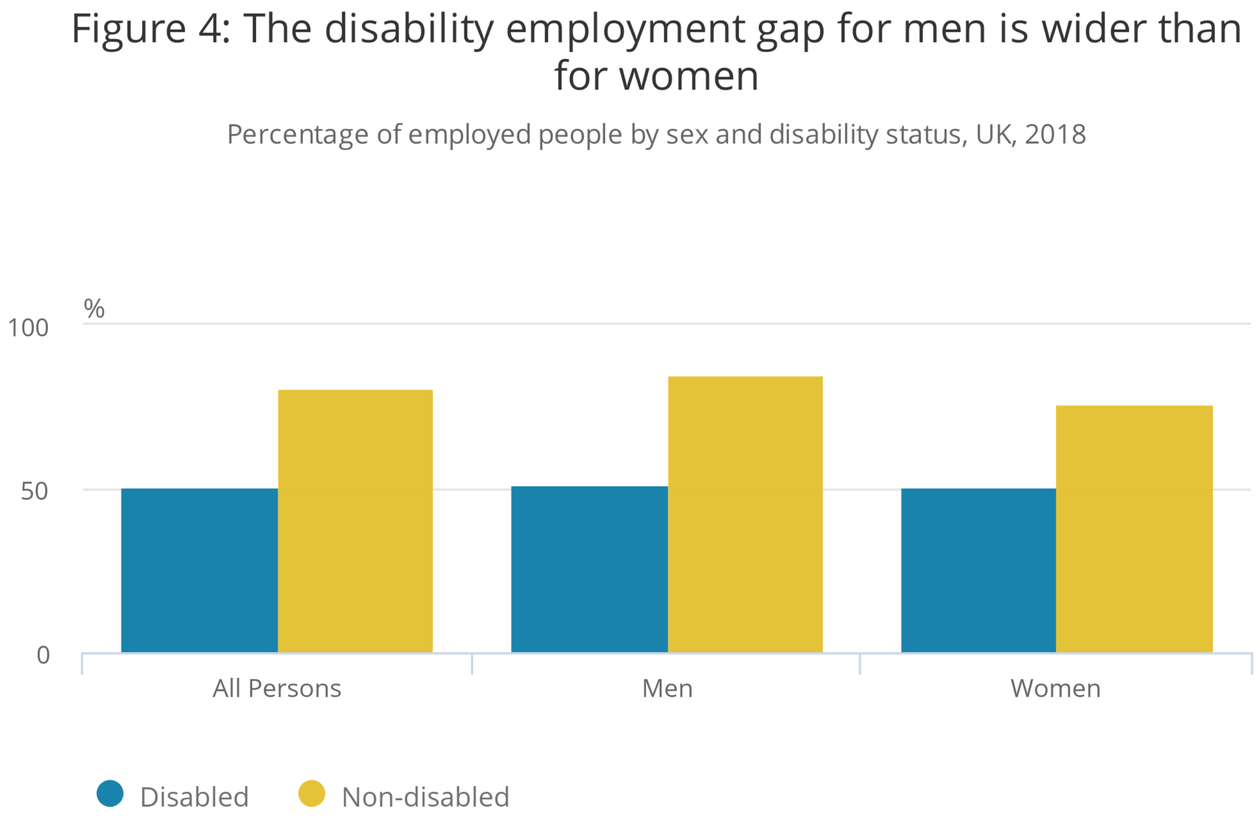
We have also published information on the [education outcomes], the [well-being] and [the social participation] of disabled people to give a wider picture of their experiences.

5 . Half of disabled people aged 16 to 64 years were employed

In 2018, of disabled people aged 16 to 64 years, 50.9% were in employment, whilst 80.7% of non-disabled people were in employment.

Figure 4: The disability employment gap for men is wider than for women

Percentage of employed people by sex and disability status, UK, 2018



Source: Office for National Statistics – Annual Population Survey

Employment rates for disabled men and women were similar at 51.7% and 50.4%. However, the disability employment gap was wider for men, largely driven by higher employment rates for non-disabled men compared with non-disabled women.

Disabled people were more likely to be self-employed than non-disabled people, at 15.2% and 13.9% respectively. Figure 4 shows the breakdown of those employed. This has remained relatively constant since 2014.

Disabled people were more likely to be economically inactive[1] (44.3%) than non-disabled people (16.3%). Of disabled people who were inactive, 57.2% said that their disability was the reason why they were inactive. The most common reason for inactivity given by non-disabled people was because they were students, at 39.1%. Other main reasons for being inactive for both disabled and non-disabled people were that they were either retired or looking after their family or home.

Of disabled people with a mental impairment, 53.3% were economically inactive in 2018. This was the highest out of the three categories of impairment, while physical impairment was the lowest, at 38.7% economically inactive, and other impairments were at 46.2%.

More information on the employment characteristics of disabled people can be found in [Disability and employment](), which includes a more detailed breakdown of impairments.

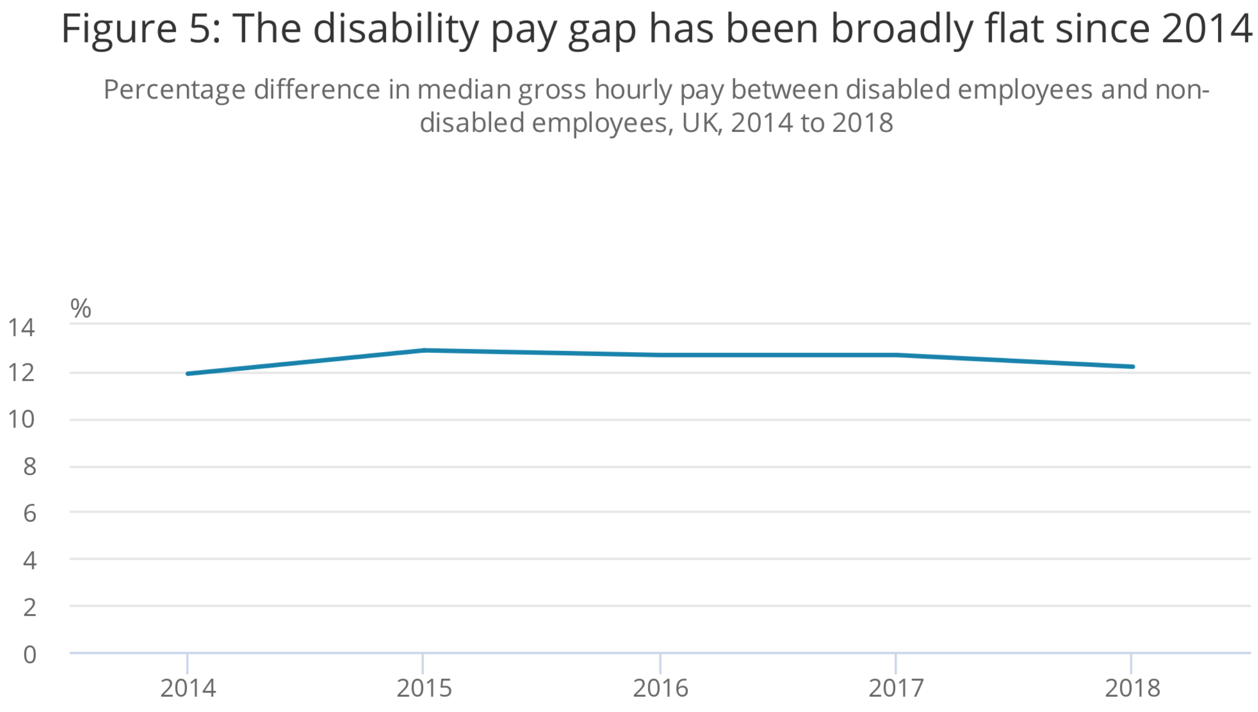
Notes for: Half of disabled people aged 16 to 64 years were employed

1. The definition of economically inactive is found in a guide to labour market statistics.

6 . Disabled employees earned on average 12.2% less than non-disabled people

Figure 5: The disability pay gap has been broadly flat since 2014

Percentage difference in median gross hourly pay between disabled employees and non-disabled employees, UK, 2014 to 2018

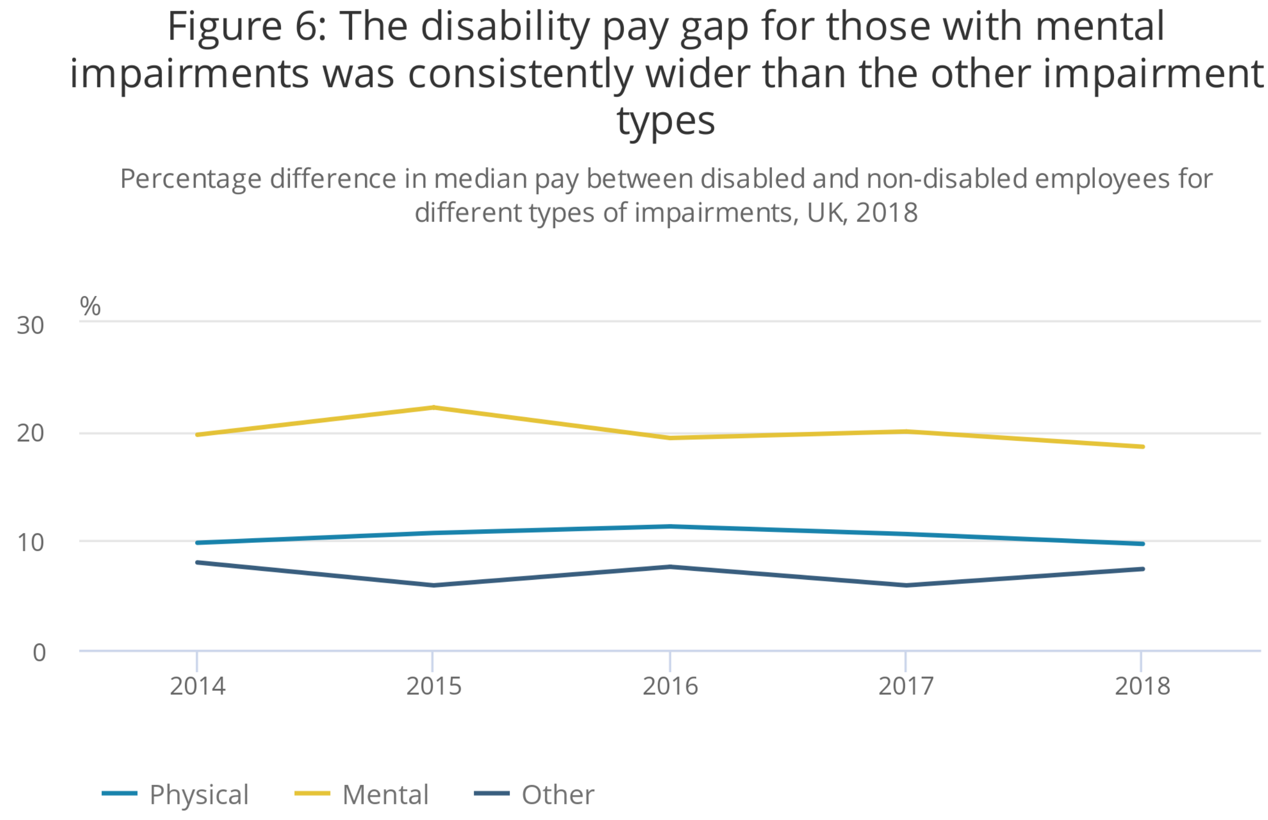


Source: Office for National Statistics – Annual Population Survey

In 2018, average pay for disabled employees was 12.2% lower than average pay for non-disabled employees. In monetary terms, for every £1 that a non-disabled employee would earn in 2018, on average a disabled employee would earn 88 pence. Since 2014, the disability pay gap has remained relatively stable, mostly varying between 12% and 13% (Figure 5).

Figure 6: The disability pay gap for those with mental impairments was consistently wider than the other impairment types

Percentage difference in median pay between disabled and non-disabled employees for different types of impairments, UK, 2018



Source: Office for National Statistics – Annual Population Survey

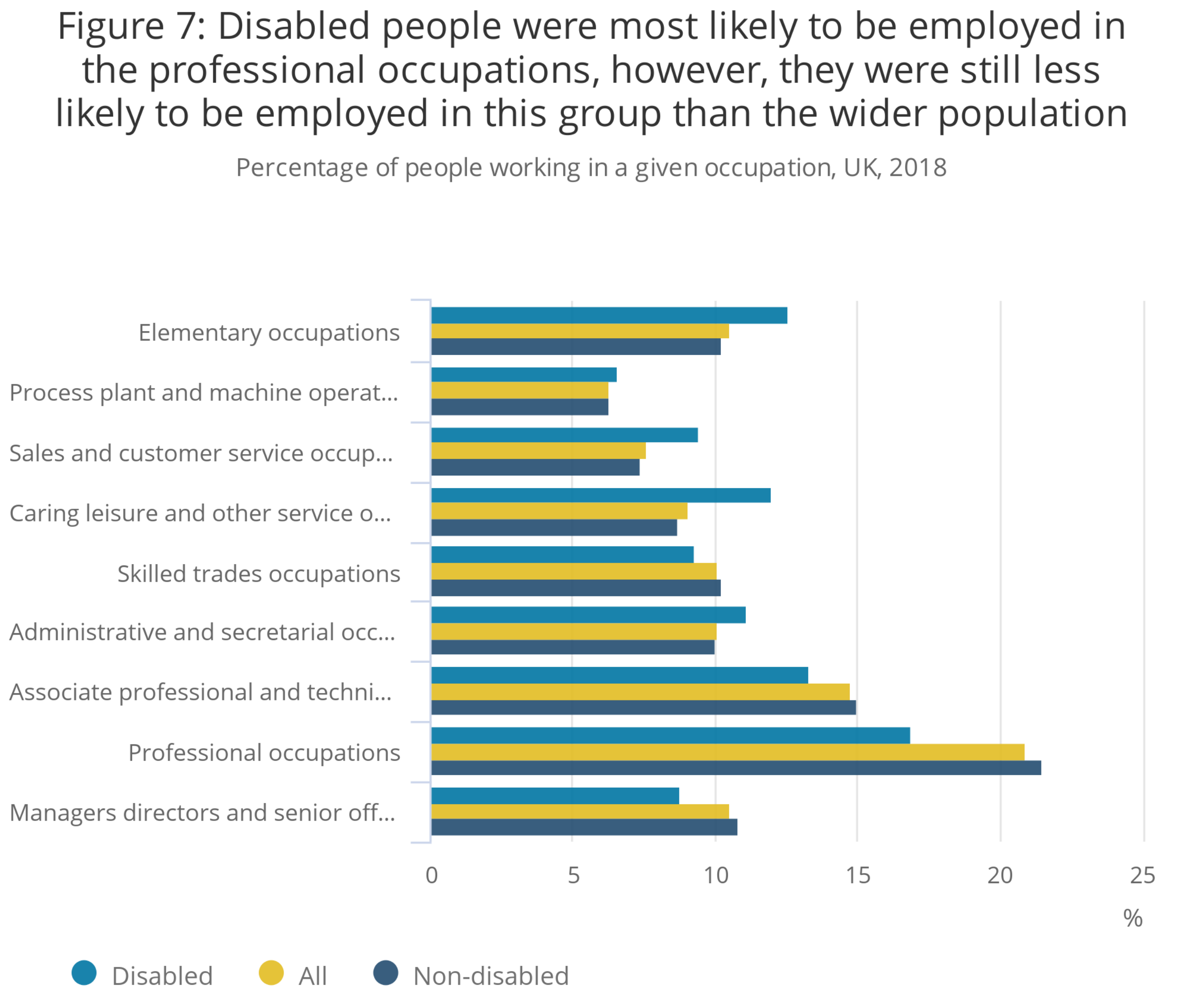
In 2018, the pay gap was widest for those with a mental impairment, at 18.6% (Figure 6). This was followed by those with a physical impairment at 9.7%, whilst those with other impairments had the narrowest pay gap of 7.4%.

Across all three impairment types, there is little evidence of consistent change since 2014. These gaps reflect differing average pay for employees by impairment type. In 2018, employees with mental impairments earned £9.82 per hour whereas those with physical impairments earned £10.90 per hour and those with other impairments earned £11.18 per hour.

Previous Office for National Statistics (ONS) analysis of both the gender and ethnicity pay gaps has shown that differences in occupation often have an important role to play in explaining pay gaps. Figure 7 suggests occupational differences are important in understanding disability pay gaps too.

Figure 7: Disabled people were most likely to be employed in the professional occupations, however, they were still less likely to be employed in this group than the wider population

Percentage of people working in a given occupation, UK, 2018



Source: Office for National Statistics – Annual Population Survey

In 2018, disabled employees were generally under-represented, compared with non-disabled employees, in the higher skilled and typically higher paying occupation groups. For example, looking at the largest occupation group "professional occupations", 21.5% of non-disabled employees held occupations within this group, whereas only 16.9% of disabled employees worked within this group. Conversely, disabled employees had higher than average representation in the lower skilled and typically lower paying occupation groups such as elementary occupations.

Taken together, these differences in representation across the different occupation groups may help to explain some of the differences in average pay between disabled and non-disabled employees. These over- and under-representations might be related to education outcomes for disabled people being different to non-disabled people. For example, analysis published on 2 December 2019 shows [disabled people were less likely to have a degree] than non-disabled people in 2018; this may then have the potential to limit their choice of occupations.

Whilst exploring occupational differences we can also look at disability pay gaps within occupation groups.

Figure 8: The pay gap is largest for managers, directors and senior officials in 2018

Median hourly earnings and percentage difference in median pay between disabled and non-disabled employees for occupation, UK

Download the data

Across the occupation groups, the widest pay gap of 13.1% was seen for managers, directors and senior officials. The remaining occupation groups have disability pay gaps in the range of 1 to 5%.

Managerial occupations are split between high skilled and upper middle skilled [1]. In 2018, disabled people in a managerial occupation were more likely than non-disabled people to be employed in the upper middle skilled occupations rather than high skilled occupations.

Recent analysis of the latest Annual Survey of Hours and Earnings (ASHE) data show that occupations classed as high skilled are paid more than those classed as upper middle . Some examples of high skilled occupations within this group would be a finance director or senior police officer, whilst upper-middle skilled occupations would include shopkeepers and (wholesale and retail) proprietors or property, housing and estate managers. We are unable to make this comparison for other occupation groups as all occupations within them are classified to the same skill level.

The narrowest pay gaps were seen in the low and lower-middle skill occupation groups such as elementary or caring, leisure and other service occupations. We have already seen that these groups were where disabled employees were more likely to be employed than non-disabled employees.

Looking at the median pay values in Figure 8, these narrowest pay gaps were likely to be influenced by the National Living Wage, which for the financial year (April to March) ending 2019 was £7.83 (for those aged 25 years and over).

Given the lower levels of median pay in these occupation groups, the National Living Wage should provide a legal lower bound for hourly pay. This would compress the distribution of pay for disabled and non-disabled employees alike, reducing the scope for pay to diverge in comparison with higher-paid occupation groups. Previous analysis (see Figure 4) of the Labour Force Survey by the ONS showed that there was a greater density of disabled employees immediately around the National Living Wage, compared with non-disabled employees.

We will explore the effects of occupation alongside other characteristics further in our regression and decomposition analysis in Section 8.

Notes for: Disabled employees earned on average 12.2% less than non-disabled people

1. Occupations can be broken down into four skill levels: high, upper middle, lower middle and low skilled.

7 . Disabled people in London were paid on average 15.3% less than non-disabled people

In addition to occupation, we have also produced pay gaps by other personal and demographic characteristics.

Figure 9: London had the largest disability pay gap in 2018

Percentage difference in median pay between disabled and non-disabled employees, UK

Source: Office for National Statistics – Annual Population Survey

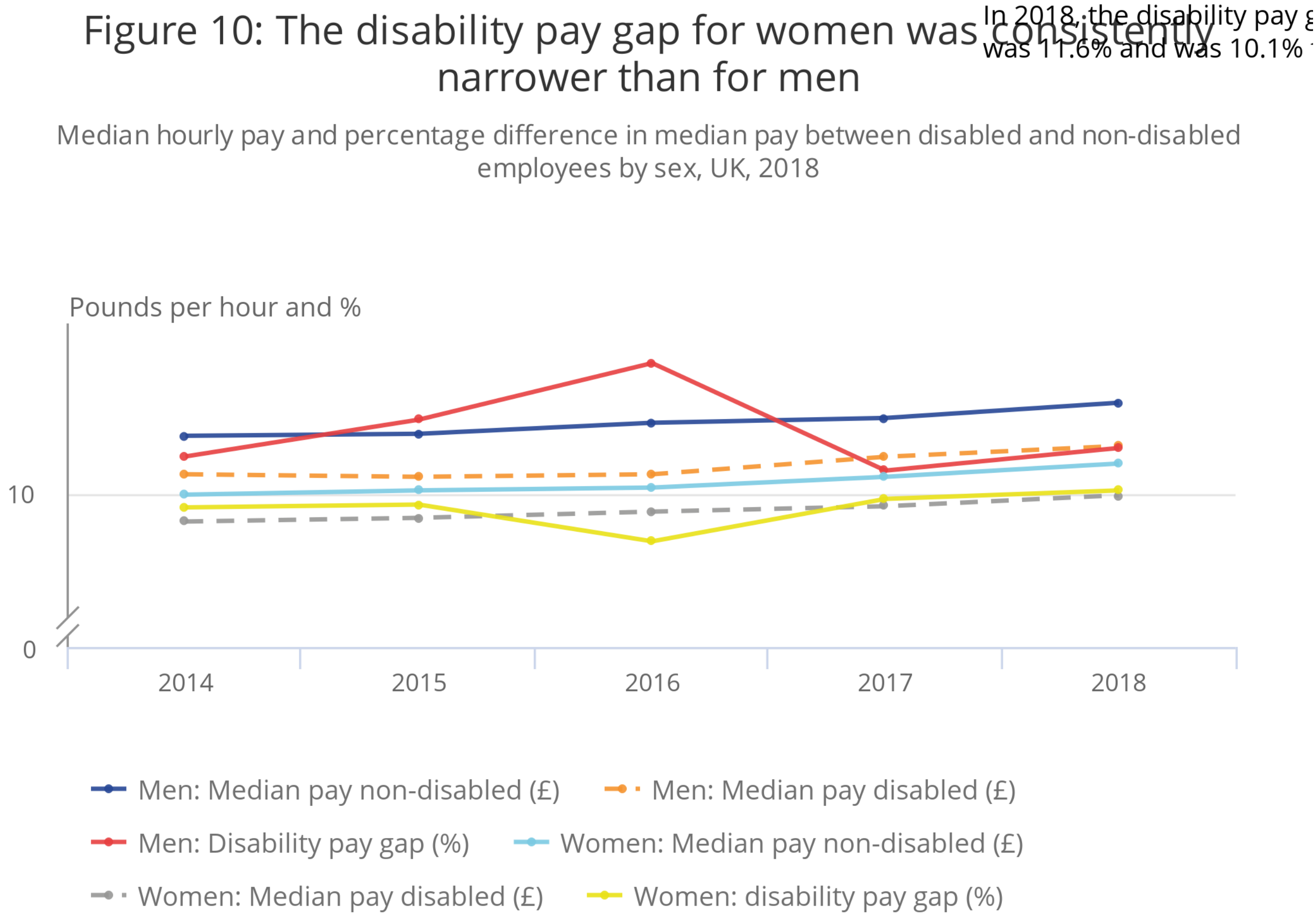
In 2018, disabled employees living in London were paid 15.3% less than non-disabled employees (Figure 9). Since 2014, London has always had one of the three widest pay gaps across the English regions and constituent countries of the UK.

Scotland had the narrowest pay gap in 2018, at 8.3%. However, the data for Scotland do show some volatility as in 2017 Scotland had the second-widest pay gap at 13.3%. This makes it difficult to identify consistent change over time.

Over the same period, Northern Ireland consistently had one of the narrowest pay gaps: in 2014 disabled employees living in Northern Ireland were paid 5.7% less than non-disabled employees. Though this has widened to 8.7% less in 2018, Northern Ireland still had the second-narrowest pay gap. A time series for all English regions and UK constituent countries can be seen in the datasets.

Figure 10: The disability pay gap for women was consistently narrower than for men

Median hourly pay and percentage difference in median pay between disabled and non-disabled employees by sex, UK, 2018



Source: Office for National Statistics – Annual Population Survey

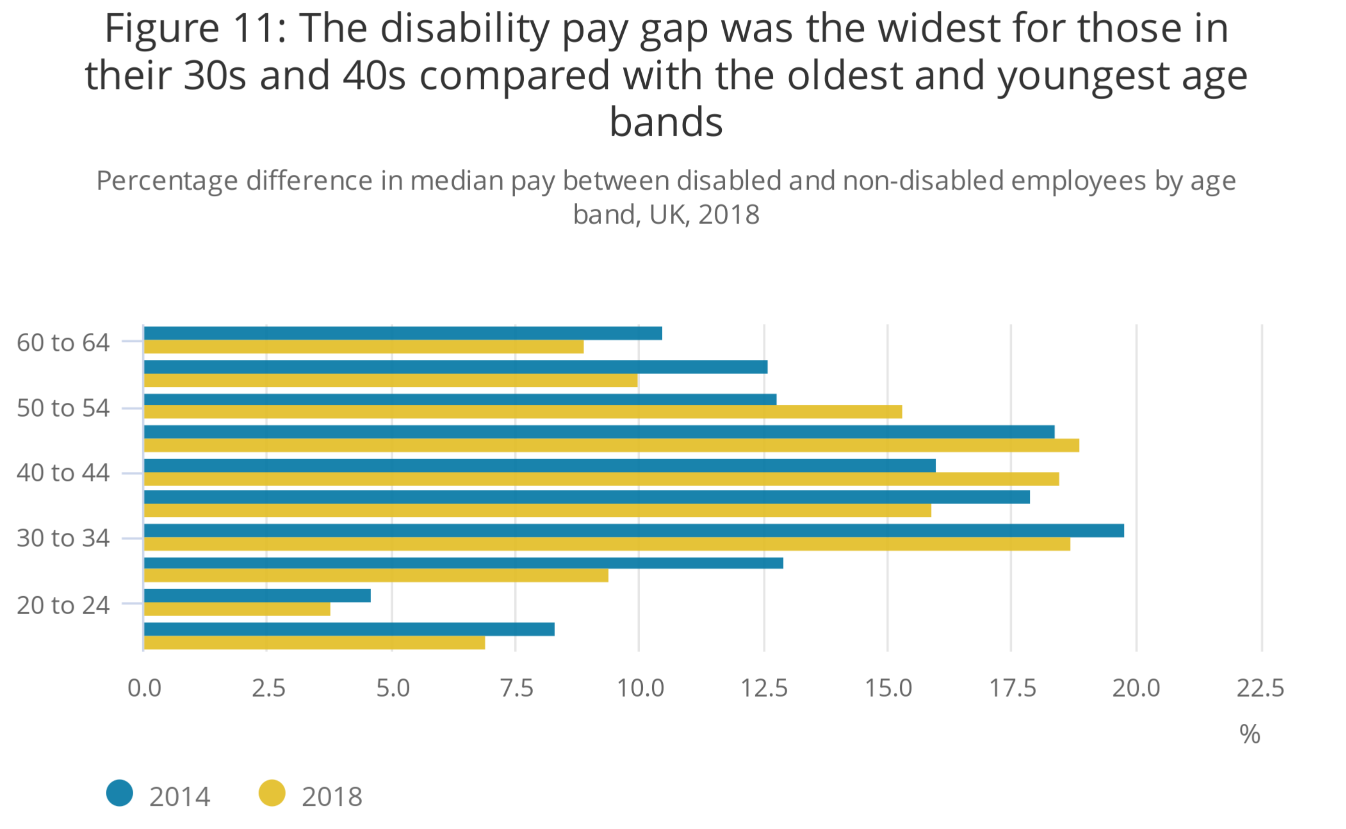
Disabled female employees were paid on average 10.1% less than non-disabled female employees in 2018; this was narrower than the pay gap between disabled and non-disabled male employees who had a pay gap of 11.6% (Figure 10).

This relates to disabled male employees being paid on average £11.67 an hour whilst non-disabled male employees were paid on average £13.20 per hour. For disabled women, average pay was £9.93 an hour, whereas for non-disabled women average pay was £11.05 an hour.

Over time, though we see some volatility for men[1], the trend has been broadly flat since 2014. Although the disability pay gap for women was smaller than for men, over the time series, women were consistently paid lower than men on average.

Figure 11: The disability pay gap was the widest for those in their 30s and 40s compared with the oldest and youngest age bands

Percentage difference in median pay between disabled and non-disabled employees by age band, UK, 2018



Source: Office for National Statistics – Annual Population Survey

The disability pay gap also varies with age. We see that pay gaps were at their narrowest for the youngest age groups (Figure 11). Employees within these age groups may be employed in more casual employment or will be at a relatively early stage of their career where average pay tends to be lower. Again, the National Living Wage could be providing a lower bound for earnings in these age groups for disabled and non-disabled employees, reducing the scope for wider pay gaps.

The widest gaps were seen for employees aged in their 30s and 40s, where average pay typically peaks for most employees.

Since 2014, we have seen a narrowing of the pay gap across most age groups, and in particular for the youngest age groups. However, this has been offset to a certain extent by a widening of pay gaps for those aged in their late 40s and early 50s.

We have also published disability pay gaps by ethnicity, working pattern and employment sector in the associated [data download](). The datasets also contain a time series for each breakdown back to 2014.

Notes for: Disabled people in London were paid on average 15.3% less than non-disabled people

1. The widening of the pay gap for men in 2016 was because of non-disabled men having an increase in average hourly earnings of 39 pence between 2015 and 2016. This was followed by an increase in average earnings for disabled men by 62 pence between 2016 and 2017 causing the pay gap to narrow.

8 . Adjusting for factors that affect the level of pay narrows the pay gap

There are many different factors that affect the level of pay a person receives. These factors then contribute to the pay gaps observed. For instance, we have already seen that differences in occupation influence average pay for disabled and non-disabled employees.

Given that the population of disabled employees has different characteristics on average compared with the non-disabled employee population, and that these factors affect the level of pay, it would be beneficial to try and control for these effects to try to isolate the impact that disability status has on pay.

Regression is a statistical technique that allows us to model the relationship between a dependent variable, for example, hourly pay, and one or more explanatory variables, for example, disability status, age or sex. In this analysis we have modelled log hourly pay as our dependent variable against the following explanatory variables:

- disability status
- impairment
- age and age squared - proxies for experience
- sex
- working pattern
- ethnicity
- region
- sector
- marital status
- occupation
- qualification status
- country of birth

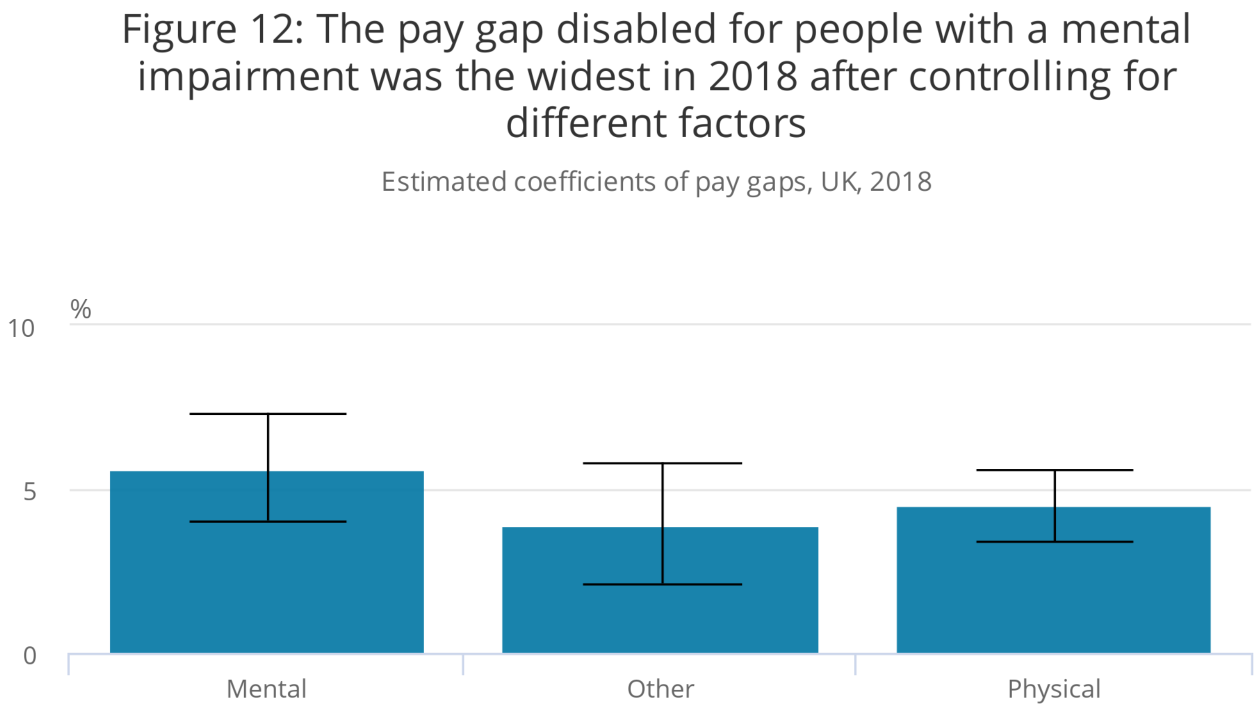
It is recognised these independent variables do not cover all the factors that can affect a person's pay, but they are the most relevant of the ones available on the dataset. An assessment of the overall model quality and specification is given in the Quality and methodology section. Modelling in this way allows us to control for the effects of the explanatory variables on pay. In effect we can try to observe the pay gap between disabled and non-disabled employees under the assumption that they have otherwise identical characteristics.

The model uses the statistical estimation method of ordinary least squares, which estimates the percentage difference of mean pay rather than median pay[1]. To adjust for means being affected more by extreme outliers than medians, we have removed the top 1% and bottom 2% of the pay distribution for all employees.

Keeping all other factors constant, disabled people with a mental impairment still experienced the widest pay gap of 5.6%, whilst those with a physical impairment experienced a pay gap of 4.5% and for those with other impairments the gap was 4.9% (Figure 12). Whilst we can observe that these adjusted pay gaps are statistically significantly different from 0 in their own right, they are not statistically significantly different from each other.

Figure 12: The pay gap disabled for people with a mental impairment was the widest in 2018 after controlling for different factors

Estimated coefficients of pay gaps, UK, 2018



Source: Office for National Statistics – Annual Population Survey

Compared with the “raw” pay gaps discussed in Section 6, we see narrower, though still statistically significant, pay gaps across all three impairment types. This narrowing of the gaps is at least in part because we are now controlling for an array of other characteristics, which can vary between disabled and non-disabled employees.

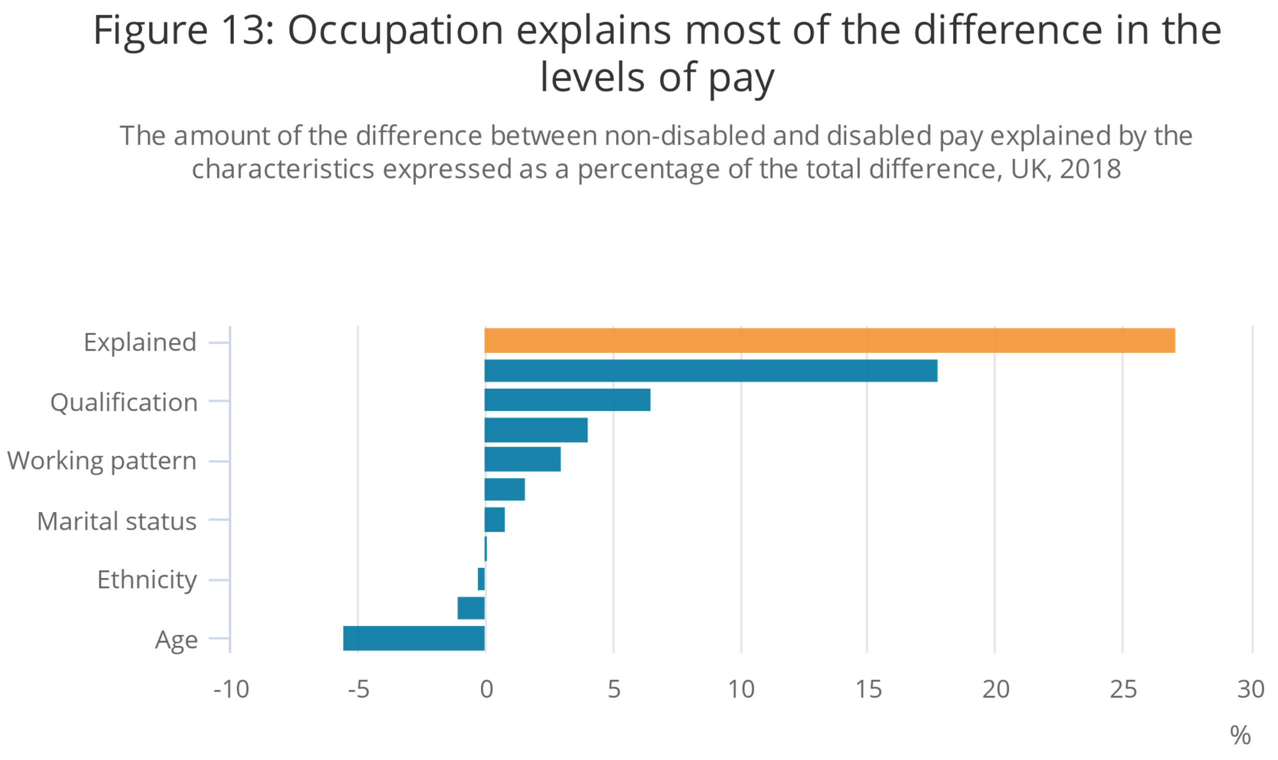
To determine which of these characteristics has most influence on the pay gap we can use the Blinder-Oaxaca decomposition. This decomposes the pay gap into what can be explained by the different average characteristics of disabled and non-disabled employees, and what cannot be explained by differences in average characteristics.

The explained element assumes that each characteristic has the same impact on pay for disabled and non-disabled employees and that differences in average pay are caused solely by the different average levels of that characteristic held by disabled and non-disabled employees. For example, if we assume that growing a year older has the same effect on pay for a non-disabled person as it does on pay for a disabled person (all other characteristics held constant), the decomposition then calculates whether or not disabled people would earn more or less on average, given that on average they are older. Once this is calculated for each characteristic in the model, the remainder of the pay gap falls within the unexplained component.

The components in Figure 13 reflect each characteristic's contribution to explaining the difference in log hourly earnings between non-disabled and disabled people. A positive value reflects that non-disabled people have favourable characteristics (in terms of pay), whereas the opposite is true for a negative value, whereby disabled employees on average possess more favourable (in terms of pay) levels of a characteristic than non-disabled employees. We have removed impairment type from the decomposition as all disabled people have an impairment whilst a small proportion of the non-disabled population report having a non-limiting health condition (about 15%).

Figure 13: Occupation explains most of the difference in the levels of pay

The amount of the difference between non-disabled and disabled pay explained by the characteristics expressed as a percentage of the total difference, UK, 2018



Source: Office for National Statistics – Annual Population Survey

Thus 27.1% of the difference between the mean non-disabled pay and disabled pay can be explained by the characteristics we have used to model pay. The largest positive contribution to the difference in pay came from occupation, explaining 17.8% of the gap, reflecting the findings seen in Figure 7. Other important factors include qualification, explaining 6.5%. This is supported by Office for National Statistics (ONS) analysis also published on 2 December 2019 showing that non-disabled people on average had higher level qualifications than disabled people.

The largest negative contribution to the decomposition comes from age. Given that pay tends to increase with age, this suggests that disabled employees are older on average than non-disabled employees (the mean age of a disabled person was 44.2 and for a non-disabled person 39.1). All else equal, we would therefore expect disabled employees to out-earn their non-disabled counterparts.

We have only explained 27.1% of the gap between disabled and non-disabled pay by the characteristics in the model. 72.9% of the gap that is still unexplained. Adding extra characteristics to the model may help us to explain more, for example, more information on the characteristics of the employee such as their skills and more detailed information on their experience, however, these are not available on the dataset.

We have also tested whether the proportion of the pay gap that can be explained is different when comparing non-disabled employees with employees who specifically have a work-limiting disability. Our findings show that there is a slight increase in the portion of the pay gap that is explained (from 27.1% to 28.9%); however, there is little difference in the relative importance of specific characteristics.

Differences in average occupation and education are still the largest and most important measurable factors, which can explain the pay gap.

Notes for: Adjusting for factors that affect the level of pay narrows the pay gap

1. Mean pay is an average where all the individual pay recorded is added together and divided by total number of respondents, whilst the median pay is where we order all the pay from lowest to highest and take the value of the pay that is in the middle.

9 . Author and acknowledgements

Author: Matthew Mayhew

The author would like to thank the following people for their advice and contributions to this article: Tom Evans, Andrea Lacey, Hamish Anderson, Mike Welsby, Harry Beeching, Tom Odell, Helen Colvin, Catherine Putz and Vahé Nafilyan as well as members of the Cross-Whitehall analytical group for disability.

10 . References

Anna Ardinaz-Badia and Josh Rawlings (2018). Understanding the gender pay gap in the UK. ONS

Hamish Anderson (2019). Ethnicity pay gaps in Great Britain: 2018. ONS

A. S. Blinder (1973). Wage Discrimination: Reduced Form and Structural Estimates. Journal of Human Resources 8(4), pages 436 to 455

S. Gaure. lfe: Linear group fixed effects. R package version 2.8-3, 2019

Frank E Harrell Jr, with contributions from Charles Dupont and many others. (2019). Hmisc: Harrell Miscellaneous. R package version 4.2-0.

Marek Hlavac (2018). oaxaca: Blinder-Oaxaca Decomposition in R. R package version 0.1.4.

Joel Jones (2019). Disability, well-being and loneliness. ONS

Joel Jones (2019). Disabled people and their participation in society. ONS

T. Lumley (2019) "survey: analysis of complex survey samples". R package version 3.35-1.

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria

R. Oaxaca (1973). Male-Female Wage Differentials in Urban Labor Markets. International Economic Review. Volume 14 Number 3, pages 693 to 709

Office for National Statistics (2018). Labour Force Survey - user guidance

Catherine Putz (2019). Disability and employment. ONS

Catherine Putz and Myer Glickman (2019). Measuring disability: comparing approaches. ONS

Indiana Sparkes (2019). Disability and education. ONS

Alexander Walker (2019). openxlsx: Read, Write and Edit XLSX Files. R package version 4.1.0.1.

Hadley Wickham (2016) ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York

Hadley Wickham (2019). forcats: Tools for Working with Categorical Variables (Factors). R package version 0.4.0.

Hadley Wickham, Romain François, Lionel Henry and Kirill Müller (2019). dplyr: A Grammar of Data Manipulation. R package version 0.8.3.

Hadley Wickham, Jim Hester and Romain Francois (2018). readr: Read Rectangular Text Data. R package version 1.3.1.

Hadley Wickham and Evan Miller (2019). haven: Import and Export 'SPSS', 'Stata' and 'SAS' Files. R package version 2.1.1.

11 . Quality and methodology

Data sources

Though this analysis makes use of the Annual Population Survey (APS) it should be noted that the primary source of data for earnings analysis in the UK is the Annual Survey of Hours and Earnings (ASHE). As a business survey, ASHE collects detailed information on the composition and distribution of earnings among employees, however, as a business survey, it collects only a limited range of personal characteristics regarding individual employees. This limits its usefulness in analysing earnings for instance by education and/or by different protected characteristics including disability.

As a result, the Labour Force Survey (LFS) (a quarterly version of the APS) is still heavily used as a source of data on earnings. Though it is accepted that the accuracy and detail of earnings information captured by the LFS falls short of that obtained by ASHE, the greater range of personal and household characteristics broaden its potential uses. However, one drawback of earnings analysis on the LFS is that the achieved sample is relatively small. This is because earnings questions are asked only to employees and only in 40% of the interviews carried out in each quarter.

Furthermore, earnings questions on the LFS are known to have poor response rates. The achieved sample for the LFS earnings questions is usually around 9,000, compared with approximately 150,000 respondents on ASHE. This limited sample size then restricts the extent to which you can perform multivariate analysis of earnings on the LFS, particularly where the variables of interest have many categories.

Therefore, for the analysis of earnings presented in this article, a new income weight has been calculated for the APS. The APS combines responses from the quarterly LFS and Annual Local Labour Force Surveys for England, Wales and Scotland. Though the APS has always collected information on earnings, until now there has never been an appropriate weight included for earnings analysis.

The income weight is calculated in a similar way to the LFS income weight. More information on this can be found in the Volume 6 LFS user guide. The main differences are that there are six calibration groups used to calculate the APS income weight, while for the LFS income weight there are four.

Finally, it should also be noted that, though the APS has a much-improved sample size compared with the LFS, it still suffers from some shortcomings when compared with ASHE. For instance, as a survey of businesses, ASHE is thought to capture more accurate earnings information as employers can consult payroll records when responding to the survey. In comparison, earnings information collected in the LFS and APS is self-reported and as such is likely to be subject to a higher degree of recall error.

Methodology

Model specification (variable name)

The variables included in the model are:

- log of hourly pay log(hourpay)
- disability status (disea)
- impairment (health)
- age (age)
- age^2 (agesq)
- sex (sex)
- working pattern (ftpt)
- ethnicity (ethukeul)
- region (govtof)
- sector (publicr)
- marital status (marsta)
- occupation (sc10mmj)
- qualification status (hiqul15d)
- country of birth (cryox7)

Given that most of the variables are categorical variables, we have included them as dummy variables. We must exclude one of the levels to avoid perfect multi-collinearity. Collinearity is where one of the variables can be derived from the rest of the variables in the model. For example, if our categorical variable has three levels, A, B, and C, then if our model is:

y= constant + A+B+C

then A can be derived if B and C are set to zero. In this example A is our reference level.

The reference levels in our model are:

- disability status – not Equality Act disabled
- impairment – no impairment
- sex – male
- working pattern – full-time
- ethnicity – white
- region – London
- sector – public
- marital status – single, never married
- occupation – professional occupations
- qualification status – degree or equivalent

Reference levels were selected if they satisfy one of two criteria, either they had the highest proportion of respondents in that category, for example, disability status and occupation, or they are natural choice that would aid interpretation, for example, marital status and qualification status.

The dependent variable is the log of hourly pay. This is because the distribution of pay is positively skewed, there is a higher density in the lower values of pay than in the higher values of pay. Taking the log of the hourly pay helps to make the distribution more symmetric and like a normal distribution so the assumption used in regression are more valid. The rest of the variables in the model are the independent variables.

Both age and age squared are used in the model to approximate for a non-linear relationship between age and log(pay). A linear relationship between age and pay would infer that, for each year old a person gets, their pay would on average increase by the same amount. This is not the case in the APS data.

Some of the independent variables might interact with each other. Interaction means that effect of one variable is dependent on the values of a second variable. When this happens, we add a term to the model, which the two variables of interest are multiplied. For example, if variable x_1 interacts with variable x_2 then the model is as follows:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + \varepsilon$$

Adding interaction terms to a model drastically changes the interpretation of the all the coefficients in the model. If we have no interaction, then $_1$ would be interpreted as the unique effect that x_1 has on y. But with the interaction the effect that x_1 has on y is different dependent on the value of x_2. The effect of x_1 is now not limited to $_1$ but also depends on the values of $_3$ and x_2.

The effect of x_1 is represented by everything that is multiplied by x_1 in the model. This is:

$$\beta_1 + \beta_3 x_2$$

$_1$ is now interpreted as the effect of x_1 on y only when x_2=0

For our model we have interacted:

- disability with impairment
- sex with working pattern
- ethnicity with country of birth

We have estimated our regression model using survey-weighted ordinary least squares (SWOLS) method. Ordinary least squares estimates the coefficients of the model by minimising the squares of the residuals (the difference between our model estimates and the observed value). SWOLS extends this to take into account the complex sample design of the APS so that our estimates are representative of the population and not our sample.

We need to take some care when interpreting the coefficients for each variable. As the dependent variable is log transformed, the coefficient is the effects on the log-scale of the variable. To interpret the coefficient in a meaningful way we take the exponential of the coefficient, which then can be interpreted as the percentage change in the level of pay.

For example, if the estimate of the coefficient of x_1 is 0.1 then the effect on pay is:

exp (0.1) = 1.105

which shows that each additional unit of x_1 increase pay by 10.5% with all other variables held constant.

However, there are caveats that must be considered when interpreting estimates using the OLS method. For example, predictor variables will have been excluded from the model because of their unavailability in the model, for example family background. These excluded variables will have an effect on the explanatory power of the model. It might be possible that the functional form of the model could be improved, for example, the relationship between log(hourly pay) and age, which would improve the accuracy of our estimates.

Blinder-Oaxaca decomposition

The Blinder-Oaxaca decomposition uses the following property of regressions to estimate the mean log hourly pay for disabled and non-disabled people:

$$\overline{\log y^D} = \hat{\beta}_0^D + \sum_k \hat{\beta}_k^D \bar{x}_k^D$$

$$\overline{\log y^{ND}} = \hat{\beta}_0^{ND} + \sum_k \hat{\beta}_k^{ND} \bar{x}_k^{ND}$$

Where y^D is the pay for disabled people, ^{D} are the estimated coefficients for disabled people and x are the characteristics used to model pay. Those indexed with ND are for non-disabled people.

Our analysis takes the effect of the characteristic on pay regardless of disability status as the baseline and the “non-discriminatory” wage. With this we estimate a counterfactual wage, what disabled people would earn on average if the characteristics had the same effect on their pay as they did on non-disabled pay as follows:

$$\overline{\log y^D}^* = \hat{\beta}_0^{ND} + \sum_k \hat{\beta}_k^{ND} \bar{x}_k^D$$

We can use this to decompose the difference between mean pay for non-disabled people and disabled people () as follows:

= U + E

where:

$$U = \overline{\log\ y^{D*}} - \overline{\log\ y^D}$$

$$E = \overline{\log\ y^{ND}} - \overline{\log\ y^{D*}}\ U \text{ is the unexplained part and } E \text{ is the explained part.}$$

In practice we do not need to calculate the counterfactual wage as the unexplained part and explained part can be calculated directly from the two regression equations:

$$U = \left(\hat{\beta}_0^{ND} - \hat{\beta}_0^D\right) + \sum_k \left(\hat{\beta}_k^{ND} - \hat{\beta}_k^D\right) \bar{x}^D$$

$$E = \sum_k \hat{\beta}_k^{ND} \left(\bar{x}_k^{ND} - \bar{x}_k^D\right)$$