The Representativeness of the Annual Survey of Hours and Earnings and its Implications for UK Wage Policy

John Forth[1,6,7], Alex Bryson[2,6,7], Van Phan[3], Felix Ritchie[3],
Carl Singleton[4,7], Lucy Stokes[5] and Damian Whittard[3]

August 2025
Forthcoming in British Jounral of Industrial Relations

Abstract

The Annual Survey of Hours and Earnings (ASHE) is based on an annual one per cent sample of employee jobs and provides many of the UK's official earnings statistics. These statistics are produced using official weights designed to make the achieved sample in each year representative of the population of employee jobs in Britain by gender, age, occupation, and region. However, we show that jobs in small, young, private-sector organisations remain significantly under-represented after applying these weights. To address this issue, we develop new weights and demonstrate their importance through policy-relevant examples. Our new estimates suggest that the bite of the National Living Wage is greater than previously reported, and the gender pay gap is wider. We conclude that a new official review of the methodology for ASHE is merited, to improve the accuracy and reliability of data informing earnings analysis and research in the UK.

Keywords: Earnings; Non-response bias; Attrition; Survey weighting; Low pay; National Living Wage; Gender pay gap

JEL codes: C81, C83, J31

[1] Corresponding author. Email: john.forth@citystgeorges.ac.uk. Faculty of Management, Bayes Business School, City St Georges, University of London, 106 Bunhill Row, London, EC1Y 8TZ, United Kingdom.
[2] University College London, UK.
[3] University of the West of England, UK.
[4] University of Stirling, UK.
[5] Competition and Markets Authority, UK.
[6] Research Associate, Economic Statistics Centre of Excellence (ESCoE), UK.
[7] Research Fellow, Institute of Labor Economics (IZA), Germany.

We gratefully acknowledge funding from ADR UK (Administrative Data Research UK) and the Economic and Social Research Council (Grant No. ES/T013877/1). The work is based on analysis of the research-ready datasets from the *Annual Survey of Hours and Earnings (ASHE)* and *Business Structure Database*, accessed through the ONS Secure Research Service. The use of the ONS data in this work does not imply the endorsement of the ONS or data owners in relation to the interpretation or analysis of the statistical data. This work uses research datasets which may not exactly reproduce National Statistics aggregates. National Statistics follow consistent statistical conventions over time and cannot be compared to these findings. We thank Susan Purdon, Eduin Latimer, and participants at the 2023 Work and Pensions Economics Group Conference, 2023 Low Pay Commission Research Symposium, and 2024 Annual Conference of the Economic Statistics Centre of Excellence (ESCoE) for valuable comments. Lucy Stokes completed this research whilst employed at the National Institute of Economic and Social Research, UK. This work represents the views of the individual authors, and not the views of the Competition and Markets Authority. The authors have no interest or relationship, financial or otherwise, that might be perceived as influencing their objectivity. The ASHE data that support the findings of this study are available to accredited researchers through the ONS (Office for National Statistics) Secure Research Service at http://doi.org/10.57906/x25d-4j96. The Business Structure Database is available via the same route at https://doi.org/10.57906/7kh0-0910. Stata code to derive the adjusted weights is available within the ONS Secure Research Service and may also be downloaded from the website of the Wage and Employment Dynamics Project (https://www.wagedynamics.com). The project involves the use of anonymised personal data and received ethical approval from the University of the West of England.

1 Introduction

The Annual Survey of Hours and Earnings (ASHE), conducted by the UK Office for National Statistics (ONS), is an important source of official statistics on earnings and working hours in the UK labour market (see, for example, ONS, 2023a, 2023b, 2024b). ASHE is also widely used to inform policymaking and policy evaluation across government. It is used *inter alia* by the Low Pay Commission (LPC) to monitor the impact of the minimum wage, by the Office for Equality and Opportunity to monitor relative pay levels for employees with protected characteristics, and by the Office for the Pay Review Bodies to inform public sector pay settlements. The ASHE dataset is also an important resource for academic research into the UK labour market (e.g. Stokes et al., 2017; Jewell et al., 2020; Forth et al., 2025). The issued sample for ASHE comprises a one per cent simple random sample of employee jobs, drawn from tax and social security records. Information on each sampled employee's earnings and hours is provided by their employer, who is statutorily required to respond. Nevertheless, the annual yield from the issued sample is typically around two-thirds.

ASHE was introduced in 2004 to replace the New Earnings Survey (NES) after an official review identified that the NES was unrepresentative of the population of employee jobs (ONS, 2002; Bird, 2004). Although ASHE shares many common features with the NES, a key innovation was the introduction of cross-sectional survey weights, which adjust the profile of the achieved sample so that it is representative of the population of employee jobs in terms of gender, age, occupation, and region. However, the absence of any employer characteristics from the weighting scheme raises the possibility that the achieved sample may have residual biases, causing it to over- or under-represent jobs from particular types of employers. We link data from ASHE to the UK's official business register to investigate this issue. We find that jobs in smaller organisations, younger organisations and those in the private sector are under-represented in the annual samples from ASHE relative to their prevalence in the wider economy, even after applying the official ASHE weights. To address this, we use a raking approach (Deville et al., 1993) to derive new cross-sectional weights that take account of these observed biases.

We use these new cross-sectional weights to re-estimate two important and highly policy-relevant sets of earnings statistics for the UK. First, we re-estimate the coverage rate and bite of the National Minimum Wage and National Living Wage. ASHE is the main data source used by the ONS and the Low Pay Commission to estimate the incidence of low pay, making it a vital component in decision making over the future level of the minimum wage and broader policies aimed at supporting living standards. We find that the percentage of jobs paid at or below the National Living Wage (NLW) is under-estimated by around one fifth if the cross-sectional response biases that we identify in ASHE are not accounted for. The bite of the NLW is also under-estimated, such that the Government's targets for this measure have been reached more quickly than previously thought. Second, we re-estimate the size of the gender pay gap. ASHE is not only used by ONS to report national statistics on the UK gender pay gap but also by Eurostat and others for cross-country comparisons, such as in the World Economic Forum's annual Gender Gap Report (e.g., Pal et al., 2024). Our findings suggest that the UK gender pay gap has been consistently under-estimated over the past 20 years, by a small but noteworthy margin of around one percentage point.

Our findings contribute to the literature on rates and patterns of low pay and gender pay inequality in the UK (e.g. Dickens et al., 2015; Aitken et al., 2019; Low Pay Commission, 2022; Giupponi et al., 2024; Jewell et al., 2020; Jones and Kaya, 2023). They also contribute to a broader literature on the nature, detection, and removal of non-response biases in business surveys (e.g. Willimack et al., 2002; Willimack and Snijkers, 2013). More broadly, our findings contribute to concerns about the quality of UK labour market statistics (Giles, 2023; Office for Statistics Regulation, 2023, 2024) and indicate that a new official review of the methodology for ASHE is merited, to improve the quality of data informing earnings analysis and research in the UK.

2 Background

When introduced in 2004, ASHE used the same sampling frame as the NES and collected many of the same data items. However, key elements of the survey methodology were changed in line with recommendations made in the National Statistics Quality Review of the Distribution of Earnings Statistics (ONS, 2002). The review identified that statistics generated from the NES were likely to be biased because the survey missed significant numbers of employees that changed job during the three months that typically elapsed between sample selection in January and the survey reference date in April. Additionally, responses to the NES were not weighted to the population of employee jobs. The revised sample design and weighting approach developed for ASHE aimed to address both these issues, with the explicit aim of making ASHE the definitive source on the distribution of earnings in the UK (see Bird, 2004; Pont, 2007).

2.1 ASHE sample design

The target population for ASHE is all employee jobs in the UK, across all industries and occupations. The sample for the survey is drawn from the UK's official Pay-As-You-Earn (PAYE) register. PAYE is the system used by HM Revenue and Customs (HMRC) to collect income tax and social security contributions for employee jobs. Employers are legally required to operate PAYE if the earnings of any of their employees reaches the National Insurance (NI) Lower Earnings Limit (£123 per week in 2023/24). They must then report payments and deductions for all their employees to HMRC on or before each payday. The PAYE register, therefore, provides a comprehensive and up-to-date record of employee jobs in the UK. A one per cent sample of jobs is drawn from the register by selecting all PAYE-registered jobs held by employees with a National Insurance (NI) number ending in a particular two digits. If a sample member holds multiple jobs, all are selected.

An initial sample of jobs is selected in January of each year. A second extract is then taken in April to identify instances where a sample member has started a new job or changed employer since the initial sample was drawn. The issued sample may then be considered a one per cent random sample of all employee jobs in existence in the first quarter of the survey year. The sample design also means that individuals with eligible NI numbers are selected into the annual issued sample each year that they are in PAYE employment.

The survey itself is completed by employers. To obtain their contact details, the sample drawn from the PAYE register is matched against the ONS Inter Departmental Business Register (IDBR). Survey questionnaires are typically dispatched to employers in the second half of April and ask for information on the paid working hours and earnings of the sampled employee job, for the pay period that includes a specific reference date (typically the second or third Wednesday of April). A specific date is chosen so that all respondents refer to the same point in time in a given survey year. Some larger organisations have a Special Arrangement (SA) in place with the ONS to provide their data electronically; these employers have internal systems that extract and return information on all relevant employees as of the survey reference date.

Employers are generally asked to return their data within one or two months. Reminders follow a set timetable each year: three reminders are sent to employers who respond via special arrangements (in June and July), and one reminder is sent to all other employers (in June).

Fieldwork for ASHE is conducted for Great Britain by ONS and for Northern Ireland by the Northern Ireland Statistics and Research Agency (NISRA). Whilst some official estimates produced from the survey data cover the UK (e.g. ONS, 2023a, 2023b, 2024b), the research-ready dataset made available to us (ONS, 2024a) only covers Great Britain and so all subsequent discussion in this paper refers to Great Britain only.

2.2 Response rate and weighting scheme

The target population for ASHE in Great Britain rose from around 24.1 million jobs in 1997 to around 31.4 million in 2023. As noted above, ASHE is based on an issued sample comprising one per cent of these jobs. Completion of ASHE is mandatory under the Statistics of Trade Act 1947, but not all sampled businesses respond. Analysis of 2004 data by Pont (2007) found that "good data" were collected for 68% of the issued sample (noting that some returned questionnaires related to individuals exempt from the survey, and that some questionnaires were not useable due to insufficient quality). An ONS review of ASHE in 2010 indicated that the anticipated yield for ASHE (based on the latest survey at the time) stood at 63% of employee jobs (ONS, 2010). Our own calculations show that the yield in Great Britain averaged around 63% across the period from 1997 to 2019 (Figure 1); it has averaged only 46% since the onset of the COVID pandemic.

[FIGURE 1 HERE]

There are limits to the time and resources available to pursue employers to return questionnaires. Pont (2007) reported on the results of two intensive follow-up exercises run in 2003 and 2004. While these exercises managed to yield additional responses, they also revealed that additional chasing is insufficient to persuade a "hardcore" of employers to respond (Pont, 2007: 723). To our knowledge, there are no published figures on the number of employers prosecuted for not responding to ASHE. Information on the completion of ONS business surveys more generally indicates that the ONS Enforcement Unit deals with thousands of cases of non-completion per year, but few of these reach court or result in prosecution. The stated aim of ONS is to encourage and assist employers to comply, rather than penalise, where possible (ONS, 2015).

ONS do not release information on the fieldwork outcomes for individual sample members. However, they derive weights to make the achieved sample cross-sectionally representative and provide these to researchers with the core ASHE dataset. There are two stages in constructing these weights (ONS, 2018). In the first stage, individual cases are assigned a design weight (DWEIGHT) based on whether they belong to: (i) the original questionnaire despatch; (ii) the group that moves jobs between sample selection and questionnaire despatch; (iii) the group that joins the workforce after the original sample selection; or (iv) the group with special arrangements for response (where employers respond electronically). In the second stage, the design-weighted sample is post-stratified to population estimates taken from the UK Labour Force Survey (LFS) across 108 post-strata, based on the cross-classification of:

- occupation (nine groups) – major groups from the 2010 Standard Occupational Classification
- age bands (three groups) – 16 to 21 years, 22 to 49 years and 50 years and over
- sex (two groups) – male and female
- region (two groups) – London and South East, and the rest of the UK.

The LFS is itself weighted to mid-year population estimates derived from the Census and other administrative data sources (ONS, 2024e). Whilst there have been concerns about the quality of the LFS, particularly since the COVID pandemic, various adjustments have been made to the weighting methodology to ensure that it is as representative as possible (ONS, 2020, 2021, 2024f) and these have helped to improve the coherence between the LFS and other labour market indicators (ONS, 2025), such that the LFS remains the best overall source of information on the profile of employees in the UK.

The resulting ASHE weight generated from the two stages outlined above (variable CALWGHT) allows estimates to be produced from the ASHE data that are notionally representative of the population of UK employees. In addition, a low pay weight (LPCALWGHT) reweights the dataset after excluding employees whose pay is affected by absence during the survey reference period (a filter commonly used by the Low Pay Commission in their analysis of minimum wage employment).

In the research community, it is widely agreed that survey weights should be applied to make samples more representative of the population when generating descriptive statistics from surveys that deviate from simple random sampling or are affected by non-response. However, the use of weights in multivariate analyses is more contentious. In such cases, the decision to apply weights must be made on a case-by-case basis, as the potential drawbacks may sometimes outweigh the benefits (see Solon et al., 2015, for a detailed discussion). In this study, we focus exclusively on descriptive population statistics, as most of the official statistics produced from ASHE fall into this category.

3 Data and Methods

3.1 Assessing the representativeness of the annual ASHE cross-sectional samples

As noted earlier, the standard ASHE weighting scheme devised by ONS seeks to ensure that the weighted data are representative of the employee population by occupation, gender, age, and region. However, as the survey is completed by employers, our further investigation into the cross-sectional representativeness of the annual ASHE samples focuses on employer characteristics.

For an initial investigation of whether response rates vary across employer types, we first examine the population of enterprises recorded each year in the Business Structure Database (BSD) (ONS, 2024c; Evans and Welpton, 2009). The BSD is an annual snapshot of the Inter-Departmental Business Register (IDBR), a comprehensive list of UK enterprises maintained by the ONS and used by the government for statistical purposes. Employment information on the IDBR is updated periodically from administrative sources, such as HMRC PAYE and Value Added Tax (VAT) records, as well as ONS surveys, such as the annual Business Register and Employment Survey. The BSD snapshot is extracted in March of each year. The BSD offers the most comprehensive, research-ready dataset on the population of employee jobs, providing reliable unit-level information about the characteristics of the employers providing those jobs.

We use the number of employees recorded for enterprise j in year t in the BSD (N_{jt}) to compute the expected probability that enterprise j will appear in the issued sample for ASHE in year t. Since the issued sample for ASHE comprises a 1% simple random sample of employee jobs, this probability $\hat{y}_{jt}$ can be expressed as follows (Upward, 2007)[1]:

$$\hat{y}_{jt} = 1 - 0.99^{N_{jt}} \qquad \text{Eq. 1}$$

[1] $\hat{y}_{jt}$ will be under-estimated in cases where an employee holds more than one job with a given enterprise. However, the degree of any under-estimation is likely to be very small, since we find that such instances account for less than 1% of all jobs in ASHE in a typical year.

We then use the unique ONS enterprise reference number (ENTREF), which is present on both BSD and ASHE, to identify which of the enterprises in each year of the BSD can be linked to one or more job records in the corresponding year of ASHE.[2] This allows us to create a binary variable y_{jt}, coded as 1 if enterprise j in year t of the BSD actually appears in the achieved sample for ASHE in year t, and 0 otherwise. The rate at which enterprises respond to ASHE, relative to expectations, can therefore be expressed as:

$$r_{jt} = \left(\frac{y}{\hat{y}}\right)_{jt} \qquad \text{Eq. 2}$$

We use OLS to regress this value on a vector $\boldsymbol{x}_{jt}$ of enterprise demographic characteristics (categorical indicators of enterprise size, workplace region, legal status, industry and enterprise age) and year fixed effects, γ_t.

$$r_{jt} = \alpha + \boldsymbol{x}'_{jt}\boldsymbol{\beta} + \gamma_t + \varepsilon_{jt} \qquad \text{Eq. 3}$$

In this firm-level regression, if the probability that a firm responds to ASHE is uncorrelated with its demographic characteristics (after accounting for its probability of selection), then we would expect that $\boldsymbol{\beta} = \boldsymbol{0}$.

As discussed later in Section 4 (Results), we reject this null hypothesis and so construct an adjustment to the ASHE weights that takes account of residual employer-related response biases within each year. We construct the weighting adjustment by applying the standard ASHE weights and then undertaking a raking procedure (Deville et al., 1993) to compute a year-specific adjustment factor (xs_adj_{it}) for each employee i in year t of the ASHE. This adjustment brings the weighted ASHE sample of employees closer into line with the BSD profile of employees in respect of a set of specified employer characteristics that are suggested by the analysis of Equation 3. Raking is a common method of generating weights, involving an iterative process of adjustments to obtain weights that align with marginal population totals across a number of different variables. We use the `-svycal, rake-` command within Stata, with control totals obtained from the BSD, following the approach set out by Valliant and Dever (2018: 59). This generates a new set of weights:

$$wxs_{it} = weight_{it} \times xs\_adj_{it} \qquad \text{Eq. 4}$$

where $weight_{it}$ is either the standard ASHE cross-sectional weight (CALWGHT) or the ASHE low pay weight (LPCALWGHT) and wxs_{it} is the adjusted version of that weight (named CSWEIGHT and LPCSWEIGHT respectively). Each new weight is then scaled so that $\sum_i wxs_{it} = \sum_i weight_{it}$ within each year.[3]

Finally, we use these new weights to generate descriptive statistics on the distribution of gross hourly earnings, comparing our new estimates with those obtained using the official ONS weighting scheme. We then provide two policy-relevant examples of the importance of our

[2] We find that a small proportion of the ENTREFs in ASHE (less than 2% in any one year) are not present in the BSD. We judge that the discrepancy probably arises from differences in the timing of ASHE and the BSD annual snapshot. These cases are dropped from our analysis.

[3] In view of the concerns about the quality of the LFS, one might consider applying the raking procedure directly to the ASHE design weights (DWEIGHT), thereby dispensing with the LFS adjustment. However, one then has no means of calibrating the profile of ASHE by occupation, gender or age, since these employee characteristics are not measured in the BSD. Consequently, this approach generates an ASHE profile that departs in notable ways on employee characteristics from the profiles generated under CALWGHT or CSWEIGHT (Table S3 in the Supplementary Appendix).

adjustments by re-estimating the incidence of low pay and the scale of the gender pay gap under our new weighitng scheme.

4 Results

4.1 Representativeness of annual ASHE cross-sections

Table 1 presents the results of estimating Eq. 3 in Section 3.1 using the BSD. Although the amount of variance explained in this illustrative analysis is low, some patterns of systematic non-response are apparent in the table, particularly in respect of employer size and legal status. Relative to expectations, the response rate among the largest firms is around 50 percentage points higher than among the smallest firms. And relative to expectations, it is around 15 percentage points higher among organisations in central government than among private limited companies.

[TABLE 1 HERE]

Table 2 compares the profile of employment in ASHE with that observed in the population of enterprises as captured by the BSD. Estimates are presented for 2023, as a representative example of the situation throughout the whole series. In contrast to the statistics in Table 1, this analysis utilises the existing ASHE weights. This is important as they may reduce any employer-related response biases in the unweighted sample through any correlations between employer characteristics and the occupation, gender, age, and regional distribution of employment. Column 1 shows the percentages of employee jobs in each category as observed in the BSD; column 2 shows the equivalent percentages in ASHE after applying the standard ASHE weights (CALWGHT).[4] Column 3 shows the differences between the ASHE weighted estimates and the BSD estimates, which can be interpreted as a measure of the bias in the ASHE estimates generated under the standard weighting scheme, assuming that the estimates generated from the BSD are unbiased. The figures in Column 3 indicate that employee jobs in larger firms are over-represented in the ASHE sample, even after applying the standard ASHE weights. Jobs in private sector enterprises (especially companies) and younger enterprises are under-represented, while those in public sector enterprises and older enterprises are over-represented.

[TABLE 2 HERE]

The rows of Column 4 in Table 2 show the squared value of the bias for each category of each variable, which is then averaged over all categories of the variable at the base of each set of values. This averaged value allows us to estimate the extent of bias across variables with different numbers of categories; larger values indicate that the profile of ASHE is more biased across that particular variable. The largest values are seen in respect of workforce size, legal status and firm age.[5] While there are some differences between the BSD and ASHE in terms of the distribution of jobs across region and industry (jobs in Education and Health and Social Work are over-represented in ASHE compared to the profile observed in the BSD, for example), the magnitude of the bias is relatively small across these dimensions.[6]

As described in Section 3.1, we use a raking procedure to adjust the standard ASHE weights, aiming to account for residual employer-related response biases. It is usually helpful to be

[4] Strictly speaking, values for the BSD show the percentage of employees rather than employee jobs, but as noted earlier, the percentage of employees holding multiple jobs with the same employer is very small.
[5] In interpreting these findings, one must bear in mind that the values in Table 2 focus on one variable at a time. Firm age is correlated with size and legal status, and hence the particularly pronounced differences by age seen in Table 2 are likely to be partly accounted for by size and legal status in the regression analysis presented in Table 1.
[6] There are, nevertheless, some concerns about measurement errors in ASHE's workplace location data (see Whittard et al., 2023).

parsimonious in the variables used for this adjustment. While a greater number of variables (and categories) will reduce any sample bias, one faces the risk of small cells. Utilising a greater number of categories (strata) may also increase the variance of the weighting values, thereby reducing the precision of estimates generated using the new weighting scheme – also known as the 'design effect' (Kish, 1965, 1992).

Informed by the analysis presented in Table 1 and Table 2, we focus on three variables for our adjustment: employer size (1-4 employees; 5-49 employees; 50-4,999 employees; 5,000 or more employees); legal status (private sector; public sector/non-profit); and firm age (less than 10 years old; 10-29 years; 30 or more years). We also trim the weights following the approach recommended by Valliant and Dever (2018: 157), capping the maximum and minimum weight values at points equal to the median value of the weight plus or minus three times the value of the interquartile range.

Table 3 compares the performance of this new adjusted ASHE weight (CSWEIGHT) with that of the original weight (CALWGHT). The adjusted weight delivers a reduction in bias compared to the original ASHE weighting on all dimensions considered (an average bias of 1.0 across all employer characteristics, compared with 9.0 for CALWGHT). The Kish (1965, 1992) design effect, which measures the impact of the weighting scheme on the variance of estimated generated from the sample, shows that the adjusted weighting scheme incurs only a small loss of precision (a design effect of 1.29, compared with 1.15 for CALWGHT).[7]

[TABLE 3 HERE]

To check that our adjusted weight does not inadvertently distort the weighted profile of ASHE based on employee characteristics (especially those targeted under the original weighting scheme), Supplementary Appendix Table S2 compares the ASHE sample profiles under the original and our adjusted weights for key employee characteristics. We see that the weighted profiles by employee characteristics are very similar under both weighting schemes.

Table 4 illustrates the impact of the weighting adjustment on descriptive statistics for nominal gross hourly earnings for selected years of ASHE (2004, 2013 and 2023, representing the beginning, middle and end of the current data series). The main effect of the weighting adjustment is a leftward shift in the hourly earnings distribution each year. Mean gross hourly earnings are lower by 17 pence per hour in 2004 (z=3.87; $p<0.001$), 38 pence per hour in 2013 (z=4.48; $p<0.001$), and 44 pence per hour in 2023 (z=6.23; $p<0.001$).[8] The impact is greater in the upper half of the distribution, and so the interquartile range (IQR) becomes narrower in absolute terms each year under CSWEIGHT, with the effect of the weighting adjustment being greater in later years. Wage inequality – measured by the IQR – is therefore lower under the new weighting scheme, although the impact on the p90: p10 ratio – another measure of inequality – is minimal. The final columns of Table 4 show the growth in nominal wages across the whole period (2004 to 2023). Growth in mean nominal wages is slightly lower under the new weights but this difference

[7] Around 4 per cent of weighting values were trimmed when generating CSWEIGHT. The untrimmed weights delivered a lower degree of bias (an average across all employer characteristics of 0.6, with the benefits seen particularly in terms of enterprise size (0.4) and enterprise age (1.2)), but this came at the expense of a higher design effect (1.5).

[8] z-scores are computed using a difference-of-means test for two estimates, accounting for the design-adjusted sampling error of each estimate.

is very small (70% versus 69%). The leftward adjustment to the distribution is thus proportionately similar across years.[9]

[TABLE 4 HERE]

The leftward shift in the wage distribution is primarily caused by the increased representation of jobs in smaller and younger firms under the revised weighting scheme, since these pay less, on average, than larger and older firms. The increased representation of private sector firms also contributes, but to a lesser extent. A shift-share analysis of the change in the estimated mean wage due to the revised weighting can suggest the importance of each employer characteristic in isolation. For instance, only adjusting the profile of jobs by enterprise size, fixing the mean wages within the size categories at their levels before applying the revised weights, would reduce the overall estimated mean wage in 2023 by 31 pence (from £19.50 to £19.19), against a total adjustment from the revised weights of 44 pence. Similarly, adjusting only the enterprise age profile of the sample according to the revised weights would reduce the overall mean wage in 2023 by 23 pence, adjusting only the private-public profile reduces it by 7 pence.[10]

5 Implications for estimates of the incidence of low pay

We use the new weights to assess the implications of sample bias in ASHE for estimates of low pay. Improving estimates of the incidence of low pay was a key motivation for the introduction of ASHE, and it remains the primary data source used by the Low Pay Commission, ONS, and others to estimate patterns of pay and the impact of minimum wage legislation (Low Pay Commission, 2022, ONS, 2023b, Cominetti et al., 2022).

First, we investigate the coverage rate of the minimum wage under the alternative weighting schemes. This coverage rate is a key benchmark as it indicates the share of all jobs that are paid at the wage floor. It indicates the reach of the minimum wage across the labour market and how this changes as the minimum wage is uprated.

Second, we investigate the bite of the minimum wage, which is the value of the minimum as a percentage of median wages, also knows as the Kaitz Index (Kaitz, 1970). Since the introduction of ASHE in 2004, the UK Government has progressively increased the bite of the National Minimum Wage (NMW). In 2016, the Government also introduced a higher National Living Wage (NLW) for employees aged 25 and over, with eligibility subsequently extended to employees aged 23-24 in 2021 and those aged 21-22 in 2024. When the NLW was introduced, the Government set a target for it to reach 60 percent of median hourly earnings by October 2020, and subsequently revised the target to two-thirds of median hourly earnings by October 2024 (DBEIS, 2020). We assess progress towards this target.

In Section 4, we found that our adjusted cross-sectional weights shifted the cross-sectional wage distribution to the left. As a result, we would expect the coverage rate and bite of the minimum wage to be higher under our adjusted weight compared to the standard ASHE weights.

[9] To test the sensitivity of our results to the use of a more detailed set of raking targets, we constructed an alternative weight using the nine categories of enterprise size, eight categories of legal status and six categories of enterprise age presented in Table 2. This approach delivered a small additional reduction in bias (an average over all employer characteristics of 0.9 rather than 1.0 as reported in Table 3) with no penalty in terms of an increased design effect due to the trimming of extreme weights. However, the impact on the wage distribution was minimal when compared with the more parsimonious approach (for instance, the estimated mean wage for 2023 remained at £19.06 under this alternative scheme) and so we retain the more parsimonious weighting scheme for the remainder of the paper.

[10] The shift-share analysis is presented in Table S3 in the Supplementary Appendix.

5.1 The coverage rate and bite of the National Minimum Wage and National Living Wage

Using the cross-sectional low-pay weights supplied by ONS (LPCALWGHT), the share of jobs paid at or below the NMW or NLW rose from 2.7% in 2004 to 4.3% in 2015 (Figure 2).[11] The introduction of the NLW in 2016 pushed the incidence further upwards, reaching 6.7%. The coverage rate declined slowly through to 2019, after which it fell more rapidly, reaching 4.9% in 2023. Applying our adjusted cross-sectional weights (LPCSWEIGHT) raises the estimated share of jobs paid at or below the NMW or NLW by approximately one fifth in each year. This raises the share to 8.1% in 2016 and 6.3% in 2023. The differences between the estimates from the two sets of weights are statistically significant from zero in every year of the series.

[FIGURE 2 HERE]

Regarding the bite of the NMW and NLW, the original ASHE cross-sectional weights estimate it to have risen from 51.4% in 2004 to 64.1% in 2023 (Figure 3). According to these estimates, the UK Government appeared to meet its target of a 60% bite in 2020. However, when the median wage is instead estimated using our adjusted cross-section weights, the bite is found to have risen from 52.7% in 2004 to 66.7% in 2023. The revised estimates show that the 60% target was reached in 2018 – two years earlier than previously thought – and the target of two-thirds median wages was met by 2023 – one year ahead of schedule.

[FIGURE 3 HERE]

6 Implications for estimates of the gender pay gap

In our second policy-relevant example, we use our improved weighting scheme to generate new estimates of the gender pay gap (GPG). This potentially has implications beyond national statistics, since the ASHE has been used recently in a swathe of new research that examines the role of firms and employers in gender wage gaps (e.g., Jewell et al., 2020; Jones & Kaya, 2023; Phan et al., 2023; Pham et al., 2024). ASHE has also been used to evaluate the impacts since 2018 of the UK Government's gender pay gap transparency law (Blundell et al., 2024).

The ONS headline measure of the GPG compares the gross hourly earnings (excl. overtime; henceforth referred to as the 'wage') of men and women observed working full-time in the ASHE dataset, applying the original ONS-produced cross-section weights (CALWGHT) to generate the statistics. The ONS computes the headline GPG by subtracting the median wage of women from the median wage of men and then dividing this difference by the median wage of men. Thus, the GPG gives the difference between average male and female hourly pay as a proportion of average male pay. As above for low pay, we only compare estimates using the different weighting schemes for Great Britain, so cannot compare directly to the UK-level estimates that are published annually by ONS.

Our alternative weights for ASHE suggest that the ONS has been consistently under-estimating the UK gender pay gap over the past 20 years, by a small but noteworthy margin of around one

[11] We follow the Low Pay Commission (LPC) in using a measure of gross hourly earnings that includes basic pay, bonus or incentive pay, and pay received for other reasons, but excludes overtime and shift premium pay. We also follow the LPC in measuring the employee's wage against the NMW or NLW rate that applied in April of the relevant year, even if (from 2016 onwards) the NLW may have been uprated part-way through the pay period reported in ASHE. And we follow the LPC in measuring the share of jobs paid at or below the NMW and NLW plus 5 pence. Our estimates for 2020 and 2021 make no attempt to adjust the wages of employees on furlough, however, and so are lower than those reported by the LPC who make adjustments which raise the pay of such employees to arrive at an estimate of what the employee would be paid if they were not furloughed.

percentage point (see Figure 4 and Table 5). In general, applying the original ONS weights to the unweighted data raises the median wage of men in ASHE more so than it does for women, leading to a larger gender pay gap than the raw averages in the unweighted data would imply. However, our alternative weights then lead to a downwards revision in median wages for both men and women, but more so for women. This is because the original ONS weighting scheme tends to under-represent smaller private sector firms, especially for women, and over-represents larger and public sector employers, where the level of pay is generally higher and the differences between men and women within jobs are generally smaller. Looking beyond the mean and median, the sampling bias that is corrected by our alternative weights has, in recent years, larger effects on estimates of relative gaps between the 25th and 75th percentiles of the male and female wage distributions, compared with the 10th and 90th percentiles (see Figure 5).

[TABLE 5 HERE]

[FIGURE 4 HERE]

[FIGURE 5 HERE]

7 Summary and conclusions

The Annual Survey of Hours and Earnings (ASHE) provides many of the UK's official earnings statistics, based on an issued sample comprising 1% of all employee jobs. Its predecessor, the NES, was criticised for lacking survey weights to address known response biases. ASHE improved upon this by introducing cross-sectional weights that adjust the profile of the achieved annual samples to ensure they are representative of the population of employee jobs in terms of gender, age, occupation, and region (see Bird, 2004; Pont, 2007).

Our analysis reveals that jobs in smaller, younger, and private-sector enterprises remain under-represented in the annual achieved samples, even after applying those weights. We develop an alternative set of cross-section weights aimed at reducing these biases. Further, we demonstrate the relevance of our findings for our understanding of the labour market through two policy-relevant examples. First, we re-estimate the incidence of low pay, showing that the incidence of minimum wage employment is under-estimated under the original cross-sectional weighting scheme. We further show that the bite of the minimum wage is higher than previously estimated, such that the Government's targets for this measure have been reached more quickly than previously thought. Second, we re-estimate the size of the gender pay gap, showing that the gap has been consistently under-estimated over the past 20 years, by around one percentage point.

Our findings contribute to the literature on rates and patterns of minimum wage employment and gender pay inequality in the UK (e.g. Dickens et al., 2015; Aitken et al., 2019; Low Pay Commission, 2022; Jewell et al., 2020; Jones and Kaya, 2023). They also contribute to the broader literature on the nature, detection and removal of non-response biases in business surveys (e.g. Willimack et al., 2002; Willimack and Snijkers, 2013), by showing the value of investigating non-response patterns from multiple perspectives and using multiple auxiliary sources for calibration.

We focus exclusively on descriptive statistics, as most of the official statistics produced from ASHE fall into this category. Such statistics are a frequent point of reference for bodies such as the Low Pay Commission, the Office for Equality and Opportunity and the Office for the Pay Review Bodies, and so we judge that our findings have a great deal of relevance for their work.

Nevertheless, policy makers and the broader research community also make extensive use of ASHE to estimate average partial effects such as the relationship between gender and earnings for employees of similar age and education, or to identify causal effects such as the impact of an increase in the NLW on job retention. Further research would be needed to understand the impact of ASHE weighting adjustments on such analyses. As in the case of descriptive statistics, the impact will vary from case to case though, in general terms, the bias in estimated parameters due to unrepresentative samples is often less severe in the multivariate case, since the structure of the sample may be partialled out to some extent through the use of design-related characteristics as control variables.[12] Analysts wishing to test the sensitivity of specific estimates may access the code needed to derive the adjusted weights via the website of the Wage and Employment Dynamics Project and also within the ONS Secure Research Service (see Data Availability Statement).

Based on our findings, we conclude that new official review of the methodology for ASHE is merited. An earlier review of the methodology for its predecessor, the New Earnings Survey, led to significant improvements (see Bird, 2004; Pont, 2007). A new review would be able to address the issues we have identified by re-examining the approach to sampling, response-chasing and weighting in ASHE. Our findings suggest that the weighting methodology should be extended to include employer characteristics. The review should also consider extending the research-ready dataset to include fieldwork outcomes for all members of the issued sample, so that analysts can better understand patterns of non-response. There is also potential to improve the coherence of data sources by making greater use in sampling and weighting of data from HMRC's PAYE Real Time Information system (see Forth et al., 2025). Any such review of ASHE would provide an opportunity to enhance the quality of data used for earnings analysis and research in the UK. This could have particularly significant implications for government policy in range of areas, including minimum wages, the earnings gaps experienced by employees with protected characteristics, and the pay of public sector workers..

[12] This is sometimes referred to as the 'model-based' approach to estimation, contrasting with the 'design-based' approach of applying weights (see Lohr, 1999). See Solon et al. (2015) for further discussion of the use of weights when moving beyond descriptive statistics.

References

Aitken A, Dolton P and Riley R (2019) "The impact of the introduction of the National Living Wage on employment, hours and wages", NIESR Discussion Paper No. 501, London: National Institute of Economic and Social Research.

Bird D (2004) "Methodology for the 2004 Annual Survey of Hours and Earnings", Technical Report, *Labour Market Trends*, Office for National Statistics, pp. 457-464.

Blundell, J., Duchini E, Simion S and Turrell A (2025) "Pay Transparency and Gender Equality", *American Economic Journal: Economic Policy*, 17, 2: 418-45.

Cominetti N, Costa R, Datta N and Odamtten F (2022) *Low Pay Britain 2022*, London: Resolution Foundation.

DBEIS (2020) *National Living Wage and National Minimum Wage: Low Pay Commission Remit 2020*, London: Department for Business, Energy and Industrial Strategy.

Deville J, Särndal C and Sautory O (1993) "Generalized raking procedures in survey sampling", *Journal of the American Statistical Association,* 88: 1013–1020.

Dickens R, Riley R and Wilkinson D (2015) "A re-examination of the impact of the UK National Minimum Wage on employment", *Economica*, 82*,* 328: 841-864.

Evans P and Welpton R (2009) "Business structure database: the Inter-Departmental Business Register (IDBR) for research", *Economic and Labour Market Review*, 3, 6: 71–5.

Forth, J., Singleton, C., Bryson, A., Phan, V., Ritchie, F., Stokes, L. and Whittard, D. (2025a) "The impact of a rising wage floor on labour mobility across firms", *British Journal of Industrial Relations*, in print.

Forth, J., Bryson, A. and Palmou, C. (2025b) "A roadmap for developing a new linked employer-employee infrastructure for the UK", Economic Statistics Centre of Excellence (ESCoE) Technical Report TR-28.

Giles, C (2023) "The UK's 'official' labour data is becoming a nonsense", *Financial Times*, 25th October.

Giupponi G, Joyce R, Lindner A, Waters, T, Wernham, T and Xu, X (2024). "The employment and distributional impacts of nationwide minimum wage changes", *Journal of Labor Economics*, 42, S1: S293-S333.

Jewell S L, Razzu G and Singleton C (2020) "Who works for whom and the UK gender pay gap", *British Journal of Industrial Relations*, 58, 1: 50-81.

Jones M and Kaya E (2023) "The UK gender pay gap: Does firm size matter?" *Economica*, 90, 359: 937-952.

Kaitz, H (1970) "Experience of the past: The National Minimum, Youth Unemployment and Minimum Wages", *US Bureau of Labor Statistics Bulletin*, 1657, 30-54.

Kish L (1965) *Survey Sampling*, New York: John Wiley.

Kish L (1992) "Weighting for unequal P_i", *Journal of Official Statistics*, 8, 2: 183.

Lohr, S. (1999) *Sampling: Design and Analysis*, Pacific Grove, CA: Duxbury Press.

Low Pay Commission (2022) *The National Living Wage Review (2015-2020)*, London: Low Pay Commission.

NISRA (2024) *Quarterly Employment Survey Historical Tables, March 2024: Table 5.11 – Employee Jobs by Broad Industry Sector, December 1978 to March 2024*, Belfast: Northern Ireland Statistics and Research Agency.

Office for Statistics Regulation (2023) *Review of Experimental Labour Market Statistics produced by the Office for National Statistics (ONS)*, London: Office for Statistics Regulation.

Office for Statistics Regulation (2024) *Review of Reintroduced LFS-derived Labour Market Statistics*, London: Office for Statistics Regulation.

ONS (2002) *Review of Distribution of Earnings Statistics*, National Statistics Quality Review Series Report No. 14, London: Office for National Statistics.

ONS (2015) *Business Surveys*, https://www.ons.gov.uk/aboutus/transparencyandgovernance/freedomofinformationfoi/businesssurveys (last accessed 18th July 2025).

ONS (2018) *Annual Survey of Hours and Earnings, Low pay and Annual Survey of Hours and Earnings pension results QMI*, London: Office for National Statistics.

ONS (2020) *Coronavirus and its Impact on the Labour Force Survey*, London: Office for National Statistics.

ONS (2021) *Labour Force Survey Weighting Methodology*, London: Office for National Statistics.

ONS (2023a) *Employee earnings in the UK: 2023*, London: Office for National Statistics.

ONS (2023b) *Low and High Pay in the UK: 2023*, London: Office for National Statistics.

ONS (2024a) *Annual Survey of Hours and Earnings - UK, dataset*, Office for National Statistics Metadata Catalogue, DOI: 10.57906/x25d-4j96

ONS (2024b) *Gender pay gap in the UK: 2024,* London: Office for National Statistics.

ONS (2024c) *Business Structure Database - UK*, Office for National Statistics Metadata Catalogue, DOI: 10.57906/7kh0-0910.

ONS (2024d) *JOBS01: Workforce Jobs Summary*, *June 2024,* London: Office for National Statistics.

ONS (2024e) *Labour Force Survey User Guide - Volume 1: Background and Methodology*, London: Office for National Statistics.

ONS (2024f) *Impact of Reweighting on Labour Force Survey Key Indicators: December 2024*, London: Office for National Statistics.

ONS (2025) *Labour Force Survey Quality Update: May 2025*, London: Office for National Statistics.

Pal K K, Piaget K and Zahidi S (2024) Global Gender Gap Report 2024. World Economic Forum.

Pham T, Schaefer D and Singleton C (2024) Unequal hiring wages and their impact on the gender pay gap. IZA Discussion Papers 17285, Institute of Labor Economics (IZA).

Phan V, Singleton C, Bryson A, Forth J, Ritchie F, Stokes L and Whittard D (2023) Accounting for firms in gender-ethnicity wage gaps throughout the earnings distribution. Economics Discussion Papers em-dp2023-16, Department of Economics, University of Reading.

Pont M (2007) "Coverage and non-response errors in the UK New Earnings Survey", *Journal of the Royal Statistical Society: Series A*, 170, 3: 713-733.

Solon G, Haider S and Wooldridge J (2015) "What are we weighting for?", *Journal of Human Resources*, 50, 2: 301-316.

Stokes L, Bryson A, Forth J and Weale M (2017) "Who fared better? The fortunes of performance pay and fixed pay workers through recession", *British Journal of Industrial Relations*, 55, 4: 778-801.

Upward, R. (2007) "Linking the ASHE and the BSD", University of Nottingham, mimeo.

Valliant R and Dever J (2018) *Survey Weights: A Step-by-Step Guide to Calculation* (Vol. 1). College Station, TX: Stata Press.

Whittard, D., Ritchie, F., Phan, V., Bryson, A., Forth, J., Stokes, L. and Singleton, C., (2023) "The perils of pre-filling: lessons from the UK's Annual Survey of Hours and Earning microdata" *Statistical Journal of the IAOS*, *39*, 3: 661-677.

Willimack D, Nichols K and Sudman S (2002) "Understanding unit and item nonresponse in business surveys", in R Groves, D Dillman, J Eltinge and R Little (eds) *Survey Nonresponse*, New York: Wiley.

Willimack D and Snijkers G (2013) "The business context and its implications for the survey response process", in Snijkers G, Haraldsen G, Jones J and Willimack D (eds) (2013) *Designing and Conducting Business Surveys*, New York: Wiley.

Table 1: Probability that a BSD employer appears in ASHE relative to expectations, OLS regression results

	Coefficient		Std. error
Enterprise size (ref. 1-4 employees)			
5-9	0.153	***	0.002
10-49	0.259	***	0.002
50-249	0.326	***	0.003
250-499	0.416	***	0.004
500-999	0.475	***	0.005
1000-2499	0.515	***	0.006
2500-4999	0.529	***	0.008
5000 plus	0.519	***	0.011
Workplace region (ref. London)			
North East	0.030	***	0.006
North West	-0.009	***	0.003
Yorkshire & Humberside	0.015	***	0.004
East Midlands	0.014	***	0.004
West Midlands	-0.001		0.003
South West	0.011	***	0.003
East	-0.038	***	0.002
South East	0.023	***	0.003
Wales	0.009	*	0.005
Scotland	0.009	**	0.004
Legal status (ref. Private company)			
Sole proprietor	0.065	***	0.003
Partnership	0.071	***	0.005
Public corporation	0.084	**	0.039
Central government body	0.191	***	0.016
Local authority	0.248	***	0.028
Non-profit making body	0.124	***	0.007
Industry (SIC07) (ref. Manufacturing)			
Agriculture, forestry, fishing	-0.069	***	0.008
Mining and quarrying	-0.075	***	0.020
Electricity, gas	-0.111	***	0.008
Water supply	-0.018	*	0.010
Construction	-0.031	***	0.004
Wholesale, retail	-0.009	**	0.004
Transport and storage	-0.039	***	0.004
Accommodation & food	-0.055	***	0.004
Information & communication	-0.038	***	0.004
Financial and insurance	0.026	***	0.007
Real estate	-0.053	***	0.005
Professional, scientific, technical	-0.025	***	0.004

Admin and support	-0.030	***	0.004
Public administration	0.190	***	0.046
Education	0.046	***	0.006
Health and social work	0.069	***	0.006
Art, entertainment, recreation	-0.027	***	0.006
Other service activities	0.060	***	0.006
Enterprise age (ref. 1 year or less)			
2-4 years	-0.010	***	0.002
5-9 years	-0.006	***	0.002
10-19 years	0.042	***	0.002
20-29 years	0.091	***	0.003
30+ years	0.116	***	0.004
Year (ref. 2004)			
2005	0.000		0.002
2006	-0.033	***	0.002
2007	-0.077	***	0.003
2008	-0.091	***	0.003
2009	-0.081	***	0.003
2010	-0.080	***	0.003
2011	0.004		0.003
2012	-0.031	***	0.003
2013	-0.037	***	0.003
2014	-0.027	***	0.003
2015	-0.011	***	0.003
2016	-0.022	***	0.003
2017	-0.032	***	0.003
2018	-0.011	***	0.003
2019	-0.036	***	0.003
2020	-0.055	***	0.003
2021	-0.060	***	0.003
2022	-0.067	***	0.003
2023	-0.078	***	0.003
Constant	0.177	***	0.005
N observations(enterprises)	*54,292,380*		
R-squared	*0.002*		

Base: all enterprises with at least 1 employee, located in England, Scotland or Wales.
Notes: author calculations using BSD dataset. See Equation (3.) ***, **, * indicate significant differences from zero, two-sided tests, at the 1%, 5% and 10% levels, respectively. Robust standard errors.

Table 2: Employer characteristics associated with response to ASHE, 2023

	BSD (employee-weighted), per cent	ASHE weighted (CALWGHT), per cent	Bias in ASHE (2)- (1)	Bias squared
	(1)	(2)	(3)	(4)
Enterprise size (number of employees):				
1-4	11.8	4.2	-7.7	58.9
5-9	6.2	4.2	-2.0	4.1
10-49	15.5	14.6	-0.9	0.8
50-249	14.1	15.1	1.1	1.1
250-499	5.8	6.5	0.7	0.5
500-999	6.0	7.3	1.3	1.6
1000-2499	7.2	8.3	1.0	1.1
2500-4999	6.8	8.3	1.4	2.0
5000 plus	26.5	31.6	5.1	25.8
Average bias (mean):				*10.7*
Legal status:				
Company	72.9	64.5	-8.4	69.8
Sole proprietor	2.1	1.2	-1.0	0.9
Partnership	1.9	2.0	0.1	0.0
Public corporation	0.7	0.8	0.1	0.0
Central government body	9.9	14.2	4.2	18.1
Local authority	5.9	8.6	2.7	7.3
Non-profit making body	6.6	8.8	2.2	4.8
Average bias (mean):				*14.4*
Enterprise age:				
1 year or less	2.2	0.7	-1.4	2.1
2-4 years	8.5	3.3	-5.2	27.1
5-9 years	10.7	6.3	-4.4	19.8
10-19 years	14.2	11.7	-2.5	6.5
20-29 years	13.3	14.3	0.9	0.9
30 years plus	51.1	63.8	12.7	161.0
Average bias (mean):				*36.2*
Workplace location:				
North East	3.6	3.6	0.0	0.0
North West	10.1	9.9	-0.2	0.1
Yorkshire	7.7	8.0	0.3	0.1
East Midlands	7.2	7.1	0.0	0.0
West Midlands	8.5	8.2	-0.3	0.1

South West	9.9	10.5	0.6	0.3
East	21.3	18.2	-3.2	10.0
London	13.7	15.0	1.2	1.5
South East	7.5	8.4	0.9	0.8
Wales	3.5	3.7	0.2	0.0
Scotland	7.0	7.5	0.5	0.3
Average bias (mean):				*1.2*
Industry (SIC(2007) Section):				
Agriculture, forestry, and fishing	1.0	0.6	-0.4	0.2
Mining and quarrying	0.2	0.1	-0.1	0.0
Manufacturing	7.5	8.9	1.5	2.2
Electricity, gas, air cond. supply	0.4	0.4	0.0	0.0
Water supply, sewerage, waste	0.6	0.7	0.1	0.0
Construction	5.1	3.7	-1.4	1.9
Wholesale, retail, repair of vehicles	14.6	13.9	-0.7	0.4
Transport and storage	4.6	3.9	-0.6	0.4
Accommodation and food service	8.0	5.4	-2.6	6.7
Information and communication	4.5	4.3	-0.3	0.1
Financial and insurance activities	3.3	3.8	0.5	0.3
Real estate activities	2.0	1.6	-0.5	0.2
Professional, scientific, and technical	8.3	8.0	-0.4	0.1
Admin and support services	9.4	6.3	-3.1	9.4
Public admin and defence	3.7	5.3	1.5	2.3
Education	9.7	13.9	4.2	17.3
Health and social work	12.8	15.5	2.7	7.2
Art, entertainment, and recreation	2.2	1.9	-0.3	0.1
Other service activities	1.9	1.7	-0.2	0.0
Activities of households as employers	0.1	0.1	-0.1	0.0
Average bias (mean):				*2.4*
N observations	*3,127,074*	*148,573*		

Base: all employee jobs.

Note: Figures for BSD are employee-weighted estimates, for those enterprises that are recorded as having at least one employee in the BSD. ASHE estimates exclude those employee jobs for whom it was not possible to match to an enterprise record in BSD (affecting 0.5 per cent of the original ASHE sample in 2023).

Table 3: Bias and design effect, original and adjusted ASHE weights, 2023

	Original ASHE weight (CALWGHT)	Adjusted ASHE weight (CSWEIGHT)
Average bias – enterprise size	10.7	1.5
Average bias – legal status	14.4	0.3
Average bias – enterprise age	36.2	2.5
Average bias – workplace region	1.2	1.0
Average bias – industry sector	2.4	0.6
Average bias across all characteristics	9.0	1.0
Kish design effect	1.15	1.29

Base: all employee jobs.

Notes: author calculations using ASHE dataset, 2023. Number of observations: 148,573.

Table 4: Employee nominal gross hourly earnings (pence per hour) under different weighting schemes, 2004, 2013 and 2023

	2004			2013			2023			Growth (2023/2004)		
	CALWGHT	CSWEIGHT	Diff.	CALWGHT	CSWEIGHT	Diff.	CALWGHT	CSWEIGHT	Diff.	CALWGHT	CSWEIGHT	Diff.
Mean	1145	1128	-17	1461	1423	-38	1950	1906	-44	1.70	1.69	-0.01
p10	505	500	-5	653	645	-8	1069	1051	-18	2.12	2.10	-0.01
p25	623	611	-13	800	774	-26	1196	1171	-25	1.92	1.92	0.00
p50	887	864	-22	1132	1094	-39	1565	1505	-60	1.76	1.74	-0.02
p75	1374	1338	-35	1744	1683	-61	2290	2220	-70	1.67	1.66	-0.01
p90	2018	1993	-25	2530	2481	-49	3235	3189	-46	1.60	1.60	0.00
IQR	750	728	-23	944	909	-36	1094	1050	-44	1.46	1.44	-0.02
p90/p10	3.99	3.99	-0.01	3.88	3.85	-0.03	3.03	3.03	0.01	0.76	0.76	0.00
N. obs.	*147,138*	*147,138*		*168,934*	*168,934*		*138,162*	*138,162*				

Base: all employee jobs paid on adult rates, where no loss of pay due to absence.

Notes: author calculations using ASHE dataset. Diff. = Difference (CSWEIGHT – CALWGHT), pence per hour. IQR = interquartile range (p75-p25). Growth = ($Estimate_{2023}$/$Estimate_{2004}$). The measure of gross hourly earnings includes basic pay, bonus or incentive pay and pay received for other reasons, but excludes overtime and shift premium pay (ASHE variable: HRPAYX)

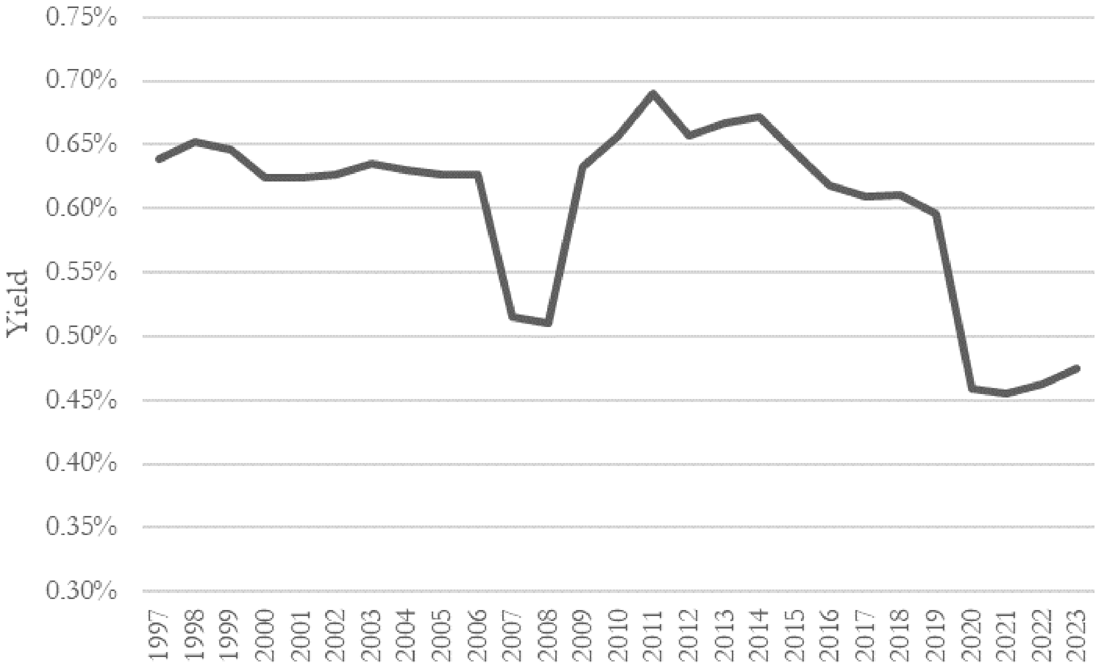
Table 5: Gender Pay Gap in Great Britain for Median and Mean Gross Hourly Earnings (excl. overtime), full-time employees only

Mean gross hourly earnings (pence per hour)							
	Men	Women	GPG	Men	Women	GPG	Difference ("bias") ppts
Year	CALWGHT	CALWGHT	CALWGHT	CSWEIGHT	CSWEIGHT	CSWEIGHT	
2004	1381	1129	-18.3%	1362	1097	-19.5%	-1.2%
2009	1641	1361	-17.1%	1609	1320	-18.0%	-0.9%
2014	1718	1459	-15.1%	1679	1422	-15.3%	-0.3%
2018	1892	1613	-14.8%	1860	1580	-15.0%	-0.3%
2019	1944	1674	-13.9%	1909	1639	-14.1%	-0.2%
2020	2067	1785	-13.6%	2046	1759	-14.0%	-0.4%
2021	2047	1787	-12.7%	2014	1754	-12.9%	-0.2%
2022	2081	1840	-11.6%	2053	1807	-12.0%	-0.4%
2023	2216	1966	-11.3%	2186	1934	-11.5%	-0.2%
Median gross hourly earnings (pence per hour)							
	Men	Women	GPG	Men	Women	GPG	Difference ("bias") ppts
Year	CALWGHT	CALWGHT	CALWGHT	CSWEIGHT	CSWEIGHT	CSWEIGHT	
2004	1100	940	-14.6%	1071	905	-15.5%	-1.0%
2009	1302	1141	-12.3%	1262	1091	-13.6%	-1.3%
2014	1369	1232	-10.0%	1325	1188	-10.3%	-0.3%
2018	1487	1354	-8.9%	1448	1306	-9.8%	-0.9%
2019	1546	1399	-9.5%	1506	1350	-10.4%	-0.9%
2020	1673	1514	-9.5%	1634	1459	-10.7%	-1.2%
2021	1658	1518	-8.4%	1609	1458	-9.4%	-1.0%
2022	1693	1556	-8.1%	1652	1502	-9.1%	-1.0%
2023	1814	1665	-8.2%	1771	1611	-9.0%	-0.8%

Source: ASHE

Notes: All the statistics here do not use observations in ASHE where the employer reports a loss of pay for the employee in the April reference period. They also only use employees on an adult rate of pay. These are the same criteria applied by ONS when constructing national statistics using ASHE.

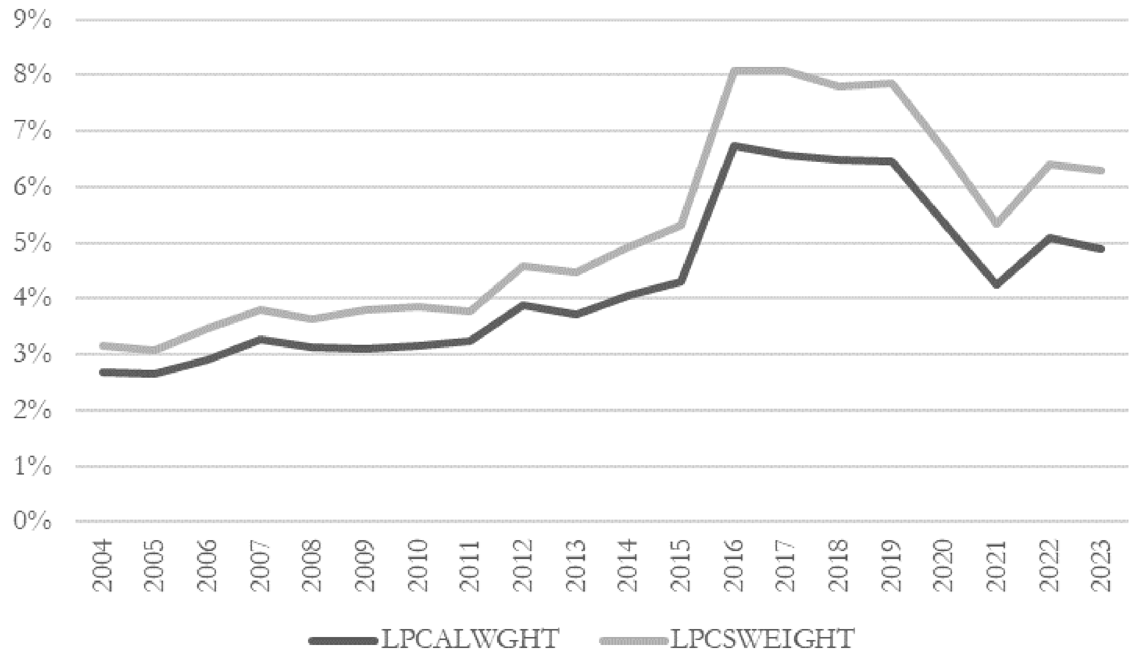
Figure 1: ASHE yield (estimated), by year



Base: all employee jobs in Great Britain.

Notes: Yield estimated by dividing the number of observations in the ASHE research dataset for Great Britain into the estimated total number of employee jobs for Great Britain in the March quarter of each year. Yield is lower in 2007 and 2008 due to cost-saving measures that induced a temporary 20% reduction in the size of the issued sample (ONS, 2007). Underlying values provided in Table S1.

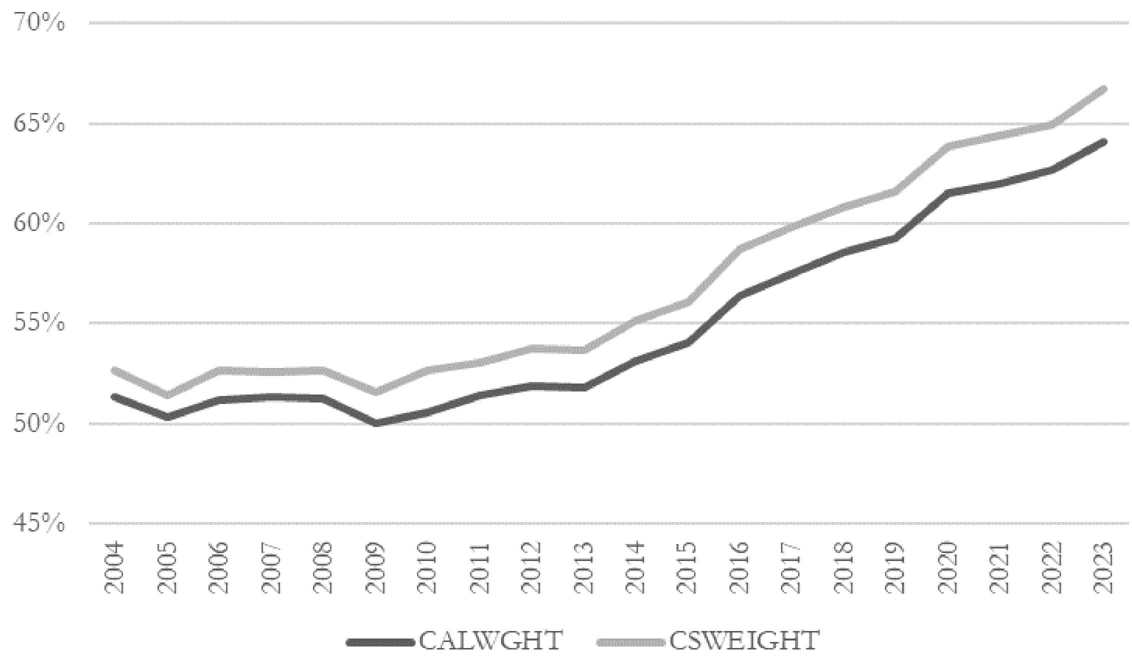
Figure 2: Coverage rate of the adult National Minimum Wage (2004-2015) and National Living Wage (2016-2023) under alternative weighting schemes



Source: ASHE

Notes: All employees aged 25+ (2004, … , 2020) or 23+ (2021, … , 2023), excluding those with loss of pay due to absence, unless due to furlough (2020 and 2021 only). LPCALWGHT is the original ASHE low pay weight. LPCSWEIGHT is our adjusted weight. Differences between the two series are statistically significant from zero at the 0.1% level in each year. See footnote 15 for our approach to measuring coverage.

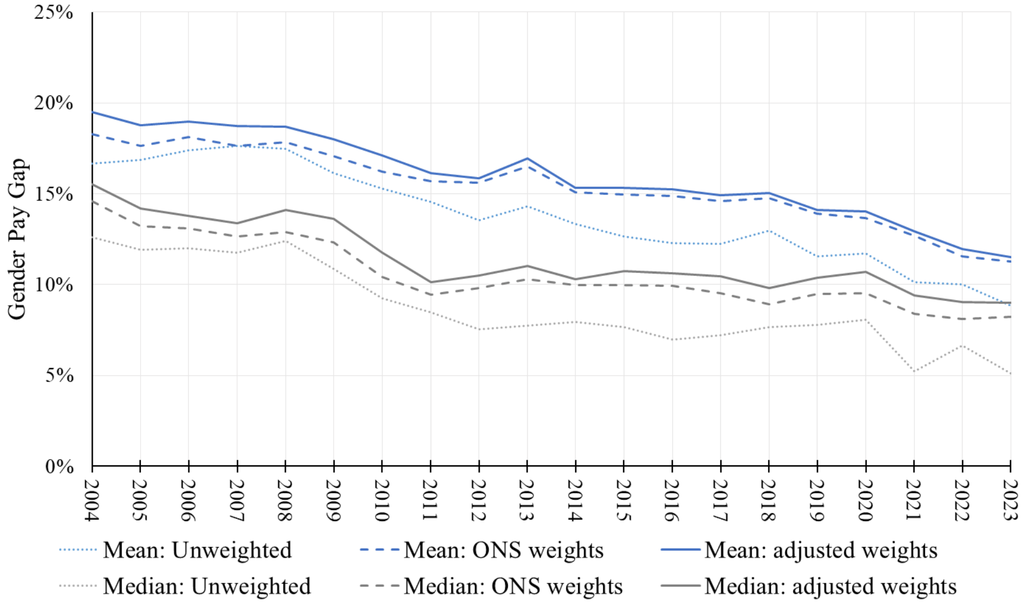
Figure 3: Bite of the adult National Minimum Wage (2004-2015) and National Living Wage (2016-2023) – Kaitz Index – under alternative weighting schemes



Source: ASHE

Notes: Bite estimated as the minimum wage expressed as a percentage of the median wage. Median wage estimated for all employees aged 25+ (2004-2020) or 23+ (2021-2023), excluding those with loss of pay due to absence, unless due to furlough (2020 and 2021 only). CALWGHT: median estimated using ASHE standard weight. CSWEIGHT: median estimated using our adjusted standard weight.

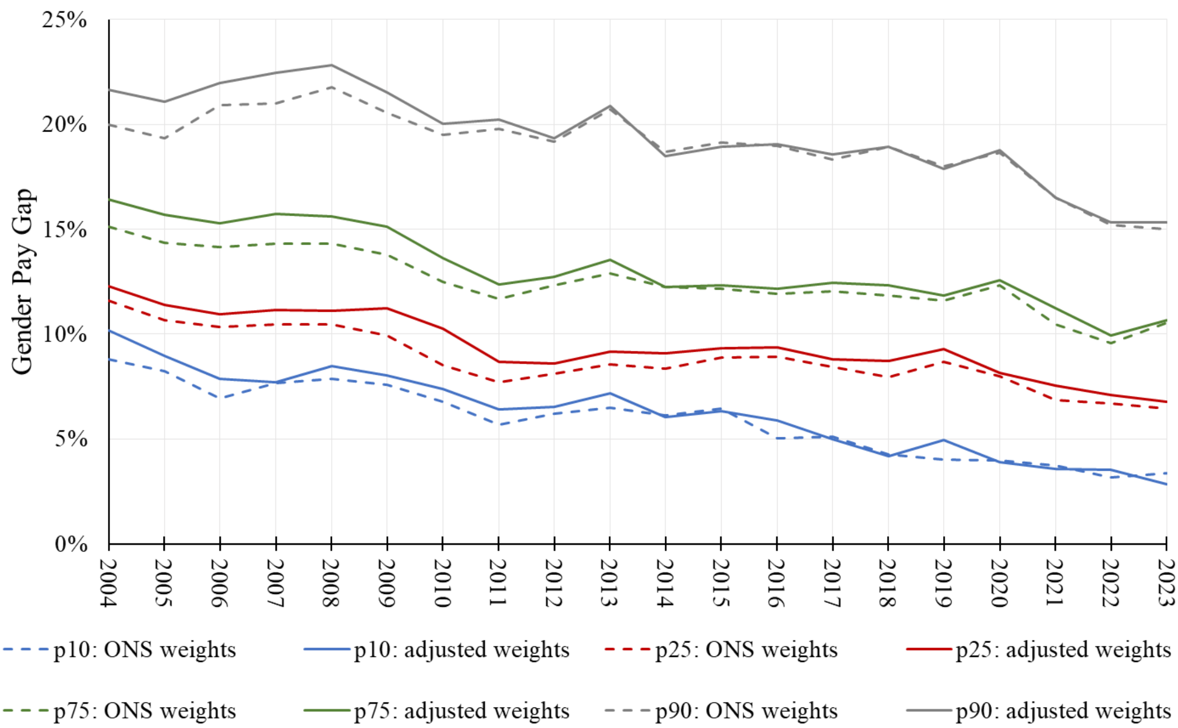
Figure 4: Gender Pay Gap in Great Britain for Median and Mean Gross Hourly Earnings (excl. overtime), full-time employees only



Source: ASHE

Notes: See Table 5.

Figure 5: Gender Pay Gap for Gross Hourly Earnings (excl. overtime), full-time employees only: beyond the median (comparing the 10th, 25th, 75th and 90th percentiles of male and female hourly earners)



Source: ASHE
Notes: See Table 5.

Supplementary Appendix

Table S1: Values used to estimate ASHE yield shown in Figure 1, by year

	Employee jobs: UK ('000s)	Employee jobs: NI ('000s)	Employee jobs: GB ('000s) - estimated	ASHE responses: GB	ASHE yield: GB - estimated
1997	24,697	588	24,109	153,950	0.64%
1998	25,332	609	24,723	161,378	0.65%
1999	25,641	621	25,020	161,750	0.65%
2000	26,097	637	25,460	158,965	0.62%
2001	26,516	649	25,867	161,358	0.62%
2002	26,825	661	26,164	163,821	0.63%
2003	26,866	669	26,197	166,431	0.64%
2004	27,166	681	26,485	166,794	0.63%
2005	27,580	696	26,884	168,343	0.63%
2006	27,831	706	27,125	169,933	0.63%
2007	28,042	715	27,327	140,936	0.52%
2008	28,293	732	27,561	140,703	0.51%
2009	27,899	714	27,185	171,891	0.63%
2010	27,363	708	26,655	175,131	0.66%
2011	27,411	699	26,712	184,501	0.69%
2012	27,703	690	27,013	177,464	0.66%
2013	27,709	695	27,014	180,082	0.67%
2014	28,346	709	27,637	185,762	0.67%
2015	29,198	722	28,476	183,475	0.64%
2016	29,704	730	28,974	179,022	0.62%
2017	30,112	743	29,369	178,943	0.61%
2018	30,252	761	29,491	180,185	0.61%
2019	30,601	774	29,827	177,930	0.60%
2020	30,911	778	30,133	138,385	0.46%
2021	30,394	769	29,625	134,696	0.45%
2022	31,408	796	30,612	141,675	0.46%
2023	32,275	813	31,462	149,372	0.47%

Sources: Employee jobs (UK) are for the March quarter of each year, sourced from ONS (2024d). Employee jobs (NI) for the same quarter are sourced from NISRA (2024). Employee jobs (GB) estimated by subtracting 'employee jobs (NI)' from 'employee jobs (UK)'. Number of ASHE responses (GB) sourced from ONS (2024a).

Table S2: Profile of ASHE sample by employee characteristics, original and alternative ASHE weights, 2023

	ASHE weighted (CALWGHT), per cent	ASHE adjusted weight (CSWEIGHT), per cent	Bias under CALWGHT (2)-(1)	Bias squared
	(1)	(2)	(3)	(4)
Gender:				
Female	49.7	49.1	-0.6	0.4
Male	50.3	50.9	0.6	0.4
Average bias (mean):				*0.4*
Hours:				
Part-time	27.5	29.7	2.2	5.1
Full-time	72.5	70.3	-2.2	5.1
Average bias (mean):				*5.1*
Age group:				
16-19	3.5	4.0	0.5	0.2
20-24	8.0	8.4	0.5	0.2
25-29	11.9	12.0	0.1	0.0
30-34	12.6	12.6	0.0	0.0
35-39	12.0	12.0	-0.1	0.0
40-44	11.2	11.1	-0.1	0.0
45-49	10.6	10.3	-0.2	0.1
50-54	10.8	10.5	-0.3	0.1
55-59	9.7	9.4	-0.3	0.1
60-64	6.6	6.5	-0.1	0.0
65 plus	3.1	3.2	0.1	0.0
Average bias (mean):				*0.1*
Occupation:				
Managers, directors and senior officials	10.4	11.2	0.8	0.6
Professional	28.0	24.8	-3.2	10.4
Associate professional and technical	15.1	14.4	-0.8	0.6
Administrative and secretarial	10.7	11.4	0.7	0.6
Skilled trades	6.3	7.1	0.8	0.6
Caring, leisure and other service	8.2	8.3	0.1	0.0
Sales and customer service	6.3	6.7	0.4	0.2
Process, plant and machine operatives	4.5	4.7	0.2	0.0
Elementary	10.4	11.4	1.0	1.0
Average bias (mean):				*1.6*
N observations	*149,372*	*149,372*		

Base: all employee jobs.

Notes: author calculations using ASHE dataset, 2023.

Table S3: Bias across employer and employee characteristics under various alternative weighting schemes, 2023

	Original ASHE weight (CALWGHT)	Adjusted ASHE weight (CSWEIGHT)	DWEIGHT with BSD calibration only
Average bias – enterprise size	10.7	1.5	2.2
Average bias – legal status	14.4	0.3	0.4
Average bias – enterprise age	36.2	2.5	2.9
Average bias – workplace location	1.2	1.0	2.2
Average bias - industry	2.4	0.6	0.8
Average bias (all employer characteristics)	*9.0*	*1.0*	*1.5*
Average bias – employee gender		0.4	7.5
Average bias – employee hours		5.1	38.3
Average bias – employee age		0.1	0.4
Average bias – employee occupation		1.6	18.1
Average bias (all cemployee haracteristics)		*1.1*	*10.8*

Base: all employee jobs.

Notes: author calculations using ASHE dataset, 2023. Number of observations: 148,573.

Table S4: Shift-share analysis showing the impact of changing the ASHE profile across a single employer characteristic on mean gross hourly earnings, 2023

	Mean wage (pence) (CALWGHT)	Sample share (CALWGHT)	Sample share (CSWEIGHT	Shift in share (3)-(2)	Mean wage (1) * shift in share (4)
	(1)	(2)	(3)	(4)	(5)
Enterprise size (employees):					
1-4	1578	0.04	0.09	0.05	77
5-49	1745	0.19	0.22	0.03	58
50-4,999	2034	0.46	0.42	-0.04	-84
5,000 or more	2006	0.31	0.27	-0.04	-82
		1.00	1.00	0.00	-31
Legal status:					
Private	1926	0.67	0.76	0.08	163
Public/non-profit	2005	0.33	0.24	-0.08	-170
		1.00	1.00	0.00	-7
Enterprise age:					
1 year or less	1765	0.10	0.18	0.08	145
2-4 years	1882	0.26	0.28	0.02	46
5-9 years	2010	0.64	0.53	-0.11	-215
		1.00	1.00	0.00	-23

Base: all employee jobs paid on adult rates, where no loss of pay due to absence.
Notes: author calculations using ASHE dataset, 2023. The measure of gross hourly earnings includes basic pay, bonus or incentive pay and pay received for other reasons but excludes overtime and shift premium pay (ASHE variable: HRPAYX).