

Information paper

Quality and Methodology Information

General details

Title of output: Annual Survey of Hours and Earnings, Low Pay and Annual Survey of Hour and Earnings Pension results
Abbreviated title: ASHE
Designation: National Statistics
Geographic coverage: UK
Date of last SQR or QMI*: April 2011
Contact details: earnings@ons.gov.uk

Executive summary

This report relates to the Annual Survey of Hours and Earnings (ASHE), conducted by ONS. ASHE is the most comprehensive source of earnings information in the United Kingdom. It provides information about the levels, distribution and make-up of earnings and hours paid for employees by gender and full-time/part-time working. Estimates are available for various breakdowns including industries, occupations, geographies and age-groups within the UK. ASHE is used to produce hours and earnings statistics for a range of weekly, annual and hourly measures. ASHE is the official source of estimates for the number of jobs paid below the national minimum wage and is also used to produce estimates of the proportions of jobs within workplace pension categories.

ASHE replaced the New Earnings Survey (NES) as ONS's main source of information on earnings in 2004 and brought improvements to the coverage of employees, imputation for item non-response and the weighting of earnings estimates.

ASHE is based on a 1% sample of employee jobs taken from HM Revenue & Customs (HMRC) Pay As You Earn (PAYE) records. Information on earnings and hours is obtained from employers and treated confidentially. ASHE does not cover the self-employed nor does it cover employees not paid during the reference period. Results are published annually via the ONS website.

Information relating to ASHE outputs, methodology and uses can be found on the [guidance and methodology page on the ONS website](#)¹.

There have been changes to the ASHE methodology over the years, which have resulted in breaks in the series in 2004, 2006 and 2011. These changes are described below in the 'How the output is created' section.

This document contains the following sections:

- Output quality;
- About the output;
- How the output is created;
- Validation and quality assurance;
- Concepts and definitions;
- Other information, relating to quality trade-offs and user needs; and
- Sources for further information or advice.

* Quality and Methodology Information' (QMI) replaced 'Summary Quality Reports' (SQR) from 04/11

Output quality

This document provides a range of information that describes the quality of the output and details any points that should be noted when using the output.

ONS has developed [Guidelines for Measuring Statistical Quality](#)²; these are based upon the five European Statistical System (ESS) quality dimensions. This document addresses these quality dimensions and other important quality characteristics, which are:

- Relevance;
- Timeliness and punctuality;
- Coherence and comparability;
- Accuracy;
- Output quality trade-offs;
- Assessment of user needs and perceptions; and
- Accessibility and clarity.

More information is provided about these quality dimensions in the sections below.

About the output

Relevance

(The degree to which the statistical outputs meet users' needs.)

Through consultation with the ASHE user base, it has been established that ASHE data meet a vast and diverse range of user needs throughout government, professional organisations and the wider public.

Some examples of government uses, for which the department has confirmed that ASHE data are fit for purpose are:

- Labour Market Division (ONS) - statistics used in various analyses of conditions in the labour market, feeding into publications;
- HMRC - various routine uses including investigation of changes to rates of taxation;
- Department for Work and Pensions (DWP) - analysis of pension scheme membership, contributions and persistency of saving for policy development, monitoring and evaluation;
- Department for Business, Innovation and Skills (BIS) & the Low Pay Commission (LPC) - data used in review body remits to analyse pay comparability. Data used for low pay and National Minimum Wage (NMW) briefing, policy analyses and analyses of annual leave, agency workers and pay generally. Data used to analyse the structure of earnings, monitor the impact of NMW and recommend future rates;
- Office of Manpower Economics (OME) - data used in the provision of services to public sector Pay Review Bodies and the Police Negotiating Board, which make recommendations to Government on pay for public sector workers, with an estimated paybill in excess of £80bn (figure correct at July 2011); and
- Department of Health (DH) - Data used to estimate wage differentials at Primary Care Trust level and feed into NHS funding. ASHE data are used in the process by which £85 billion is allocated to Local Health Authorities (figure correct at July 2011).

Some examples of non-government uses, for which users have confirmed that ASHE data are fit for purpose are:

- Trade Unions - statistics used to support pay negotiations;
- media - statistics covered widely in various media organisations on television, radio, in newspapers and online;
- legal - figures used to inform judgements about payments for loss of earnings or where home care is required; and
- academia - data used in a range of economic and labour market research.

Users have said that the primary strengths of the ASHE dataset and the associated suite of outputs are:

- size & coverage - the ASHE dataset contains information on approximately 180,000 jobs in all industries, occupations and regions, making it the most comprehensive source of earnings information in the UK and enabling a vast range of analyses;
- quality - alternative sources of earnings information such as the Labour Force Survey (LFS) rely on self-report or proxy data, which are known to be less reliable than information from employers' administrative systems; and
- uniqueness - for many uses, ASHE is the key data source, and for some uses it is the only data source.

Users have said that the primary weaknesses of the ASHE dataset and the associated suite of outputs are:

- lack of personal demographic information such as ethnicity, religion, education, disability and pregnancy;
- timing and periodicity - for some users, more regular and timely results would be beneficial;
- ASHE does not cover the self-employed; and
- the quality of estimates at low levels of disaggregation can be poor.

A summary of users' responses to consultation about the relevance of the ASHE data and outputs is included in the Assessment of user needs and perceptions' section.

Timeliness and punctuality

(Timeliness refers to the lapse of time between publication and the period to which the data refer. Punctuality refers to the gap between planned and actual publication dates.)

The survey reference date for ASHE is in April of each year. Provisional results for ASHE and Low Pay, which contains estimates for the number of jobs paid below the national minimum wage, are published in November of the same year and revised results are published in November of the following year. (The ASHE workplace pensions statistics are published in February in the following year. ASHE releases usually meet target publication deadlines, though there have been some exceptions to this where it has been necessary to delay publication by a few weeks in order for ONS to undertake additional quality assurance.)

For more details on related releases, the [UK National Statistics Publication Hub](#)³ is available online and provides 12 months' advance notice of release dates. If there are any changes to the pre-announced release schedule, public attention will be drawn to the change and the reasons for the change will be explained fully at the same time, as set out in the [Code of Practice for Official Statistics](#)⁴.

How the output is created

Sampling and data collection

The survey uses a random sample of 1% of all employee jobs from HMRC's PAYE system, taken in January of the reference year. The sample is drawn in such a way that many of the same individuals are included from year to year, thereby allowing longitudinal analysis of the data. The sample is matched against ONS's Inter-Departmental Business Register (IDBR) in order to obtain contact and address details for the employers. Information on the hours paid and earnings of employees is obtained from employers and treated confidentially. The survey has a reference date in April and asks about individuals who were employees at that time. The reference date changes each year depending on when Easter falls. A second extract is taken from the PAYE system in April in order to identify people who have either joined the labour market or changed jobs since the January sample was taken.

ONS has a special arrangement with some very large employers for them to provide electronic returns extracted from their employee records in April. These employees are selected on the same basis as the regular ASHE sample.

Since ASHE is a survey of employee jobs, it does not cover the self-employed or any jobs within the armed forces. Given the survey reference date in April, the survey does not fully cover certain types of seasonal work, for example employees taken on for only summer or winter work.

Validation is carried out on returned data that is regarded as incomplete or potentially inaccurate, based on automatic comparisons with data for similar jobs or against data for the same job in previous years. In these cases, respondents may be re-contacted by ONS in order to verify the information that has been provided.

Weighting

Returned data are weighted to UK population totals from the LFS based on classes defined by occupation, region, age and sex. There are two processes involved in the weighting of responses for ASHE. The first allocates individual cases a design weight to adjust for non-response. For this purpose, responses are treated as being in one of four strata, depending on whether they were part of the original questionnaire despatch, one of the later supplementary surveys or have a special arrangement in place with ONS to return their data electronically.

For the second part of the weighting, the final file of responses is post-stratified to population estimates taken from the LFS in 108 post-strata. These post-strata are defined as a cross-classification of:

- occupation (9 groups) - major groups from Standard Occupational Classification 2000;
- age-band (3 groups) - 16-21, 22-49, 50+;
- gender (2 groups) - male and female; and
- region (2 groups) - London & South East and the rest of the UK.

In order to produce estimates for the number of jobs falling below the National Minimum Wage, the dataset is re-weighted to exclude employees whose earnings were affected by absence during the reference pay period.

Imputation

Since the introduction of weighting for ASHE data, the problem of item non-response (that is, where a questionnaire is returned by a respondent, but in an incomplete form) has become a significant issue when processing ASHE data. A method of imputation, 'donor imputation', has been adopted. In this process, records with similar characteristics are sought to act as 'donors' for missing variables.

Further information relating to ASHE methodology may be found on the [Annual Survey of Hours and Earnings methodology and guidance page of the ONS website](#)⁵.

Results

ASHE analyses for weekly and hourly earnings relate to employees on adult rates whose earnings for the survey pay period were not affected by absence. They do not include the earnings of those who did not work a full week, and whose earnings were reduced because of, for example, sickness. ASHE analyses for annual earnings relate to employees on adult rates of pay who have been in the same job for more than one year.

Estimates on the structure and distribution of earnings by a range of different breakdowns covering gender, occupation, industry, age and geography are produced from the survey. ASHE data are also used to produce estimates on the number of jobs paid below the national minimum wage and a range of statistics on membership of workplace pension schemes.

Statistical disclosure control

[Statistical disclosure control methodology](#)⁶ is applied to all outputs produced from ASHE. This ensures that information attributable to an individual or individual organisation is not identifiable in any published outputs. The [Code of Practice for Official Statistics](#)⁴, and, specifically, the principle on confidentiality, set out practices for how we protect data from being disclosed. The principle includes the statement that ONS outputs should - 'ensure that official statistics do not reveal the identity of an individual or organisation, or any private information relating to them, taking into account other relevant sources of information'.

Firstly, to protect individual earnings data, a frequency count is taken and all cells that are based on a count of fewer than three individuals are suppressed. Secondly, to protect employers' pay information, a dominance rule is applied within each cell, which uses the contribution from the

largest employer and the overall standard error of the estimate to deduce whether information about the employer can be derived to any accuracy.

Given the nature and complexity of ASHE outputs it is not possible to use a practical method to check for issues of secondary suppression. Instead, ASHE applies a policy where no sample counts are released, only weighted sample counts rounded to the nearest 1,000. This gives users enough information about the sample size for a cell for them to make quality inferences, without giving sufficient information to derive data by difference with any degree of certainty. Although in some circumstances a figure can be derived by difference, it would be impossible to tell how many individuals contributed to the figure.

Key information

| | |
|----------------------------|-----------------------------------------------------------------------------------------------------------------------------------|
| What it measures | Estimates of the structure and distribution of earnings and paid hours for employee jobs in the UK |
| Frequency | Annual |
| Sample size (numbers used) | 180,000 employee jobs |
| Periods available | 2004-2013 (ASHE inc supplementary surveys) 1997-2004 (ASHE excl supplementary surveys) 1970-2003 (NES unweighted estimates) |
| Sample frame | HMRC PAYE system |
| Sample design | 1% simple random sample with a longitudinal element |
| Weighting | Employee jobs are weighted to population totals taken from the LFS based on classes defined by occupation, region, age and sex |
| Imputation | Donor imputation carried out for missing data items on partial returns. No non-responder imputation carried out |
| Outliers | No outlier treatment. All validated values are used in estimates |

Validation and quality assurance

Accuracy

(The degree of closeness between an estimate and the true value.)

Estimates from this survey are subject to various sources of error. Total error consists of two elements, the sampling error and the non-sampling error.

Sampling error

This occurs because estimates are based on a sample rather than a census. ASHE estimates this error through coefficients of variation (cv) which are published alongside all ASHE outputs. The cv is the ratio of the standard error (se) of an estimate to the estimate itself, expressed as a percentage. Generally, if all other factors are constant, the smaller the cv the higher the quality of the estimate.

In published tables ASHE uses colour coding as a quick reference guide to the cv of the estimates; estimates with cvs less than or equal to 5% are published with no colour fill; estimates with cvs between 5% and 10% are published with a light green background; estimates with cvs between 10% and 20% are published with a dark green background; cells for which estimates have been suppressed on quality or disclosure grounds are also filled in dark green.

| Key |
|----------------------------------------|
| cv ≤ 5% |
| cv > 5% and ≤ 10% |
| cv > 10% and ≤ 20% |
| x = unreliable cv > 20% or unavailable |

High-level coefficients of variation for the 2013 ASHE estimates are shown below:

Coefficients of variation for estimates of median gross weekly earnings and hourly earnings (excluding overtime), UK, April 2013.

| | | CV(%) | | |
|----------------------------------------------------------|-------|---------------|-----------|-----------|
| | | All employees | Full-time | Part-time |
| Median gross weekly earnings | Men | 0.2 | 0.3 | 0.7 |
| | Women | 0.2 | 0.3 | 0.5 |
| | All | 0.2 | 0.2 | 0.4 |
| Median gross hourly earnings (excluding overtime) | Men | 0.2 | 0.3 | 0.5 |
| | Women | 0.2 | 0.3 | 0.4 |
| | All | 0.2 | 0.2 | 0.3 |

It should be noted that at low levels of disaggregation, high coefficients of variation imply estimates of low quality. For example, for an estimate of £400 with a cv of 10%, the true value is likely to lie between £321.60 and £478.40. This range is given by the estimate $\pm 1.96 \times se$. Where these ranges for different estimates overlap, interpretation of differences between the relevant domains becomes more difficult.

Non-sampling error

ASHE statistics are also subject to non-sampling errors. For example, there are known differences between the coverage of the ASHE sample and the target population (that is, all employee jobs). Jobs that are not registered on PAYE schemes are not surveyed. These jobs are known to be different from the PAYE population in the sense that they typically have low levels of pay. Consequently, ASHE estimates of average pay are likely to be biased upwards with respect to the actual average pay of the employee population. Non-response bias may also affect ASHE estimates. This may happen if the jobs for which respondents do not provide information are different from the jobs for which respondents do provide information. For ASHE, this is likely to be a downward bias on earnings estimates since non-response is known to affect high-paying occupations more than low-paying occupations.

Finally, ASHE results tables do not account for differences in the composition of different 'slices' of the employee workforce. For example, figures for the public and private sectors include all jobs in those sectors and are not adjusted to account for differences in the age, qualifications or seniority of the employees or the nature of their jobs, all factors which may affect how much employees earn. Various procedures are in place to minimise errors in returned data. Returns undergo a range of checks which include validation against previous returns and expected values, selective editing (a technique for prioritising suspicious values for follow-up based on their impact on published results) and re-contacting businesses for verification. Similar checks are also made at the aggregate level for key results.

Revisions

Provisional results are published in the November following the survey reference date. Revised results are then published one year later alongside the following year's provisional results. The revised results take account of late returns to the survey and amendments to data resulting from validating returns to the current year's survey.

Revisions are usually quite small, with revision at the UK level typically around 0.1%. However, estimates for domains with smaller sample sizes are susceptible to larger revisions.

Coherence and comparability

(Coherence is the degree to which data that are derived from different sources or methods, but refer to the same topic, are similar. Comparability is the degree to which data can be compared over time and domain, for example geographic level.)

The Average Weekly Earnings (AWE) statistics, based on the Monthly Wages and Salaries Survey (MWSS) of about 9,000 employers, is the lead measure of short-term changes in average earnings in Great Britain. Figures are available with industrial breakdowns and public/private sector splits. No information is available on occupation, hours worked, and other characteristics of the workforce.

The AWE and ASHE are not directly comparable on all measures of earnings. The closest measure that can be derived and compared for these surveys is for mean gross weekly pay in Great Britain. In the year to April 2013, the ASHE estimate of mean gross weekly pay for all employees (regardless of whether they worked full-time or part-time) was £504, up 2.2% on the previous year. The corresponding estimate from the AWE was £485, up 3.9% from April 2012. However, it should be noted that the relatively high percentage increase for AWE partly reflects unusually high bonus payments in April 2013 as a result of some businesses deferring some bonuses that would normally have been paid in March 2013 to April 2013.

The LFS collects information on the earnings and normal and actual hours worked of about 15,000 people aged 16 and over each quarter. In addition it collects data on a wide range of personal characteristics, including education level and ethnic origin. This information is collected from the employees by means of a household survey. Consequently, there are a number of important differences between LFS and ASHE since ASHE data usually come directly from company records. This means that ASHE collects the actual amounts paid and the number of hours that the employee has been paid for. It classifies the employee's industry from the business perspective, which in the case of ASHE is the main business of the enterprise. In addition, descriptions of individuals' occupations are provided by the business. In contrast, LFS data is collected from the employee, or a member of the employee's household. This can lead to proxy and estimated responses, for example if someone responds on behalf of a partner who is not available at the time. This enables the production of statistics on levels and distribution of earnings similar to those from ASHE but with lower precision because of the much smaller sample size and less reliable data collection methods (since self-report and proxy data are not as reliable as data from employers' administrative systems).

Concepts and definitions

(Concepts and definitions describe the legislation governing the output, and a description of the classifications used in the output.)

Respondents are required by law to provide information for ASHE in accordance with the Statistics of Trade Act 1947.

As well as their use in a range of National Statistics on earnings, low pay and workplace pensions membership, the data are also the primary source for the UK results for the [European Labour Costs Survey \(LCS\)](#)⁷ and the [European Structure of Earnings Survey \(SES\)](#)⁸, which are required under regulation to Eurostat, the European statistical agency.

ASHE uses the principal UK coding schemes for classifying occupation and industry. Occupations are coded using the Standard Occupational Classification (SOC) 2010 and industry using the Standard Industrial Classification 2007 (SIC 2007). Details of these classifications can be found on the [classifications section of the ONS website](#)⁹.

Other information

Assessment of user needs and perceptions

(The processes for finding out about uses and users, and their views on the statistical products.)

Government users

ONS has established a survey user group which meets annually and includes representatives from 23 teams in 17 government departments. This group discusses changes to the questionnaire, methodology and survey outputs. ONS also meets with some of these users as and when necessary.

ONS conducted a triennial review of ASHE in 2011 which included a user questionnaire. With few exceptions, users said that they regarded the ASHE data and outputs as 'good' or better in respect of their accuracy, relevance, timeliness, coherence, comparability and accessibility.

Non-government users

ONS conducted four strands of research into the uses made of ASHE by non-government users between 2012 and 2013. The outcome of this work is discussed in a paper, 'Non-government uses of ASHE', which was published on the [ONS website](#)¹ in March 2013. Feedback from these users indicated that the majority were satisfied or very satisfied with the extent to which ASHE outputs met their needs.

Sources for further information or advice

Accessibility and clarity

(Accessibility is the ease with which users are able to access the data, also reflecting the format in which the data are available and the availability of supporting information. Clarity refers to the quality and sufficiency of the release details, illustrations and accompanying advice.)

ONS's recommended format for accessible content is a combination of HTML web pages for narrative, charts and graphs, with data being provided in usable formats such as CSV and Excel. The ONS website also offers users the option to download the narrative in PDF format. In some instances other software may be used, or may be available on request. Available formats for content published on the ONS website but not produced by the ONS, or referenced on the ONS website but stored elsewhere, may vary. For further information please refer to the contact details at the beginning of this document.

For information regarding conditions of access to data, please refer to the links below:

- [Terms and conditions \(for data on the website\)](#)¹⁰;
- [Copyright and reuse of published data](#)¹¹;
- [Pre-release access \(including conditions of access\)](#)¹²; and
- [Accessibility](#)¹³.

In addition to this Quality and Methodology Information, Basic Quality Information relevant to each release is available in the background notes of the [ASHE statistical bulletin](#)¹⁴, the [Low Pay statistical bulletin](#)¹⁵ and the [ASHE Pensions statistical bulletin](#)¹⁶.

Useful links

[ASHE guidance and methodology](#)¹.

[All editions of the ASHE publication](#)¹⁷.

[All editions of the Low Pay publication](#)¹⁸.

[All editions of the ASHE pensions publication](#)¹⁹.

References

| | |
|-------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1. ASHE guidance and methodology page | http://www.ons.gov.uk/ons/guide-method/method-quality/specific/labour-market/annual-survey-of-hours-and-earnings/index.html |
| 2. Guidelines for measuring statistical quality | http://www.ons.gov.uk/ons/guide-method/method-quality/quality/guidelines-for-measuring-statistical-quality/index.html |
| 3. UK National Statistics publication hub | http://www.statistics.gov.uk/hub/index.html |
| 4. Code of Practice for Official Statistics | http://www.statisticsauthority.gov.uk/assessment/code-of-practice/index.html |
| 5. Specific methodology and guidance | http://www.ons.gov.uk/ons/guide-method/method-quality/specific/labour-market/annual-survey-of-hours-and-earnings/annual-survey-of-hours-and-earnings/index.html |
| 6. Statistical disclosure control methodology | http://www.ons.gov.uk/ons/guide-method/method-quality/general-methodology/statistical-disclosure-control/index.html |
| 7. European Labour Costs Survey (LCS) | http://epp.eurostat.ec.europa.eu/portal/page/portal/labour_market/labour_costs |
| 8. European Structure of Earnings Survey (SES) | http://epp.eurostat.ec.europa.eu/portal/page/portal/microdata/ses |
| 9. Current standard classifications | http://www.ons.gov.uk/ons/guide-method/classifications/current-standard-classifications/index.html |

| | |
|---------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 10. Terms and conditions (for data on the website); | http://www.ons.gov.uk/ons/site-information/information/terms-and-conditions/index.html |
| 11. Copyright and reuse of published data | http://www.ons.gov.uk/ons/site-information/information/creative-commons-license/index.html |
| 12. Pre-release access (including conditions of access) | http://www.ons.gov.uk/ons/guide-method/the-national-statistics-standard/code-of-practice/pre-release-access/index.html |
| 13. Accessibility | http://www.ons.gov.uk/ons/site-information/information/accessibility/index.html |
| 14. ASHE statistical bulletin | http://www.ons.gov.uk/ons/rel/ashe/annual-survey-of-hours-and-earnings/2013-provisional-results/stb-ashe-statistical-bulletin-2013.html |
| 15. Low Pay statistical bulletin | http://www.ons.gov.uk/ons/rel/ashe/low-pay/april-2013/index.html |
| 16. ASHE Pensions statistical bulletin | http://www.ons.gov.uk/ons/rel/ashe/annual-survey-of-hours-and-earnings-pension-tables/2013-provisional-results/stb--2013-annual-survey-of-hours-and-earnings--summary-of-pensions-results.html |
| 17. All editions of the ASHE publication | http://www.ons.gov.uk/ons/rel/ashe/annual-survey-of-hours-and-earnings/index.html |
| 18. All editions of the Low Pay publication | http://www.ons.gov.uk/ons/rel/ashe/low-pay/index.html |
| 19. All editions of the ASHE Pensions publication | http://www.ons.gov.uk/ons/rel/ashe/annual-survey-of-hours-and-earnings-pension-tables/index.html |

Information paper

Summary Quality Report for the Annual Survey of Hours and Earnings (ASHE)

1 Introduction

This report is part of a rolling programme of quality reports being introduced by the Office for National Statistics (ONS). The full programme of work being carried out on [Statistical Quality¹](#) is available on the National Statistics website. Summary Quality Reports are overview notes which pull together key qualitative information on the various dimensions of quality as well as providing a summary of methods used to compile the output.

This report relates to the Annual Survey of Hours and Earnings (ASHE), conducted by ONS. ASHE provides information about the levels, distribution and make-up of earnings and paid hours for employees in all industries and occupations. ASHE replaced the New Earnings Survey (NES) from 2004. ASHE brought improvements to the coverage of employees, imputation for item non-response and the weighting of earnings estimates.

2 Summary of Quality

2.1 Relevance

The degree to which the statistical product meets user needs for both coverage and content.

ASHE measures the average level and distribution of earnings and paid hours for employee jobs. Estimates are presented by a number of break-downs including gender, occupation, industry and region

| | |
|----------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| What it measures | Estimates of the level and distribution of earnings and paid hours for employee jobs. |
| Frequency | Annual |
| Sample Size (numbers used) | 181,000 employees |
| Periods available | 2004-2010 (ASHE inc supplementary surveys) 1997-2004 (ASHE excl supplementary surveys) 1970-2003 (NES unweighted estimates) |
| Sample Frame | HMRC PAYE Register |
| Sample Design | All employees whose National Insurance Number ends in a particular pair of digits are selected. ONS treats this as a one per cent simple random sample. |
| Weighting | Employee jobs weighted together based on classes defined by occupation, region, age and gender, to population totals taken from the Labour Force Survey (LFS²). |
| Imputation | Donor imputation carried out for missing data items on partial returns. No non-responder imputation carried out. More detail is provided in Section 3 of this report. |
| Outliers | No outlier treatment. All validated values are used in estimates. |

The survey is a sample survey of employee jobs, although information is collected from employers. It is based on a one per cent random sample of jobs on the HM Revenue and Customs' Pay As You Earn (HMRC PAYE) register. It covers all employee jobs in all industries and occupations across the whole of the United Kingdom. It has a reference date in April, asking about individuals who are employees at that time. This reference date changes each year depending on when Easter falls.

ASHE analyses for weekly and hourly earnings relate to employees on adult rates whose earnings for the survey pay period were not affected by absence. They do not include the earnings of those who did not work a full week, and whose earnings were reduced because of, for example sickness.

ASHE analyses for annual earnings relate to employees on adult rates of pay who have been in the same job for more than one year.

Since ASHE is a survey of employee jobs, it does not cover the self-employed or any jobs within the armed forces. Given the survey reference date in April, the survey does not fully cover certain types of seasonal work, for example employees taken on for only summer or winter work.

Estimates by a range of different breakdowns covering gender, occupation, industry, age and geography are produced from the survey. However, because information is collected from the employer, information about social variables that are not held by the employer, such as ethnic background, household composition or level of disability cannot be produced.

ASHE abides by the principal coding schemes used for classifying occupation and industry. Occupation is coded using the Standard Occupational Classification (SOC) 2000 and industry using the Standard Industrial Classification (SIC) 2007.

ASHE uses the HMRC PAYE register as a sampling frame. Consequently a number of employee jobs are not available for selection. This happens when no employee in the business earns above the threshold requiring the business to register for PAYE. Although this means that a number of part-time employees with low earnings are not available for selection, investigations have shown that these employees are paid similar amounts to other part-time employees working in the same occupations for businesses that are above the PAYE threshold.

One of the main uses of ASHE is to estimate the number of jobs falling below the National Minimum Wage (NMW). There are a number of circumstances under which an employee would not be subject to NMW legislation, for example if the employee is an apprentice or where accommodation is received as a benefit in kind. ASHE does not collect information to identify where these circumstances apply, and so the survey does not estimate non-compliance with NMW.

ASHE is also used as the main data source for the European Structure of Earnings Survey and the European Labour Costs Survey, which are carried out every 4 years.

2.2 Accuracy

The closeness between an estimated result and the (unknown) true value.

Estimates from this survey are subject to various sources of error. Total error consists of two elements, the sampling error and the non-sampling error.

Sampling error

This occurs because estimates are based on a sample rather than a census. ASHE estimates this error through coefficients of variation (cv) which are published alongside all [ASHE outputs](#)³. The cv is the ratio of the standard error of an estimate to the estimate itself, expressed as a percentage. The smaller the cv the higher the quality of the estimate.

In published tables ASHE uses colour coding as a quick reference guide to the cv of the estimates.

| Key |
|----------------------------------------|
| cv ≤ 5% |
| cv > 5% and ≤ 10% |
| cv > 10% and ≤ 20% |
| x = unreliable cv > 20% or unavailable |

Non-sampling error

Non-sampling errors are not easy to quantify and include errors of coverage, measurement, processing and non-response. ASHE minimises any non-response bias by imputing for item non-response and weighting for unit non-response.

Various procedures are in place to minimise errors in returned data. Returns undergo a range of checks which include validation against previous returns and expected values, selective editing and re-contacting businesses for verification. Similar checks are also made at the aggregate level for key results.

Revisions

Results are published in the December following the survey reference date. Revised results are then published one year later alongside the following year's provisional results. The revised results take account of late returns to the survey and amendments to data resulting from validating returns to the current year's survey.

Revisions are usually quite small, with revision at the UK level typically around 0.1%. However estimates for domains with smaller sample sizes are susceptible to larger revisions.

2.3 Timeliness and Punctuality

Timeliness refers to the lapse of time between publication and the period to which the data refer. Punctuality refers to the time lag between the actual and planned dates of publication.

The survey reference date for ASHE is in April of each year. The ASHE Statistical Bulletin is published in December of the same year, covering a number of headline statistics and information from the survey. ASHE releases consistently meet target publication deadlines.

The [National Statistics Release Calendar](#)⁴ is available on the National statistics website and provides twelve months advance notice of releases.

2.4 Accessibility and Clarity

Accessibility is the ease with which users are able to access the data, also reflecting the format(s) in which the data are available and the availability of supporting information. Clarity refers to the quality and sufficiency of the metadata, illustrations and accompanying advice.

ASHE results are published as a number of standard web tables, in Excel format. Due to the volume of data, these tables are made available on the National Statistics website by the end of the calendar year. Additionally, all of the tables are available from the Earnings Helpdesk following the publication of the first wave of tables.

The ASHE results can be accessed on the [National Statistics Website](#)³

The Earnings Helpdesk team can be contacted by telephone: 01633 456120 or via email: earnings@ons.gov.uk⁵.

2.5 Comparability

The degree to which data can be compared over time and domain.

An earnings survey has been carried out each year since 1970. However, over the years there have been a number of changes to coverage, methodology and classification conventions used for the survey.

The ASHE results for 1997-2003 have been produced by reworking data provided for the NES using [ASHE methodology](#)⁶.

In 2004 additional supplementary surveys were introduced to improve the coverage of ASHE. These focus on employees that have started a new job between the original ASHE sample extract being produced in January and the survey reference date in April. Results for 2004-2006 are produced including these new supplementary surveys.

Data for 2004 have also been produced excluding the new supplementary surveys so that it is comparable with the data for 2003 and earlier.

In 2006 ASHE moved to the ONS standard for geographic areas using Output Areas (OAs) as building blocks to higher level geographic areas. This improves comparability with other ONS geographic results and allows further geographic analysis to be produced.

In addition to this, from 2006 the LFS moved from using seasonal quarters to calendar quarters. As ASHE uses LFS data in the calculation of aggregation weights, it was necessary to use data from LFS Quarter 2 rather than LFS Spring quarter. A consistent back-series which takes into account both the new geographies and LFS weights has been produced going back to 2004.

Two methodological changes took place during the collection and processing of the 2007 ASHE. The automatic coding tool, Automatic Coding by Text Recognition (ACTR) has been introduced for assigning Standard Occupational Classification (SOC) codes to ASHE records. This replaces the current system whereby SOC codes are carried forward for employees who were in the same job as the previous year, or manually allocated by a team of coders if the employee was either not included in ASHE for the previous year or has subsequently changed jobs. The second change is to treat large businesses that have arrangement in place with ONS to return their data electronically, as a separate stratum in the ASHE weighting. These businesses have a response rate significantly higher than the returns from the paper questionnaires sent out in the original despatch. This meant that SA records were receiving a higher weight than they ideally ought to. Treating businesses with SAs as a separate stratum allows ONS to allocate more appropriate weights to them.

In March 2007, ONS released information on its statistical work priorities over the period 2007-2008. ONS announced that the sample size of the ASHE was to be reduced by 20%. The sample size was reduced for the 2007 and 2008 survey periods, with results based on approximately 140,000 returns, down from 175,000 in 2006.

The full 1 per cent sample has since been restored and ASHE results have been produced on this basis since 2009.

2.6 Coherence

The degree to which data that are derived from different sources or methods, but which refer to the same phenomenon, are similar.

The most comparable results to those obtained from ASHE are figures from the LFS. The LFS also collects information on hours and earnings, although it is collected from the employee by means of a household survey, rather than the data being supplied by the employer, as is the case with ASHE.

Consequently there are a number of important differences between LFS and ASHE. The data collected from ASHE usually comes directly from company records. This means that ASHE collects the actual amounts paid and the number of hours that the employee has been paid for. It classifies the employee's industry from the business perspective, which in the case of ASHE is the main business of the enterprise. In addition, descriptions of individuals' occupations are provided by the business.

In contrast, LFS data is collected from the employee, or a member of the employee's household. This can lead to proxy and estimated responses, for example if someone responds on behalf of a partner who is not available at the time. The LFS also collects information on hours worked, rather than hours paid.

The Average Weekly Earnings (AWE) statistics, based on the Monthly Wages and Salaries Survey of about 9,000 employers, are the lead measure of short-term changes in average earnings in Great Britain. Figures are available with industrial breakdowns and public/private sector splits. No information is available on occupation, hours worked, and other characteristics of the workforce.

3 Summary of Methods Used to Compile the Output

Coverage

The ASHE covers employee jobs across the whole of the UK. The survey covers employees from all industries and occupations, but does not cover the self employed or the armed forces. The survey reference date falls in April, hence the employees covered are only those that are active in that month, which can distort the numbers in some seasonal occupations.

Sample Frame and Design

The ASHE sample is drawn from the HMRC PAYE register. All employees with National Insurance Numbers ending in a particular pair of digits are sampled. This can then be treated as a one per cent simple random sample of employees in all industries and occupations. This usually gives an original sample size of around 260,000.

The original sample is taken from the HMRC register in late January. However, the survey reference date is in April. Consequently a second extract is taken from the HMRC register in April that is used to identify people who have newly joined employers between those dates. These people are split between those on the April extract who have changed jobs since January and those who did not appear on the January extract at all.

In 2010, following the introduction of a new PAYE system, there was a discontinuity in the way that employments were created. As a consequence, the initial ASHE sample was approximately 35% larger than expected, implying an extra 10 million jobs in the UK workforce. In consultation with HMRC, ONS took steps to identify and remove 'suspect' records from the sample. Quality-assurance demonstrated that the final sample, which contained approximately 281,000 records, and the final dataset of responses were very similarly distributed across age, sex, industry and business size-band categories.

The ONS has a special arrangement with some very large employers for them to provide electronic returns extracted from their employee records in April, selected on the same basis as the normal ASHE sample.

Imputation

Since the introduction of weighting for ASHE methodology, the problem of item non-response (i.e. where a questionnaire is returned by a respondent, but in an incomplete form) has become a significant issue when processing ASHE data. A method of imputation, 'donor imputation', has been adopted. In this process, records with similar characteristics are sought to act as 'donors' for missing variables.

Weighting

There are two processes involved in the weighting of responses for ASHE. The first allocates individual cases a design weight to adjust for non-response. For this purpose, responses are treated as being in one of four strata, depending on whether they were part of the original questionnaire despatch, one of the later supplementary surveys or have a special arrangement in place with ONS to return their data electronically.

For the second part of the weighting, the final file of responses is post-stratified to population estimates taken from the LFS in 108 post-strata. These post-strata are defined as a cross-classification of:

- Occupation (9 groups) - major groups from Standard Occupational Classification 2000
- Age-band (3 groups) - 16-21, 22-49, 50+
- Gender (2 groups) - male and female
- Region (2 groups) - London & South East and the rest of the UK.

In order to produce estimates for the number of jobs falling below the National Minimum Wage, the dataset is re-weighted to exclude employees whose earnings were affected by absence during the reference pay period.

Statistical Disclosure

[Statistical disclosure control methodology](#)⁷ is applied to all outputs produced from the ASHE. This ensures that information attributable to an individual or individual organisation is not identifiable in any published outputs. The [Code of Practice for National Statistics](#)⁸, and specifically the Principle on Confidentiality set out practices for how we protect data from being disclosed. The Principle includes the statement that ONS outputs should “ensure that official statistics do not reveal the identity of an individual or organisation, or any private information relating to them, taking into account other relevant sources of information”.

Firstly, to protect individual earnings, a frequency count is applied suppressing all cells that are based on a count of less than three individuals. Secondly, to protect employer pay information, a dominance rule is applied, which uses the contribution from the largest employer and the overall standard error of the estimate to deduce whether information about the employer can be derived to any accuracy.

Given the nature and complexity of ASHE outputs it is impossible to use a practical method to check for issues of secondary suppression. Instead ASHE applies a policy where no sample counts are released, only weighted sample counts rounded to the nearest 1,000. This gives users enough information about the sample size for a cell for them to make quality inferences, without giving sufficient information to derive data by difference with any degree of certainty. Although in some circumstances a figure can be derived by difference, it would be impossible to tell how many individuals contributed to the figure.

4 References

| | Title of Reference | Website location |
|---|--------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1 | Statistical Quality | http://www.ons.gov.uk/ons/guide-method/method-quality/quality/index.html |
| 2 | Labour Force Survey | http://www.ons.gov.uk/ons/rel/lms/labour-market-statistics/august-2011/index.html |
| 3 | Annual Survey of Hours and Earnings | http://www.ons.gov.uk/ons/rel/ashe/annual-survey-of-hours-and-earnings/2010-results/index.html |
| 4 | National Statistics Release Calendar | http://www.statistics.gov.uk/hub/index.html |
| 5 | Earnings Helpdesk | Tel: 01633 819024; email earnings@ons.gov.uk |
| 6 | Methodology for the 2004 ASHE | http://www.ons.gov.uk/ons/guide-method/method-quality/specific/labour-market/annual-survey-of-hours-and-earnings/index.html |
| 7 | Statistical Disclosure Control Methodology | http://www.ons.gov.uk/ons/guide-method/method-quality/general-methodology/statistical-disclosure-control/index.html |
| 8 | National Statistics Code of Practice | http://www.ons.gov.uk/ons/guide-method/the-national-statistics-standard/code-of-practice/index.html |