

The effects of taxes and benefits on UK household income

Contact:
hie@ons.gsi.gov.uk


Release date:
23 October 2015

Next release:
To be announced

Table of contents

1. [Methodology Background](#)
2. [Overview](#)
3. [Executive summary](#)
4. [Output quality](#)
5. [About the output](#)
6. [How the output is created](#)
7. [Validation and quality assurance](#)
8. [Concepts and definitions](#)
9. [Sources for further information or advice](#)

1 . Methodology Background

National Statistic	
Survey name	Living Costs and Food Survey
Frequency	Annual
How compiled	Voluntary sample based survey and administrative data
Geographic coverage	UK
Sample size	5,000 households
Last revised	23 October 2015

2 . Overview

- provides estimates of household incomes, including the average amount of taxes that households pay and the value of benefits that they receive
- covers private households (not including people living in hotels, lodging houses, and in institutions such as old people's homes)
- has been produced for over 50 years

The Effects of Taxes and Benefits on Household Income (ETB) has been produced each year since 1961. Its main purpose is to provide quantitative analysis of the effects of government intervention (through taxes and benefits) on the income of households in the UK.

Data are published in both an annual statistical bulletin and a supplementary [Methodology and Coherence article](#). The final anonymised microdata are supplied to the [UK Data Archive](#) under an end user licence.

The estimates in the analysis are based mainly on data derived from the [Living Costs and Food Survey \(LCF\)](#). The LCF is an annual survey of the expenditure and income of private households.

ETB uses a number of administrative sources to improve the quality of estimates, particularly to estimate income and benefits in kind. The data cover the UK as a whole, with a number of published estimates at a regional level. The ETB also provides estimates for retired and non-retired households.

The ETB provides data and analysis on topics which are not fully covered elsewhere (the effect of indirect taxes on household income, and how this serves to reduce inequality, as well as benefits in kind)

3 . Executive summary

The [Effects of taxes and benefits on UK household income](#) (ETB) has been produced each year since 1961 and is an annual analysis looking at how taxes and benefits affect the income of households in the UK. It provides estimates of household incomes, including the average amount of taxes that households pay and also the value of benefits that they receive.

Data are from the [Living Costs and Food Survey](#) (LCF), formally known as the Expenditure and Food Survey (EFS), which is a voluntary sample survey of around 5,000 private households in the UK. Data are published in both an annual [statistical bulletin](#) and a supplementary [methodology and coherence article](#). The final anonymised microdata are also supplied to the [UK Data Archive](#) under end-user licence.

ETB remains valuable as it provides data and analysis on topics that are not sufficiently covered elsewhere (the effect of indirect taxes on household income and how this serves to reduce inequality, as well as “benefits in kind”).

This report 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
- Sources for further information or advice

4 . Output quality

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

We have developed [Guidelines for Measuring Statistical Quality](#); these are based upon the five European Statistical System (ESS) Quality Dimensions. This report addresses the quality dimensions and important quality characteristics, which are:

- relevance
- timeliness and punctuality
- comparability and coherence
- accuracy
- output quality trade-offs
- assessment of user needs and perceptions
- accessibility and clarity

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

5 . About the output

Relevance

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

The effects of taxes and benefits on UK household income (ETB) is one of the Office for National Statistics's (ONS's) longest-standing outputs and has been produced for over 50 years. Its main purpose is to provide quantitative analysis of the effects of government intervention (through taxes and benefits) on the income of households in the UK.

The data cover the UK as a whole, with a number of published estimates at a regional level. ETB also provides estimates for retired and non-retired households.

ETB data and estimates are used to provide information for policy-makers and analysts. The main users of these estimates in government are HM Treasury (HMT), the Department for Work and Pensions (DWP) and HM Revenue and Customs (HMRC). ETB has also contributed to ONS's Economic well-being release. Outside government, research organisations and academia use these data in their work on income distribution.

There are considerable difficulties in moving from estimates of government expenditure and financing published in the UK National Accounts: the Blue Book, to apportioning taxes and benefits to individual households. It is possible to get information about the types of households that receive benefits and pay taxes through the LCF. But there are other kinds of financing, such as Corporation Tax and government receipts from public corporations: no attempt is made in this analysis to apportion them to households because of the complexities involved.

Similarly, there are other items of government expenditure, such as capital expenditure and expenditure on defence and on the maintenance of law and order, for which there is no clear conceptual basis for allocation to households, or for which we do not have sufficient information to make an allocation.

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.)

Since 1994 to 1995, ETB estimates have been for the financial year, though prior to this date they were for calendar years. The publication and data are normally released in June, approximately 15 months after the end of the income reference period. A thorough review of the production process has been carried out in order to identify opportunities for producing these figures in a timelier manner.

Producing estimates of indirect taxes and social transfers in kind is the most time-consuming part of the production process due to the complexity of the methodologies and required updates. It is, however, possible to get to a measure of disposable income relatively quickly. By making changes to existing processes, it was anticipated that it would be possible to move to producing two publications:

- a concise, earlier publication focused on the distribution of the breakdown of disposable income
- the main publication, which would include the full range of analysis we currently produce on the redistribution of income through the tax and benefit systems

The earlier publication, [Household disposable income and inequality](#), was published for the first time in February 2016 with the publication of ETB following at its usual time in May. This approach is consistent with the [Code of Practice for Official Statistics](#): in particular it is consistent with Protocol 2 regarding the release of statistical reports as soon as they are judged ready, so that there is no opportunity, or perception of opportunity, for the release to be withheld or delayed; and Principle 1, Practice 4, which says that statistical reports should be published according to a timetable that takes account of user needs.

There are plans to move the LCF from reporting on a calendar year basis to a financial year basis. This provides the potential for further improvements and will remove one of the main challenges that currently exist in producing publication-ready data in a timely manner.

Notification of the provisional date on which data are due for publication is made approximately 1 year in advance. Notification of the exact date on which data are published each year is made public approximately 3 months beforehand via the [GOV.UK release calendar](#). To date, publication of ETB data has occurred without delay.

For more details on related releases, the GOV.UK release calendar provides 12 months' advanced 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.

6 . How the output is created

The estimates in this analysis are based mainly on data derived from the Living Costs and Food (LCF) Survey, which replaced the Family Expenditure Survey from 2001 to 2002 and was known as the Expenditure and Food Survey until 2008. The LCF is an annual survey of the expenditure and income of private households. People living in hotels, lodging houses and in institutions such as old people's homes are excluded. Each person aged 16 and over keeps a full record of payments made during 14 consecutive days and answers questions about hire purchase and other payments; children aged 7 to 15 keep a simplified diary.

The respondents also give detailed information, where appropriate, about income (including cash benefits received from the state) and payments of Income Tax. Information on age, occupation, education received, family composition and housing tenure is also obtained. The survey covers the whole 12-month period. The [Family spending](#) publication also includes an outline of the survey design.

The number of households in Great Britain responding in full to the LCF in 2013 was 4,761 and a further 232 households provided enough information to be included in the sample. The response rate was 48%. An additional sample of 152 households covered Northern Ireland, where the response rate was 61%. To count as a co-operating household, all members aged 16 and over must fill in the diaries for both weeks and give full details of income.

The LCF data used in this analysis are grossed so that totals reflect the total population of private households in the UK. The weights are produced in two stages. First the data is weighted to compensate for non-response (sample-based weighting). The non-response weights are then calibrated so that weighted totals match population totals for males and females in different age groups and for different regions and countries (population-based weighting). The results in the analysis are weighted so that statistics represent the total population in private households in the UK based on 2011 Census data. In 2013 to 2014, an additional calibration to Labour Force Survey (LFS) employment totals was also applied.

The basic unit of analysis used is the household. The starting point of the analysis is original income. This is the annualised income in cash of all members of the household before the deduction of taxes or the addition of any state benefits. The next stage of the analysis is to add cash benefits and tax credits to original income to obtain gross income. Income Tax, Council Tax and Northern Ireland rates, and employees' and self-employed National Insurance contributions are then deducted to give disposable income. The next step is to deduct indirect taxes, such as VAT, to give post-tax income. Finally, the analysis adds benefits that are provided "in kind" to households by government, for which there is a reasonable basis for allocation to households, to obtain final income. These "in kind" benefits include the provision of education, health services and subsidised travel and housing.

The effects of taxes and benefits on UK household income (ETB) uses a number of administrative sources to improve the quality of estimates, particularly to estimate income and benefits in kind. A full list of the administrative data used is available on the [statement of administrative sources](#).

The households are ranked by their equivalised disposable income, which the analysis uses as a proxy for standard of living. Equivalisation is a process that adjusts households' incomes to take account of their size and composition, recognising that this affects the demand on resources. For example, a couple with a child would need a higher income than a childless couple for the two households to achieve the same standard of living. The McClements scale was used before 2009 to 2010 but since then the [modified-OECD scale](#) has been used, in line with other major surveys that collect income data.

7 . Validation and quality assurance

Accuracy

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

The main types of error that affect the accuracy of data used in these analyses are:

- sampling error
- coverage error
- non-response bias
- measurement error
- systems error
- editing error

More information is provided about these accuracy dimensions in the following sections.

Sampling error

Sampling error occurs as a result of the selection of a sample to represent a population. In most analysis it is not feasible to gain data on the whole population (a notable exception would be the census). While the Living Costs and Food (LCF) sample is designed to produce the "best" estimate of the true population values for income and expenditure, a number of equal-sized samples covering the population would generally produce varying population estimates.

Sampling error is typically less for measures based on large groups of households and those that do not vary greatly between households. Conversely, it is largest for small groups of households and for measures that vary considerably between households. A broad numerical measure of the amount of variability is provided by the standard error. There will be greater sampling variability associated with estimates for decile and quintile groups, and for particular household types mainly because the sample sizes are smaller. Decile and quintile groups for particular household types are even smaller, which will increase sampling variability further.

Coverage error

Coverage error occurs when households relevant to the population being analysed are not included within the sampling frame. The LCF draws its sample using the [Small User Postal Address File](#) (PAF).

It is acknowledged that this source contains some errors in content and in coverage. A reverse record check we conducted in 1994 used census data to show that coverage in the PAF was 93.0%. When including addresses that were incomplete, but that provided sufficient detail for an interview to be conducted, PAF coverage increased to 96.6%. Three-quarters of missing addresses in 1991 were still missing in 1993, suggesting missing data was not due to a time lag.

The make-up of the missing addresses is unknown and the omission of these addresses could provide some bias in the estimates. We update the sampling frame using the PAF on a 6-monthly basis. Where an address is sampled that does not fit the survey parameters it is removed, for example, a business address. The PAF is used as the sample frame for our social surveys, therefore any error or bias will be in line with other surveys.

The survey uses a complex stratified sample that draws sample characteristics from the 2001 Census. While the census is not a sample survey it does have its own sources of non-sampling error, for example, non-completion and incorrect response. Any bias from the census will also be reflected in this analysis.

Non-response bias

Non-response includes both households not responding at all to the survey (unit non-response) and households who participate in the survey but do not provide a response to particular questions (item non-response). If non-responders and responders have the same characteristics then there will be no bias.

Respondents may not answer specific questions that households deem private or personal. This is particularly relevant for the LCF, a survey that asks a variety of questions based on household income and expenditure. The response rate to the income questions in the LCF is fairly high; where respondents do not answer income questions they are generally not included in the survey as this is a fundamental part of the LCF.

Very little imputation is done for non-response to income questions. While there are a number of alternative sources of income data, such as the Family Resources Survey (FRS), all come with their own non-sampling error. Estimates from the Survey of Personal Incomes are thought to be quite robust, particularly for cases at the higher end of the income distribution.

The LCF assigns weights to cases to correct for unit non-response in the survey sample. A comparison was made of the households responding in the 1991 Family Expenditure Survey (FES) with those not responding, based on information from the 1991 Census of Population (see "A comparison of the census characteristics of respondents and non-respondents to the 1991 FES" by K Foster, ONS Survey Methodology Bulletin Number 38, January 1996).

Results from the study indicated that response was lower than average in Greater London, higher in non-metropolitan areas and that non-response tended to increase with increasing age of the head of the household, up to age 65. Households that contained three or more adults, or where the head was born outside the UK or was classified to an ethnic minority group, were also more likely than others to be non-responding. Non-response was also above average where the head of the household had no post-school qualifications, was self-employed, or was in a manual social class group.

The data were re-weighted to compensate for the main non-response biases identified from the 1991 Census comparison. We have completed a similar comparative exercise, with the 2001 Census data, which resulted in an update of the non-response weights for the 2007 and subsequent Expenditure and Food Survey (EFS) or LCF estimates. At present, the LCF is contributing towards the 2011 Census non-response linkage project, which will enable non-response weights to be updated.

A calibration weight is also calculated, which ensures that the sample is reflective of the entire population when it is grossed to create population aggregates. This uses 2011 Census-based population projections. Weighted totals match population totals for males and females, in different age groups and for regions and countries in the UK.

Factors influencing non-response that are within our control are the survey design and the interviewer characteristics. Any proposed changes to the survey design are considered at length both to see if there is really a need for the information that is being collected and to assess the effect on respondent burden. In 2013 the average interview length was 73 minutes.

To increase the incentive to participate, we provide a book of stamps with the initial invitation. Furthermore, those who complete the expenditure diary receive a small token. From January 2010, this was a £10 high street voucher (£5 for children); prior to 2010 this incentive was given in cash. The effectiveness of these measures is frequently monitored.

Interviewers receive full training aimed at increasing participation and the accuracy of the data that is collected and interviewers call at different times of the day to attempt to maximise participation from sample households. Interviewers receive reasonable quotas to ensure that they are able to work each case effectively and maximise potential participation.

The LCF is conducted using computer-assisted personal interviewing (CAPI) in common with all ONS social surveys; this helps to eliminate item non-response occurring due to routing errors in the administering of the questionnaire. Since 2001 to 2002, proxy responses have been accepted in some cases. Proxy cases occur where one member of the household answers questions on behalf of another member of the household. The inclusion of proxy data reduces non-response but may increase error relating the accuracy of the response to the true value. The percentage of fully responding households with a proxy interview in Great Britain in 2013 to 2014 was 26%, which is just over double the 2001 to 2002 level.

Some of the data that are used in the [Effects of taxes and benefits on UK household income](#) (ETB) are subject to imputation; the methodology for this is outlined throughout this report. However, in general terms, the use of imputation is likely to result in an increase in the non-sampling error. Often these imputations make use of both administrative and survey data together. The limitations of the survey data have already been discussed. There are details of how to access a list of administrative sources used in the data sources section. Each of these data come with their own non-sampling error, although there will usually be fewer sources of such error in administrative data than in survey data.

Measurement error

Measurement error occurs when reported survey responses are different from the true value. This can occur for a variety of reasons, but the LCF take a number of steps to minimise this error. In some cases, the respondent may be unable or unwilling to provide a true answer to the question. This is particularly relevant in areas that are sensitive, related to the LCF income questions.

Respondents are encouraged to consult their payslip where possible to aid the provision of accurate information. Measurement error can also occur if the question is unclear or if participants are unable to understand the question; this is addressed in the LCF through extensive testing of new questions, including cognitive testing. A recent example of cognitive testing is a new question on combined utility expenditure, data from which is used in estimates of indirect taxation. The use of CAPI minimises collection error, but it may be off-putting compared with other methods that allow anonymity and less pressure on interviewer time. A further source of measurement error is the participant's response to the interviewer; in some cases the socio-economic characteristics of the interviewer make the participant feel uncomfortable in giving a true answer.

Assurances are given to respondents that their data will be treated in line with the [Code of Practice for Official Statistics](#) and the practicalities of what this means are explained. In the case of personal information, such as income and expenditure, this is particularly relevant; some respondents may report income that is in line with their tax returns rather than the true value. It is therefore likely that there will be some under-estimation of income.

Research suggests a larger level of under-reporting for self-employed income than income from wages and salaries. From an expenditure point of view, households may be reluctant to give true estimates of some items. As previously discussed, there is known under-reporting of alcohol, tobacco and confectionery, so an adjustment is made.

This analysis deals with some income concepts that may differ from the common perception of income. Steps have been taken to break down the questions on income to components to ensure that the data are collected on the desired level of conceptual accuracy. Exact figures are requested where possible; where these are not available estimates are allowed, notably income from self-employment and interest and dividend income. In some cases we anticipate that respondents may in reality provide a rounded figure. As previously stated, respondents are encouraged to consult documentation to increase the accuracy of their response.

Systems error

A number of the processes undertaken to conduct the LCF and the ETB analysis are automated. Therefore, there is a possibility that error could arise as a result of a mis-specification of some of these computerised processes. However, the data undergo rigorous quality assessment processes to look for any indication that an error has occurred; this enables most errors to be rectified at an early stage. Using systems saves resource and also limits non-sampling error that could occur as a result of carrying out the same processes manually. Aside from methodological improvements, the same processes are used year-on-year and are quality assured each year. Therefore, it seems unlikely that large error would emerge from the use of automated systems.

While the use of CAPI minimises data entry error it is still possible that keying errors can occur when the interviewer enters the response.

Editing error

The LCF undergoes a variety of editing checks from both the LCF and ETB teams. This is to ensure the quality of the data and to highlight and correct cases that are deemed to be in error. This process is usually automated with software flagging erroneous cases. The number of edited cases is small and changes are only made where it appears clear that there is an error. Data editing may also occur during the interview, with the interviewer flagging responses that do not appear to be consistent. The Blaise computer-assisted interviewing program that is used by ONS for social surveys will not allow an interview to proceed where a response is not allowed by the system and will flag with the interviewer responses that seem unlikely, in order that it can be queried at the point of interview.

Coherence and comparability

(Coherence is the degree to which data that are derived from different sources or methods, but that 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 LCF was previously carried out on a calendar year basis, whereas ETB is based on the financial year. As the LCF data for the final quarter of the financial year were not finalised by the time ETB was published, there were occasionally slight differences between the number of households in the calendar and financial year datasets. The LCF moved to a financial year reporting period from 2015 to 2016, which removed the differences between datasets.

The Department for Work and Pensions (DWP) publishes an analysis each year of the income distribution in their publication [Households below average income \(HBAI\)](#), based on data from the Family Resources Survey (FRS). ETB details information about its comparability with HBAI in the supplementary [Methodology and coherence report](#).

There are a number of different measures of income used, the most common of which is probably household disposable income. This is the total income households receive from employment (including self-employment), income from private pensions, investments and other sources, plus cash benefits (including the State Pension), minus direct taxes (including Income Tax, National Insurance and Council Tax). Income is normally analysed at the household level as this provides a better measure of people's economic well-being; while income is usually received by individuals, it is normally shared with other household members (for example, spouse or partner and children).

In contrast, earnings statistics generally refer to gross pay for employees (not self-employed), before tax and excluding any in-kind benefits. Earnings are typically reported at the individual level, for full-time employees.

The [McClements scale](#) was used before 2009 to 2010 for equivalisation purposes but since then the modified-OECD scale has been used, in line with other major surveys that collect income data.

The 2013 to 2014 publications were weighted based on 2011 Census data so that statistics represent the total population in private households in the UK. The weighting process also takes account of households that do not respond to the survey. Prior to 2013 to 2014 weighting was based on 2001 Census data.

8 . Concepts and definitions

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

UK income statistics (mainly produced by Office for National Statistics and Department for Work and Pensions), follow the definitions and concepts set out in the [Canberra Handbook](#) (published by United Nations Economic Commission for Europe), which is the basis of internationally agreed standards in this area. This defines income as receipts (either monetary or in kind) that are received on a regular basis and are available for current consumption.

9 . 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 \(for data on the website\)](#)
- [Copyright and reuse of published data](#)
- [Pre-release access \(including conditions of access\)](#)
- [Accessibility](#)

In addition to this Quality and Methodology Information, Basic Quality Information relevant to each release is available in the background notes of the [Effects of Taxes and Benefits Statistical Bulletin](#).

Useful links

[The Effects of Taxes and Benefits on Household Income - all editions](#)
[The Effects of Taxes and Benefits on Household Income - Historical Data, 1977-2013/134](#)
[Fifty years of the effects of taxes and benefits on household income](#)