

Wealth and Assets Survey QMI

Quality and Methodology Information for the Wealth and Assets Survey, detailing the strengths and limitations of the data, methods used, and data uses and users.

Contact:
Carla Kidd
wealth.and.assets.survey@ons.
gov.uk
+44 (0)1633 580088


Release date:
5 December 2019

Next release:
To be announced

Table of contents

1. [Methodology background](#)
2. [About this Quality and Methodology Information report](#)
3. [Important points](#)
4. [Quality summary](#)
5. [Quality characteristics of the Wealth and Assets Survey data](#)
6. [Methods used to produce the Wealth and Assets Survey data](#)
7. [Other information](#)

1 . Methodology background

National Statistic	
Survey name	Wealth and Assets Survey
Frequency	Biennial
How compiled	Longitudinal survey
Geographic coverage	Great Britain, excluding addresses north of the Caledonian Canal, the Scottish Islands and the Isles of Scilly
Last revised	5 December 2019

2 . About this Quality and Methodology Information report

This Quality and Methodology Information (QMI) report contains information on the quality characteristics of the data (including the European Statistical System's five dimensions of quality) as well as the methods used to create it.

The information in this report will help you to:

- understand the strengths and limitations of the data
- learn about existing uses and users of the data
- understand the methods used to create the data
- decide suitable uses for the data
- reduce the risk of misusing data

3 . Important points

- The latest release of [Wealth in Great Britain Round 6: 2016 to 2018](#) contains several estimates that have been adjusted for inflation for the first time within the release; all estimates within the previous Wealth in Great Britain reports are presented as current values (that is, the value at time of interview) and have not been adjusted for inflation.
- All reasonable attempts have been made to ensure that the data are as accurate as possible; however, there are two potential sources of error that may affect the reliability of estimates and for which no adequate adjustments can be made, known as sampling and non-sampling errors.

4 . Quality summary

Overview

The Wealth and Assets Survey (WAS) launched in 2006 and is a biennial longitudinal survey conducted by the Office for National Statistics (ONS). This survey measures the well-being of households and individuals in terms of their assets, savings, debt and planning for retirement. The survey also examines attitudes and attributes related to these. Data from this longitudinal survey will also provide users with the ability to measure changes of wealth in Great Britain over time. The survey is currently sponsored by a funding consortium, including the ONS; Department for Work and Pensions (DWP); HM Revenue and Customs (HMRC); Financial Conduct Authority (FCA); and Scottish Government (SG).

Approximately 30,000 households were interviewed in wave one; 20,000 in wave two; 21,000 in wave three; 20,000 in wave four; and 18,000 in wave five. Wave six of the WAS was only in the field for 21 months as a result of moving WAS to financial years (rounds); this is detailed in the Recent improvements subsection. The number of households sampled for wave six was therefore approximately 16,000.

As wealth is known to be unevenly distributed, addresses more likely to contain wealthier households were sampled at a higher rate to improve the efficiency of the sample. These addresses were identified utilising data from HMRC.

The datasets for waves one to five have been released to consortium members as well as the UK Data Archive (UKDA) under special user licence and to the [Approved Researcher Scheme](#) managed by the ONS. Given the need to maintain respondents' anonymity, certain variables have not been released to consortium members and the UKDA. To encourage the widest possible use of the data, including internationally, the latest End User Licence (EUL) datasets were created and deposited with the UKDA in December 2019.

The WAS provides valuable data on topics that are not sufficiently covered elsewhere.

Uses and users

The results of the WAS are used by the DWP, HMRC, ONS and other government departments (including the Department for Business, Energy and Industrial Strategy (BEIS) and Department of Health (DH)) as well as academics. The data provide a greater understanding of the levels and distribution of wealth in terms of pensions, property, financial and physical assets, and indebtedness.

Strengths and limitations

Self-valuation tends to yield higher estimates of worth than most other property indicators may suggest.

As wealth is highly skewed towards the top, the survey was designed to pick up the very wealthy. However, this means that the sample now contains some very wealthy outliers. All such cases are thoroughly checked and, as a result, they are included in the survey results. Given the skewed nature of wealth data and the effect that outliers can have on parametric estimates, the [Wealth in Great Britain statistical bulletin](#) and the associated background tables do not generally report mean values. Instead, they use the median values to report central tendency (this is not possible for physical wealth estimates because of how physical wealth data are collected).

Recent improvements

The survey period has now moved to a two-year, financial year-based periodicity (April to March), with this periodicity being referred to as a “round”. Therefore, round 6 covers the period April 2016 to March 2018. This move to a two-year, financial-year basis allows the WAS to be integrated with other household financial surveys that are based on financial years and analysed on a consistent basis alongside other components included within other household financial surveys (income and expenditure). An article detailing this, [Moving the Wealth and Assets Survey onto a financial years’ basis](#), was published in July 2019.

The latest round of the WAS will include derived variables to measure wealth on an individual basis. A paper on the methods used to [measure wealth on an individual level](#) was published in October 2018.

Some figures included in the latest release, [Wealth in Great Britain Round 6: 2016 to 2018](#), have been deflated to April 2016 to March 2018 average prices using the Consumer Prices Index including owner occupiers’ housing costs (CPIH) to reflect the change in the value of money over time. This is a divergence from the previous releases where all figures were in nominal terms. This change has been made to allow better comparisons over time.

5 . Quality characteristics of the Wealth and Assets Survey data

Relevance

A number of government departments joined the Wealth and Assets Survey (WAS) consortium because this survey was identified as being able to supply data on topics that were not sufficiently covered elsewhere. This survey fills a major information gap on wealth and indebtedness at a household and personal level. The pension wealth data in particular is unique owing to its detail. The longitudinal element provides a further dimension to this dataset, allowing users to analyse levels of change across all waves, from lower levels of wealth and indebtedness to households’ or individuals’ total wealth.

The survey has a large sample and almost complete coverage of Great Britain. The results of wave one to round six have been and are likely to be used by the Department for Work and Pensions (DWP); HM Revenue and Customs (HMRC); other government departments; analysts within the Office for National Statistics (ONS); and academics to provide a greater understanding of the levels and distribution of wealth in terms of pensions, property, financial and physical assets, and indebtedness.

Accuracy and reliability

Multiple quality assurance methods ensure that the WAS data are as reliable as possible. These methods are applied during the interview and after collection through outlier detection and comparisons of the data between waves and rounds. All data that are identified as possible errors are investigated and, where appropriate, altered.

Revisiting respondents in subsequent waves provides the opportunity to confirm some of the previous waves’ data. This is of particular importance for respondents whose previous waves’ interviews were given by proxy. We may revise published data to improve its accuracy following comparisons of data between waves.

Coherence and comparability

Major government surveys now use harmonised questions on important topics to ensure comparability of results. Where appropriate, WAS questions are harmonised with other government surveys. A list of harmonised questions is available via the [Harmonisation Programme](#).

Separate datasets for each wave are issued after all checks have been completed. Although many of the variables are comparable between waves, some datasets have changed as have some of the categories of responses for particular variables. When this occurs, details are provided in the user guides and variable lists.

A glossary of the main terms used in the WAS is provided in [Glossary: Wealth in Great Britain, 2006 to 2008](#).

There is limited comparable data from administrative sources or major surveys for some topics covered by the WAS. Nevertheless, using information that was available, the comparability of the results with results from other sources have been checked by the various contributors to the report as part of the validation process. This comparison will have included information from less extensive surveys, administrative data and the Financial Reporting Standards (FRS).

Accessibility and clarity

The UK Data Archive (UKDA) at the University of Essex provides access to approved researchers under special user licence.

Documentation to guide the users of these datasets has been provided to the consortium members and is available to UKDA-approved researchers.

Our 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. We also offer 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 our website but not produced by us, or referenced on our website but stored elsewhere, may vary. For further information, please email us at wealth.and.assets.survey@ons.gov.uk.

Timeliness and punctuality

The survey has been in existence for a short period and had a biennial interview wave pattern. The survey period has now moved to a two-year, financial year-based periodicity (April to March), with this periodicity being referred to as a “round”.

These data are available in the following releases:

- [Wealth in Great Britain, Main Results from the Wealth and Assets Survey 2006 to 2008](#)
- [Wealth in Great Britain Wave 2: 2008 to 2010](#)
- [Wealth in Great Britain Wave 3: 2010 to 2012](#)
- [Wealth in Great Britain Wave 4: 2012 to 2014](#)
- [Wealth in Great Britain Wave 5: 2014 to 2016](#)
- [Moving the Wealth and Assets Survey onto a financial years' basis](#)
- [Wealth in Great Britain Round 6: 2016 to 2018](#)

For more details on related releases, the [Release calendar](#) 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](#).

Building on the investment of work during wave 1 and in particular the wave 2 process, the ONS and the consortium are improving the efficiency of the data derivation, cleaning and validation process so that the publication of datasets for round 6 and beyond should be timelier.

Concepts and definitions (including list of changes to definitions)

The classifications used for the WAS are harmonised with other government surveys. These classifications are:

- Household Outcome Code
- UK Standard Industrial Classification (UK SIC)
- UK Standard Occupational Classification (UK SOC)
- Country of birth
- Nationality
- Religion
- Ethnicity

Output quality

There is a trade-off between accuracy and timeliness of data dissemination. In theory, the WAS data could be disseminated immediately after fieldwork completion. However, the format of the data and level of item or unit non-response would significantly reduce the analytical value and usability of the data.

We have decided to undertake editing and imputation of WAS data prior to its dissemination. This significantly improves the quality and usability of data available for analysis, but it delays the dissemination of data. We are actively working to reduce the amount of time that the edit and imputation stages take, to retain accuracy while ensuring the data are disseminated as soon after data collection as possible.

The longitudinal nature of the survey allows us to validate the responses provided in the previous interview at the current interview. For example, on occasion, information is provided by proxy at one wave and then in person at the subsequent wave. Personal responses are considered to be more accurate and therefore we take the opportunity to improve the quality of the previous wave's responses. This can mean that higher quality data are available at a later date, which can lead to the revision of previously published estimates.

Sampling error

All reasonable attempts have been made to ensure that the data are as accurate as possible. However, there are two potential sources of error that may affect the accuracy of estimates and for which no adequate adjustments can be made: sampling and non-sampling errors.

Sampling error refers to the difference between the results obtained from the sample and the results that would be obtained if the entire population was fully enumerated. The survey estimates are therefore likely to differ from the figures that would have been produced if information had been collected for all households or individuals in Great Britain. The extent to which survey estimates vary from their population values can be estimated, to a given level of confidence, through the calculation of confidence intervals via the standard error of the estimate.

The standard error is a measure of sampling variability, which shows the extent to which the estimates are expected to vary over repeated random sampling. To estimate standard errors correctly, the complexity of the survey design needs to be accounted for.

Some estimates of standard errors for main variables are available in the supporting tables, [Wealth in Great Britain R6: Quality Indicators](#). However, these standard error estimates do not account for imputation, which may affect variability.

Additional inaccuracies, which are not related to sampling variability, may occur for reasons such as errors in response and reporting. Inaccuracies of this kind are collectively referred to as non-sampling errors and may occur in a sample survey or a census. The main sources of non-sampling error are:

- response errors such as misleading questions, interviewer bias or respondent misreporting
- bias resulting from non-response, as the characteristics of non-responding persons may differ from responding persons
- data input errors or systematic mistakes in processing the data

Non-sampling errors are difficult to quantify in any collection. However, every effort was made to minimise their effect through careful design and testing of the questionnaire, training of interviewers, and extensive editing and quality control procedures at all stages of data processing. Imputation is another method used to improve accuracy resulting from missing observations in the dataset.

Response

Response rates are reported on a monthly basis and are based on the number of fully and partially co-operating households as a proportion of the numbers of eligible households in the sample. A response rate of 55% was achieved for wave 1, and 68% of the eligible households' sampled responded in wave 2. For wave 3 onwards, the response rates for the new and old cohorts are included in Table 1.

Regional response rates for WAS have not varied a great deal although London, in common with other social surveys, tends to exhibit markedly lower response rates.

Table 1: New and old cohort response rates, wave 3 to round 6

	New cohort	Old cohort
Wave 3	51%	73%
Wave 4	53%	70%
Wave 5	55%	69%
Round 6	46%	71%

Source: Office for National Statistics – Wealth and Assets Survey

6 . Methods used to produce the Wealth and Assets Survey data

How the output is created

This longitudinal Wealth and Assets Survey (WAS) measures the numbers and values of assets, debt and savings as well as attitudes to savings and indebtedness. Classificatory variables, including age, sex and employment status, are also covered. The first wave of the survey commenced with interviews carried out over two years from July 2006 to June 2008. For subsequent waves:

- a second wave took place two years on from initial interviews, covering the period July 2008 to June 2010
- a third wave began in July 2010, which was completed in June 2012
- a fourth wave of WAS commenced in July 2012 and was completed in June 2014
- a fifth wave of WAS commenced in July 2014 and was completed in June 2016
- wave six of WAS commenced in July 2016 but ran for only 21 months to March 2018; this was a result of moving WAS to a financial year and round basis

Information on the creation of rounds 5 6 WAS data can be found in [Moving the Wealth and Assets Survey onto a financial years' basis](#).

The WAS is a continuous survey with interviews spread evenly over the year, which helps to ensure that estimates are not biased by seasonal variations.

The survey samples private households in Great Britain, excluding north of the Caledonian Canal, the Scottish Islands and the Isles of Scilly.

The stratification of the sample for the first wave of WAS was based on regional and the 2001 Census variables and had two stages. At the first stage, a stratified sample of primary sampling units (PSUs) was drawn from a list of postcode sectors included in the small users' Postcode Address File (PAF). This list was sorted by geography (region by Metropolitan status), the proportion of households with the household reference person (HRP) in National Statistics Socio-economic classification (NS-SEC) group one to three, and the proportion of households without a car. This stratification is using judgement based on experience from optimising stratifiers for the Family Resources Survey (FRS) and Living Costs and Food Survey (LCF). The NS-SEC stratifier is the most powerful stratifier for economic social surveys. Car ownership was chosen over economic activity for the second stratifier because this is more correlated with wealth.

The second stage involved selecting 26 addresses per PSU being using systematic random sampling from the small users' PAF. The list of addresses in each PSU was sorted by postcode and street number. The sampling was carried out in such a way that the addresses flagged as expected to feature wealthier households had two and a half to three times the probability of being sampled as non-flagged addresses.

For the first two years of the first wave of the survey, 1,200 PSUs were drawn, giving a set sample of 31,200 addresses per year.

For the second and subsequent waves, all households that responded in the first wave and all households that could not be contacted in the previous wave were revisited. To ensure respondents' contact details are maintained between waves, a "keep in touch" phone call is administered approximately four months prior to the respondent's next interview. This exercise can document households that have split and those that have moved.

As the sample for each subsequent wave consists predominantly of the preceding wave's respondents, the size of the sample reduces with each wave. To mitigate the effect of attrition, a new cohort was introduced in wave three (8,000 new addresses in year one and 4,000 in year two). A further cohort of 8,000 addresses was introduced in wave four, 6,000 in wave five and a further 9,000 in round six. The new cohort improves the size of the cross-sectional sample, which is required because attrition has reduced the sample since wave one. The new cohort then may help to reduce any bias introduced by attrition as the new cohort is selected from the current population and so helps in accounting for changes in the characteristics of population over time.

Editing

An extensive range of validation checks and computer edits were applied to both the household and individual questionnaires during the computer-assisted personal interview (CAPI) and to the aggregate data file in the office.

Imputation

Imputation is an adjustment process that is used to determine and assign replacement values to resolve problems of missing, invalid or inconsistent data.

The problem of missing data in the WAS is approached in two stages: first, a deductive imputation method, followed by a statistical method. Deductive imputation was applied where a missing or inconsistent value could be deduced with certainty. Secondly, statistical imputation was carried out using a nearest-neighbour imputation method where information from a donor record that had no errors or missing values was used to replace the missing values for a recipient record. In this approach, a donor is selected from a pool of potential donors with similar characteristics based on conditional probabilities.

For longitudinal households, where an observed value is present in one wave but the other wave is missing and therefore requires imputation, an imputed value is drawn from a donor with reference to the observed value or is calculated based on observed relationships or ratios between variables in the donor record. The imputation is conducted under edit constraints to ensure that outliers and implausible relationships are not introduced into the data through the imputation process.

Outliers

As part of the data cleaning process, cross-sectional outliers were identified on all monetary variables used to compile derived variables. Large changes between waves were also identified as longitudinal outliers. Outlier thresholds were determined through analysis of the distribution of the data. Each variable was analysed and dependent upon the nature of that variable and the spread of its data, and a percentage of the highest and lowest values were identified as outliers. Not all variables had their lowest values labelled as outliers as low values can be acceptable for some variables, for example, zero values in financial accounts.

Outliers were checked for credibility through examination of other variables, including the previous waves' responses, in an attempt to find evidence to support or inform an edit to the outlier. This evidence includes the inspection of wealth, through income, assets and debts, and verification from linked variables (comparisons of mortgage value with monthly mortgage payment and remaining term). There are reasons to justify substantial longitudinal changes. Alterations to working status or household structure, in particular a split in partnership or a house move, can significantly affect the longitudinal change of many variables.

Amendments were only made to data where sufficient evidence to support an amendment existed. In waves one and two, approximately 5% of the data were investigated as outliers, of which a minority of these were amended. A more systematic approach was established for the identification of cross-sectional and longitudinal outliers from wave three onwards.

Weighting

A three-stage weighting procedure was implemented in the WAS. First, a design weight, equal to the reciprocal of the address selection probability, was constructed. Secondly, a non-response weight was created to reduce potential non-response bias. The non-response model currently includes region (GOR), a socio-economic indicator (OAC) and the HRMC-provided wealth index used to identify the wealthiest households. This applies to a new cohort. In older panels, an attrition adjustment was applied and joiners were incorporated, before calibration.

The final stage of the weighting procedure calibrated the product of the design and non-response weights to known population totals taken from official population estimates present at the time of the fieldwork period. Different sets of weights have been created so that analysis can be performed both longitudinally (person-level) and cross-sectionally (household-level and person-level) on the data.

Information on the creation of the round based weights has been included in the article, [Moving the Wealth and Assets Survey onto a financial years' basis](#).

7 . Other information

More information on the methods used to compile the output can be found in [Chapter 10, Technical details, in the Wealth in Great Britain, Main Results from the Wealth and Assets Survey, 2006 to 2008 \(PDF, 820KB\)](#) and [Chapter 4 in the Wealth in Great Britain Wave 2, 2008 to 2010 \(Part 1\)](#).

Statistical disclosure control methodology is applied to Wealth and Assets Survey (WAS) data. 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](#) and specifically the Principle on Confidentiality set out practices for how we protect data from being disclosed. The Principle includes the statement that Office for National Statistics (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”.

More information can be found in the [National Statistician's guidance on the confidentiality of official statistics](#) and on the [disclosure control policy for social survey microdata](#) page.