



# Wealth and Assets Survey User Guide Round 8



<b>OVERVIEW .....</b>	<b>5</b>
<b>USING WEALTH AND ASSETS SURVEY DATA.....</b>	<b>6</b>
Content of datafiles .....	6
Variable naming conventions .....	6
Weights .....	7
Interview Outcome codes.....	8
Longitudinal data linkage.....	8
Linking within a wave.....	9
Linking End User License (EUL) data .....	9
Longitudinal Flags.....	10
Using the “Round” Datasets .....	11
Individual level wealth .....	11
Published Totals.....	11
<b>SURVEY DESIGN .....</b>	<b>11</b>
Sampling strategy.....	11
Sample sizes of each wave .....	12
Wave structure .....	13
Mode of data collection.....	15
<b>FIELDWORK PROCEDURES .....</b>	<b>15</b>
Interviewer training .....	16
Respondent contact.....	17
Field sampling procedures .....	17
<b>RESPONSE RATES.....</b>	<b>18</b>
Keeping in Touch .....	21
<b>QUESTIONNAIRE CONTENT .....</b>	<b>21</b>
Overview .....	21
Questionnaire changes .....	22
Programming and testing .....	22
<b>OCCUPATIONAL CODING UPDATE.....</b>	<b>23</b>
<b>EDITING .....</b>	<b>24</b>
<b>IMPUTATION .....</b>	<b>25</b>
General Methodology.....	25
Donor Selection.....	26
Quality Assurance and Evaluation .....	29
Percentage of Data Imputed for Round 8.....	29
<b>WEIGHTING.....</b>	<b>30</b>
Overview .....	30
Different Types of Weights .....	30
Longitudinal R7-R8 Weights.....	31
R2-R8 survivor weights .....	32

Cross-sectional weights .....	33
Pre-calibration weights for returning panels (W1, W3, W4, W5 Q1-7, R6, R7) .....	33
Pre-calibration weights for the R8 panel .....	34
Household-level cross-sectional weights.....	35
Descriptive statistics for R8 weights .....	35
Cross-sectional weights for WAS W1-R7 .....	36
Data Quality .....	37
Sampling error .....	37
Non-sampling error.....	37
External source validation .....	38
<b>WEALTH ESTIMATES.....</b>	<b>39</b>
<b>PRIVATE PENSION WEALTH MEASURES.....</b>	<b>39</b>
<b>CONTACT DETAILS.....</b>	<b>44</b>
<b>ANNEX: IMPACT OF COVID .....</b>	<b>45</b>
Background .....	45
Incremental Changes .....	45
Initial Approach.....	45
Questionnaire Changes.....	45
Sample Design.....	47
New Methods of Making Contact.....	47
Impact Upon Analysis .....	48

## Overview

The Wealth and Assets Survey (WAS) is a longitudinal survey that interviews across Great Britain i.e. England, Wales and Scotland (excluding north of the Caledonian Canal and the Isles of Scilly).

Respondents to Wave 1 (July 2006 – June 2008) of the survey were invited to take part in a Wave 2 follow up interview two years later (July 2008 – June 2010). Respondents to Wave 2 were then invited to take part in a Wave 3 follow up interview a further two years later (July 2010 – June 2012). In addition to these re-curing follow-up interviews, a new random sample of addresses was also added for Wave 3 to Round 8.

The approximate number of achieved households in each wave/round was:

- Wave 1 achieved approximately 30,000 household interviews
- Wave 2 achieved approximately 20,000 household interviews
- Wave 3 achieved approximately 21,000 household interviews
- Wave 4 achieved approximately 20,000 household interviews
- Wave 5 achieved approximately 18,000 household interviews
- Round 6 achieved approximately 18,000 household interviews
- Round 7 achieved approximately 17,500 household interviews
- Round 8 achieved approximately 15,100 household interviews.

This periodicity, that started in July and ended in June two years later is referred to as a “Wave”, and was maintained until Wave 5, which covered the period July 2014 to June 2016.

The survey then moved to a two-year, financial year-based periodicity (April to March), referred to as a “Round”

- Round 6 covered the period April 2016 to March 2018
- Round 7 April 2018 to March 2020
- Round 8 April 2020 to March 2022.

This move allowed WAS to be integrated and analysed with other household financial surveys that report on financial years.

The economic well-being of households is often measured by its income, despite a household's resources being composed of its stock of wealth as well as its flow of income. To fully understand the economic well-being of households it is necessary to look beyond measures of household income.

The WAS addresses this gap by gathering information on the ownership of assets (financial, physical and property), pensions, savings and debt.

Round 8 of WAS was funded by a consortium of government departments:

- Department for Work and Pensions
- HM Revenues and Customs
- Scottish Government
- Office for National Statistics.

Fieldwork is undertaken by the Office for National Statistics.

Interviews in all waves were conducted using Computer Assisted Interviewing (CAI).

# Using Wealth and Assets Survey Data

## ***Content of datafiles***

The data are split into two linked files:

- a household level file containing all property and physical wealth component variables, as well as all derived variables (DV) used for the calculation of aggregated household wealth and income
- a person level file consisting of all person level financial wealth, pension wealth and income component variables and DVs.

## ***Variable naming conventions***

### *Variable categories*

Naming conventions are applied to aid identification of the theme related to a variable.

Variables belonging to the same category often start with the character that relates to the category. For example, pension variables start with the letter 'P' e.g. 'PocTyp1R8'. Most of the variable names are linked to the description of that variable e.g. 'LnLft1' – Number of monthly installments left to pay loan one.

The majority of derived variables names have the prefix 'DV'.

### *Round suffix*

Most variables in both datasets are given the suffix 'R8', to indicate they contain values collected in Round 8

### *Wave suffix*

A smaller number of variables are given the suffix 'R7', 'R6', 'W5', 'W4', 'W3', 'W2' or 'W1' to indicate they contain values from Rounds 6 or 7 or Waves 1 to 5. Most of the variables with a previous wave suffix are present to allow the datasets to be matched to those from previous Rounds and Waves of the survey, e.g. HHSERIALW1.

### *Imputation suffix*

All variables used as components for wealth DVs were subject to imputation. Variables that have had missing data imputed appear in the datasets in two versions. The version that contains only the values observed at interview will end with the suffix R8 as described above, e.g. FSInValR8. The version that contains both observed and imputed values will end with the suffix '\_i', e.g. FSInValR8\_i.

### *Aggregation suffix*

To calculate total household wealth all component DVs were aggregated to household level. To enable data users to use aggregated household level DVs on person level data, relevant DVs are also provided on the person level file, e.g. DVFBondVR8\_aggr at household level and DVFBondVR8 at person level.

### *Binary suffix*

A number of questions invite respondents to choose all answer options that apply. How this data is stored within the questionnaire will be affected by the order in which the options are selected and their number. To aid statistical imputation and analysis, responses to these questions were translated to series of binary variables for each option with "binary" included as part of the variable name e.g. responses to the SocFund question were saved to:

- socfund\_binary1r8: Received a grant for funeral expenses
- socfund\_binary2r8: Received a grant for maternity expenses

- socfund\_binary3r8: Receives student uniform grant
- socfund\_binary4r8: Received a budgeting loan or budgeting advance from DWP
- socfund\_binary5r8: Received an LA Loan or a grant

from Round 8.

Until Round 7 responses to these questions would have been recorded as collected with each of the 5 variables recording responses permitted to hold any of the answer options to the question. How the data for these questions are submitted for analysis needed to be modified from Round 8.

## Weights

To carry out cross-sectional analysis based on the individual wave data, the following table has the appropriate variable weight to apply for cross-sectional analysis. For further information on how weights are created, please see the following [Weighting](#) section of this guide.

**Table 1 – Cross-sectional weights used across each wave of the WAS**

Wave/Round	Cross-sectional Calibration Weight
1	XS_wgtW1
2	XS_calwgtW2
3	W3xswgt
4	W4xshhwgt
5	W5xshhswgt
6	R6xshhswgt
7	R7xshhwgt
8	R8xshhwgt

Some questions are asked of personal respondents only i.e. proxy responses are not permitted. To permit analysis of these questions cross-sectional non-proxy weights have been provided on the person level file since Wave 4.

**Table 2 – Cross-sectional non-proxy weights used across each wave of the WAS**

Wave/Round	Cross-sectional Non-Proxy Calibration Weight
4	W4_nonproxy_wgt
5	W5_nonproxy_wgt
6	R6_nonproxy_wgt
7	R7_nonproxy_wgt
8	R8xs_nonproxy_wgt

In contrast to cross-sectional analysis, longitudinal analysis can only be carried out on person level files. The following table presents the longitudinal variable weights to apply for longitudinal analysis.

**Table 3 – Longitudinal calibrations weights used at each wave of the WAS**

Wave/Round	Longitudinal Calibration Weight
W1W2 longitudinal weight	Longit_calwgtW2
W1W3 longitudinal weight	w1w3wgt
W2W3 longitudinal weight	w2w3wgt
W1W4 longitudinal weight	W1W4_longwgt
W3W4 longitudinal weight	W3W4_longwgt
W4-W5 longitudinal weight	W4W5_longwgt
W1-W5 longitudinal weight	w1w5wgt

R5-R6 longitudinal weight	R5R6LongWgt
R6-R7 longitudinal weight	R6R7LongWgt
R7-R8 longitudinal weight	R7R8LongWgt
R2-R5 survivor weight	SurvWgtR2R5
R2-R6 survivor weight	SurvWgtR2R6
R2-R7 survivor weight	SurvWgtR2R7
R2-R8 survivor weight	SurvWgtR2R8

### **Interview Outcome codes**

The datasets include responding households only. The variable HOutR8 gives an outcome code which provides an indication of the type of interview outcome of the household.

**Table 4 – Interview outcome codes for responding households in Round 8**

<b>Response Status</b>	<b>HOut Code</b>	<b>Code Description</b>
Fully co-operating	110	Complete interview by required respondent(s) in person
	120	Fully co-operating household, one or more interviews completed by proxy.
	121	HRP economic unit interviewed in person, one or more other interviews by proxy
	122	HRP and/ or spouse/ partner interview by proxy
Partially co-operating	211	Full response in person from HRP economic unit. One or more other interviews missing or complete
	212	Non-contact with one or more respondents
	213	Refusal by one or more respondents
	220	HRP economic unit not complete (one of two eligible adults missing; either interview incomplete)
	230	No individual interviews with HRP economic unit, but household interview completed

Although the dataset exclusively consists of responding households, not every individual in every household responds. The variable IOut1R8 indicates the interview outcome of individuals.

**Table 5 – Individual response outcome codes for Round 8**

<b>IOutR8</b>	<b>Response Status</b>
1	Full interview (in person or by proxy)
2	Partial interview (in person or by proxy)
3	Ineligible for interview – child aged 0 to 15
4	Ineligible for interview – adult aged 16 to 18 in full-time education
5	Eligible adult – refused to be interviewed
6	Eligible adult – non-contact

Please note: Although individuals with an outcome code of 5 or 6 did not give an interview they can still be included in the analysis because their values for wealth component variables have been imputed.

Also, analysts should be aware that although children have not been interviewed for this survey, the data on children's assets has been recorded against their person number in the household, not against the adult who responded to the relevant questions in this section.

### **Longitudinal data linkage**

As a longitudinal survey some users may wish to link data for individual respondents across Waves or Rounds. Due to the change from 'Wave' to 'Round' datasets, longitudinal data linkage can only be fully done between Rounds 5 to 8.

For longitudinal linkage between previous Waves 1-5, please see the [Wave 5 User guide](#).

To permit linkage all files include a single variable for linking cases. For household linking, there are separate variables for each wave; each case may have up to three variables with a valid code. For person level there is one variable used for matching a case in any wave.

For Secure Research Service (SRS)/Government License (GL) Datasets:

- always used the linked file as a base when matching variables across Waves
- use HHSerialR7/R8 for household linking
- use PIDNO for person level linking, this remains the same over the survey lifetime of a sample unit.

*To add R7 variables to the R8 person file, keeping only R8 cases:*

- sort R7R8 person level linked files by PIDNO
- sort R7 Person file by PIDNO
- sort R8 Person file by PIDNO
- match R7 and R8 files use PIDNO to match
- select the required cases e.g. for R8 cases (including linked R7 cases) use HHSerialR8 > 0

*To add R7 variables to the R8 household file, keeping all cases*

- sort the linked file by HHSerialR8
- sort R8 household file by HHSerialR8
- match files using HHSerialR8
- sort the new file by HHSerialR7
- sort R7 household file by HHSerialR7
- match files using HHSerialR7
- this will produce a linked R7R8 file with R7 and R8 variables.

### ***Linking within a wave***

To add household variables to the R8 person file, sort both files by HHSerialR8 and use this variable to match

Please Note: Person level variables cannot be directly added to the household file unless they are first aggregated to the household level to permit a 1 to 1 match to be made.

### ***Linking End User License (EUL) data***

As the EUL datasets are treated to minimize the risk of disclosure, the variables HHSerial and PIDNO have been anonymised. For household files, the variable CASE should be used when linking, in place of HHSerial. For person files, the variable Person should be used to link in place of PIDNO.

*Linking using EUL datasets:*

- use CaseR7-R8 for household linking
- use PersonR7-R8 and CaseR7-R8 for person level linking

*To add R7 variables to the R8 person file, keeping only R8 cases sort R7 by CaseR7 and PersonR7*

- sort R8 by CaseR7 and PersonR7
- match R7 and R8; using PersonR7 and CaseR7 to match
- select the required cases e.g. for R8 cases (including linked R7 cases) use CaseR8 > 0

*To add R7 variables to the R8 household file, keeping all cases sort R7 household file by CaseR7*

- sort R8 household file by CaseR7
- match files using CaseR7

## **Longitudinal Flags**

A number of longitudinal flags have been produced that may help to understand changes in the data when conducting longitudinal analysis with the linked data.

The following person level flags are only included on the person level datasets.

**Table 6 – Indicator for linkage status, shown through variable ‘TypeR8’**

<b>TypeR8</b>	<b>Variable Label</b>	<b>Linkage Status</b>
1	W5 and R8 linked	Regardless of interview eligibility and response status.
2	W7 and R8 linked	Regardless of interview eligibility and response status.
3	R8 HAK Joiner (no HAK joiners for R8)	Individual joined the household when keep-in-touch exercise was conducted.
4	R8 re-entrant – linked to earlier wave	Individual joined survey prior to R7, did not respond at R7, but re-entered survey at R8.
5	R8 New respondent	Individual joined the household when R8 interview was conducted.
6	R8 New household	Individual is part of a household that responded at R8 for the first time.
7	Individual not present at R8	This person was part of a responding household in R7 but left the household at R8 and did not respond.
8	Household not present at R8	Individual was part of a responding household in R7 but the whole household did not respond at R8.

**Table 7 – Sample member status at R8, shown through variable ‘P\_flag1R8’**

<b>P_Flag1R8</b>	<b>Sample Member Status</b>	<b>Description</b>
1	Longitudinal original sample member (LSM)	Individual was a member of a responding household in previous wave and current wave.
2	Entrant original sample member (ESM)	Individual was a member of a responding household in current wave, but household did not respond in previous wave.
3	Secondary sample member (SSM)	Individual was not a member of any household in previous wave but joined a longitudinal household in current wave.
4	Non-responding sample member (NSM)	Individual was a member of a responding household in previous wave but left the sample at current wave.

**Table 8 – Wave entrant status at Round 8, shown through variable ‘P\_flag2R8’**

<b>P_Flag2R8</b>	<b>Wave Entrant Status</b>	<b>Description</b>
1	Original Sample Member (OSM) birth entrant	Child entrant (15years or younger) born to OSM household member.
2	SSM birth entrant	Child entrant (15years or younger) born to SSM household member.
3	Other SSM entrant	Adult entrant (16years or older).

**Table 9 – Wave eligibility status at Round 8, shown through variable ‘P\_flag3R8’**

<b>P_Flag3R8</b>	<b>Wave Eligibility Status</b>	<b>Description</b>
------------------	--------------------------------	--------------------

1	Eligible adult	Aged 16 years or older and not in full-time education.
2	Ineligible adult	Aged 16 to 18 years in full-time education.
3	Ineligible child	Aged 15 years or younger.

**Table 10 – Household representative Status at Round 8, shown through variable ‘P\_Flag4R8’**

P_Flag4R8	Household Representative (HRP) Status
1	HRP in both previous and current wave.
2	HRP in previous wave, but not in current wave.
3	HRP in current wave, but not previous wave.
4	Never a HRP.

## Using the “Round” Datasets

The periodicity of the data Wealth and Assets Survey (WAS) was changed from a two-year July to June periodicity to a two-year April to March periodicity during the course of Round 6. Information regarding this shift and how to use the Round-based datasets in conjunction with the Wave-based datasets can be found in the [Round 6 User Guide](#). The Round 7 & Round 8 datasets are the first to comprise a full Round-based 24-month data collection period, spanning the periods April 2018 to March 2020 & April 2020 to March 2022.

## Individual level wealth

Individual level wealth variables are available on the person files from W3 onwards. Information on the person level wealth derived variables including the variable names and how they are derived can be found in the ‘Individual level wealth derived variable specification’. The methodology used to create the measure of individual wealth can be found within the document “[Measuring wealth on an individual level](#)”.

ONS use the household weights (see the ‘[weighting](#)’ section for further information on the weight variables) when using the person level wealth variables. This is to ensure that aggregate estimates of wealth produced using the individual level wealth variables are consistent with aggregate estimates of wealth produced using the household level wealth variables.

## Published Totals

You may find that the EUL dataset totals don’t match the totals published by ONS. This is due to treatments applied to the dataset to permit the release of microdata that complies with disclosure best practice to maintain the anonymity of all respondents. The ONS Secure Research Service (SRS) contains a version of the data unaffected by these treatments. The [linked pages](#) provide information on how you can apply to access secure data from home. More information on Assured Organisational Connectivity can be obtained by emailing [SRS.Connectivity@ons.gov.uk](mailto:SRS.Connectivity@ons.gov.uk).

## Survey Design

### Sampling strategy

The Wealth and Assets Survey collects information about private household wealth in Great Britain. The survey uses the small users Postcode Address File (PAF) as the sample frame for residential addresses in Great Britain, that is, England, Wales and Scotland, excluding the area north of the Caledonian Canal and the Isles of Scilly. The ONS copy of the PAF is updated twice a year to ensure that recently built addresses are included and demolished or derelict properties are removed quickly.

The survey estimates are designed to be representative of the GB population, therefore the WAS,

like most social surveys, uses a 'probability proportional to size' or PPS method of sampling cases. This means that the probability of an address being selected is proportional to the number of addresses within a given geographic area. This results in a higher number of addresses being selected from densely populated areas.

WAS uses a two-stage or 'clustered' approach to sampling. Firstly, postcode sectors are randomly selected from the PAF. The postcode sectors are the primary sampling units (PSUs) for the survey. Within each of these postcode sectors, 26 addresses are randomly selected. The selection uses a stratified (ordered) PAF, where addresses are listed by postcode and street number. The list of 26 addresses is split into two quotas of 13 addresses to ease the allocation (to interviewers) and management of fieldwork.

In response to the pandemic quota sizes were adjusted across the Round, first doubling to 26 addresses from June 2020, before being gradually reduced across the remainder of the Round; detail on the quota size changes made across the whole of the Round are presented in the Annex.

The sampled PSUs were allocated to months at random. This was done using a repeating random permutation which ensured that PSUs allocated to the same quarter and month were evenly spread across the original sample, while still ensuring that each sampled PSU had an equal chance of being allocated to each month. This even spread meant that monthly and, particularly, quarterly samples were balanced with respect to the regional and census-based variables used in the stratification. Adjustments were, however, made to this initial allocation to months, to more effectively match this random allocation across months to field capacity.

Although the address selection within postcode sectors is random, some addresses have a higher probability of selection than others. Over sampling of wealthy addresses occurs because wealth has a heavily skewed distribution, with a relatively small number of addresses holding considerable wealth. Wealthy households are often harder to secure responses from, so oversampling is necessary. For year 1 of Wave 1, addresses identified as having high wealth were 2.5 times more likely to be sampled than other addresses. This factor was increased to 3.0 for the second half of Wave 1 in order to further increase the number of achieved interviews with high wealth addresses.

'High' wealth addresses are identified after the postcode sectors have been established. A limited amount of information is available about the type of household resident at a particular address on the PAF and what is generally available relates to the area around the address, rather than being specific to an address. However, HMRC collects data on income and certain components of wealth in order to administer the tax system and the Self-Assessment regime. Data from HMRC on tax returns at an address level, in conjunction with average FTSE250 dividend yields from the previous calendar year, are used to estimate the value of shareholdings at a household level.

### **Sample sizes of each wave**

The following table provides a summary of the sample sizes (rounded), issued and achieved, for each of the Waves/Rounds of the Wealth and Assets Survey.

**Table 11 – Summary of sample sizes in all Waves/Rounds of the WAS.**

Wave/Round	Issued addresses	Achieved households	Achieved adults <sup>1</sup>
1	62,800	30,500	53,300
2 <sup>2</sup>	32,200	20,000	34,500
3	37,900	21,300	40,400
4	35,300	20,100	38,300
5	32,700	18,400	35,600
6 <sup>3</sup>	32,000	18,000	34,000

7	33,800	17,500	38,900
8	36,500	15,100	32,300

1. Respondents aged 16 and over.
2. For Wave 2, the achieved Wave 1 sample was issued, plus all of the non-contacts.
3. Due to the change from a July to an April start of fieldwork, the Round 6 response counts include the last 3 months of Wave 5 (April, May and June 2016).

In developing the survey, precision targets for change on key estimates were agreed in consultation with funding departments. From these, it was estimated that an overall achieved sample of approximately 32,000 households, spread evenly over the two years of Wave 1 was required. In addition to the above precision targets there was a further target to achieve a two-year sample of 4,500 households above the top wealth decile for Wave 1. This was well above the 3,200 households that would be above the top wealth decile for an equal probability sample. Oversampling the wealthiest households allows for more detailed analysis of this group and gives more precise estimates of the levels of wealth across the whole population.

A total of 32,200 addresses were issued for Wave 2.

In Wave 3, follow-up of the respondents and non-contacts at Wave 1 and Wave 2 was supplemented by the introduction of a new random sample of around 12,000 addresses.

This pattern of supplementing the longitudinal sample with a new random sample continued in each subsequent Wave and Round. For:

- Wave 4, around 8,300 addresses were added
- Wave 5, 6,000 addresses
- Round 6, 9,000 addresses
- Round 7, 13,000 addresses
- Round 8 year 1, 16,000 addresses; following a reduction in funding the decision was taken to temporarily suspend the addition of new sample for year 2.

**Wave structure**

The following diagram illustrates the longitudinal design of the Wealth and Assets Survey. Wave 1 started in July 2006 with fieldwork being spread over a two-year period. Wave 2, a follow up to Wave 1, was conducted between July 2008 and June 2010. The introduction of a new cohort of addresses in Wave 3 and subsequent Waves/Rounds is shown in blue.

All interviews have a two-yearly interval between waves, therefore providing estimates of change in relation to the same period of time. For example, Wave 1 interviews conducted during July 2006 would be repeated for Wave 2 in July 2008. It is important that this gap remains constant so that estimates of change are comparable wave on wave.

**Table 12 – Longitudinal survey design of the WAS – Wave Structure**

	Jul-06 - Jun-08	Jul-08 - Jun-10	Jul-10 - Jun-12	Jul-12 - Jun-14	Jul-14 - Jun-16	Jul-16 - Mar-18	Apr-18 - Mar-20	Apr-20 - Mar-22
Wave 1	Yr 1 Yr 2							
Wave 2		Yr 1 Yr 2						
Wave 3			Yr 1 Yr 2					
Wave 3 new cohort			Yr 1 Yr 2					
Wave 4				Yr 1 Yr 2				
Wave 4 new cohort				Yr 1 Yr 2				
Wave 5					Yr 1 Yr 2			
Wave 5 new cohort					Yr 1 Yr 2			
Round 6						Yr 1 Yr 2		
Round 6 new cohort						Yr 1 Yr 2		
Round 7							Yr 1 Yr 2	
Round 7 new cohort							Yr 1 Yr 2	
Round 8								Yr 1 Yr 2
Round 8 new cohort								Yr 1

Due to the survey periodicity moving from “Waves” (July, ending in June two years later) to “Rounds” (April, ending in March two years later), interviews using the ‘Wave 6’ questionnaire started in July 2016 and were conducted for 21 months, finishing in March 2018. From Round 7, the fieldwork returned to a full 2-year, 24-month data collection period.

Detailed information on moving the Wealth and Assets Survey onto a financial year basis was published on the [ONS website](#) in July 2019.

Following a reduction in funding for the second year of Round 8 the decision was taken to temporarily suspend the addition of new cohort from April 2021 to March 2022.

## ***Mode of data collection***

The Wealth and Assets Survey has two interview stages. The primary interview is where the WAS questionnaire is utilised; this is referred to as the ‘mainstage’ interview. The second is the Keeping in Touch Exercise (KITE) interview used to maintain respondent’s contact details. The mainstage interview data is used for analysis, whereas the KITE is important for management of the respondents’ contact information and keeping people engaged with the survey.

### ***Mainstage Interview***

Until Round 7 WAS mainstage interviews were conducted using Computer Assisted Personal Interviewing (CAPI). Face to face interviewing was the preferred choice due to the complex subject matter of the survey and the need for the interviewer to support the respondent in answering the questions. Interviewer-respondent interaction is much greater on a face-to-face survey compared with other modes such as on-line, paper and telephone. Another reason for face-to-face interviewing was the need to interview everyone aged 16 and over in the household. This is more challenging with some alternative modes of data collection.

The interview length for the WAS questionnaire provides further reason for adopting the CAPI approach. Face to face contact with respondents allowed interviewers to identify when respondents were becoming fatigued during interviews. This allowed interviewers to suggest a break from the interview, or perhaps for them to continue the interview at another time in some cases. Identifying respondent fatigue, by picking up on body language, is best done when the interview is face to face. CAPI was also considered the best approach to maximise cooperation with the survey. Response rates of face-to-face surveys tend to be higher than telephone, paper and web alternatives.

In response to the pandemic, and the lockdowns that followed, the survey switched to Computer Assisted Telephone Interviewing (CATI) as it was no longer possible to complete interviews face-to-face. For the whole of Round 8 all WAS mainstage interviews were completed by telephone. The Annex to this guide provides more detail on this and other changes prompted by the pandemic.

### ***Keep in Touch Exercise Interview***

Conversely, the KITE interview aims to collect much less information, and only from one person in each household. The questionnaire is set up to establish whether the household circumstances have changed. In the vast majority of cases there is no change to the household’s address or composition, so the interview is very short (about five minutes). The requirements of KITE are much simpler than the mainstage interview, therefore in order to reduce costs and maximise value for money, the interviews are conducted using CATI.

## **Fieldwork procedures**

The following provides a summary of:

- interviewer training prior to starting work on this survey
- how progress is monitored

- how performance is assessed during data collection
- how contact is maintained with respondents between wave.

In response to the pandemic, and the lockdowns that followed, that coincided with the start of Round 8, field work procedures were updated to ensure contact with respondents remained possible.

## ***Interviewer training***

Until early 2020, and the start of the pandemic, interviewers working on the Wealth and Assets Survey received both generic field interviewer and survey specific training.

### ***Generic interviewer training***

Until then new interviewers to ONS were placed on a six-week training programme – the Interviewer Learning Programme (ILP) - where they were equipped with the skills required for social survey interviewing. The programme coordinated the activities of managers, trainers and interviewers into a structured programme that ensured all interviewers could meet the high standards expected of an ONS interviewer. The training adopted a blended learning approach. Methods used included classroom training, instructional and activity-based workbooks, and activity-based applications that tested the interviewers' skills and knowledge base. At the end of the six weeks, interviewers continued to be supported in their personal development. This was done with the assistance of their Interviewer Manager. They were also assigned an experienced interviewer as a mentor. New interviewers shadowed mentors.

Interviewers also participated in specific training events such as Achieving Cooperation Training (known as ACT). This training package had been rolled out to the entire field force. It was managed through training days at ONS HQ and interviewer support group meetings. Quarterly meetings of Interviewer Managers and their teams were held throughout the year where training issues and refresher training was regularly addressed.

Telephone interviewers and ONS help desk operatives received equivalent training and would very often convert refusals; following the receipt of an advance letter.

### ***Survey specific training***

#### ***Telephone interviewers***

ONS telephone interviewers working on the Wealth and Assets Survey received an annual briefing on how to administer the Keep in Touch Exercise (KITE) questionnaire. This briefing, delivered by telephone operations management, covered all aspects of the KITE interview including recording changes of address, changes in household composition and updating contact details. KITE interviewers were trained to try and turn around refusals, should panel respondents' express concerns over future involvement in the survey.

#### ***Face-to-face interviewers***

Interviewers working on the Wealth and Assets Survey underwent training in two stages prior to starting any WAS interviews. Firstly, they were provided with a home-study pack to work through which provided detailed information on the purpose and design of the survey as well as questionnaire content. They were also sent training cases to complete to set scenarios outlined in the Home Study pack.

Prior to the Covid-19 Pandemic, interviewers attended a face-to-face briefing of up to 12 interviewers.

This briefing session was tailored to provide an overview of the survey, highlighting the importance of administering the filter and seal screens. The briefing was led by one or two Interviewer Managers, sometimes with support from an experienced interviewer/mentor.

The interviewers completed a post-briefing online questionnaire, via Survey Monkey, testing their knowledge of the survey. Any concerns were passed onto the interviewer's manager.

Interviewers would not start WAS work until their Interviewer Manager was assured that they were fully briefed and ready to undertake the survey. All new trainees were validated through a mock interview with an experienced member of the field management team which was treated as if a live interview. On pass this validation interviewers were allocated live work on the survey.

In the response to the pandemic and the switch from face to face to telephone interviewing additional generic and survey specific training and briefings were remotely provided to interviewers. More detail on the changes to data collection prompted by the pandemic can be found in the Annex to this guide the linked publication [Impact of COVID-19 on ONS social survey data collection methodology](#).

## ***Respondent contact***

Once the sample has been selected, either from the small users Postcode Address File (new cohort), or by maintaining panel address details (old cohort), advance letters are issued to sampled households/respondents. Advance letters are issued approximately ten days prior to the start of the monthly fieldwork period. The advance letters are intended to inform eligible respondents that they have been selected for an interview; provide information on the purpose of the interview; explain the importance of respondent's participation; and to provide contact details in case eligible respondents want to find out more.

New cohort households are issued one advance letter addressed 'Dear resident' which assumes no prior knowledge or involvement in the survey. For the old cohort, each eligible respondent is sent an advance letter, addressed specifically to them, thanking for their help in the previous interview and inviting them to take part again. The exception to this is the old cohort where the respondent was a proxy interview in the previous wave – these respondents are sent a named advance letter, but the letter assumes no prior knowledge or participation in the survey.

ONS recognises that some sectors of the community can be difficult to contact. These include (but are not limited to) metropolitan areas, flats, London, ethnic minorities and gated estates. ONS provides interviewer guidance on calling patterns designed to maximise contact as efficiently as possible. This strategy was underpinned by a Calling Pattern Checklist.

The calling strategy which achieves the highest contact rate at the lowest cost is to vary calling times. Many households will be easily contacted within the first couple of calls, but for those which are not it is important to make sure that successive visits are at different times of the day (including evenings) and on different days of the week.

Interviewers were required to call at each address in their monthly quota of 13 addresses (new cohort) at least six times with at least two calls after 6pm and a call on a Saturday. Best practice procedures whereby interviewers varied their calling times and days across the field period were employed to maximise response to the WAS.

Unable to visit sampled addresses during the pandemic, alternative strategies e.g. the use of tele-matching and an on-line portal to gather telephone numbers, were developed to permit initial contact to be made with sampled households. More detail on the changes made can be found in the Annex to this guide.

## ***Field sampling procedures***

Some occupied dwellings are not listed on the PAF. This may be because a house has been split into separate flats, only some of which are listed. If the missing dwelling could be uniquely associated with a listed address, a divided address procedure was applied to compensate for the under-coverage. In these cases, the interviewer included the unlisted part in the sample only if the associated listed address had been sampled.

Until the start of Round 8 where an interviewer discovered a concealed multi-household address in England, Wales or Scotland the same procedure was followed. The interviewer listed all the addresses found using a standard method, and then used a kish grid to select one address to be sampled.

Unable to identify concealed multi-household addresses when making contact via the phone, this approach was set aside for the duration of Round 8, with the first contacted address invited to participate.

## **Response rates**

WAS achieved an average response rate of 55 percent for Wave 1, with fieldwork being conducted between July 2006 and June 2008. The achieved sample for Wave 1 was issued for re-interview between July 2008 and June 2010, yielding an improved response of average response rate of 68 percent for Wave 2.

From Wave 3 onwards the original sample and the additional, supplementary sample have had their information collected as separate cohorts, producing average response rates for:

- Wave 3, from July 2010 to June 2012, of 65 percent
- Wave 4, from July 2012 to June 2014, of 66 percent
- Wave 5, from July 2014 to June 2016, of 65 percent
- Round 6, from April 2016 to March 2018, of 63 percent
- Round 7, from April 2018 to March 2020, of 58 percent
- Round 8, from April 2020 to March 2022, of 41 percent.

Changes to collection methods in response to the pandemic, e.g. switching from face-to-face to telephone interviewing, for the whole of Round 8 contributed to the lower recorded response rates as larger samples were issued to compensate for lower expected contact with the objective of maintaining achieved sample sizes.

The following table provides a more detailed breakdown of the outcome of cases included in the set sample for Waves 1 to 5 and Rounds 6 to 8.

**Table 13 Response Rates from Wave 1 to Round 8 by Cohort**

Outcome	Issued cases	Eligible cases	Co-operating households <sup>1</sup>	Non-contacts	Refusal to HQ	Refusal to interviewer	Total Refusal	Other non-response	Response rate	Non-contact	Refusal to HQ	Refusal to interviewer	Other non-response
<b>Wave 1</b>	61,917	55,835	30,511	3,889	3,805	15,397	19,202	1,770	55%	7%	7%	28%	3%
<b>Wave 2</b>	32,195	29,584	20,009	2,717	1,268	4,527	5,795	1,063	68%	9%	4%	15%	4%
<b>Wave 3 old</b>	25,234	21,397	15,517	1,503	809	2,868	3,677	700	73%	7%	4%	13%	3%
<b>Wave 3 new</b>	12,683	11,297	5,734	988	876	3,296	4,172	403	51%	9%	8%	29%	4%
<b>Wave 3 all</b>	37,917	32,694	21,251	2,491	1,685	6,164	7,849	1,103	65%	8%	5%	19%	3%
<b>Wave 4 old</b>	27,062	23,199	16,238	1,648	930	3,435	4,365	948	70%	7%	4%	15%	4%
<b>Wave 4 new</b>	8,269	7,417	3,894	614	561	2,060	2,621	288	53%	8%	8%	28%	4%
<b>Wave 4 all</b>	35,331	30,616	20,132	2,262	1,491	5,495	6,986	1,236	66%	7%	5%	18%	4%
<b>Wave 5 old</b>	26,739	22,795	15,622	1,783	782	3,487	4,269	1,121	69%	8%	3%	15%	5%
<b>Wave 5 new</b>	6,002	5,371	2,793	483	423	1,449	1,872	223	55%	9%	8%	27%	4%
<b>Wave 5 all</b>	32,741	28,166	18,415	2,266	1,205	4,936	6,141	1,344	65%	8%	4%	18%	5%
<b>Round 6 old</b>	22,441	19,655	13,926	1,162	625	2,720	3,345	1,222	71%	6%	3%	14%	6%
<b>Round 6 new</b>	9,712	8,676	4,033	864	655	2,684	3,339	440	46%	10%	8%	31%	5%
<b>Round 6 all<sup>2</sup></b>	32,153	28,331	17,959	2,026	1,280	5,404	6,684	1,662	63%	7%	5%	19%	6%
<b>Round 7 old</b>	19,836	18,583	12,815	1,266	559	2,793	3,352	1,420	68%	7%	3%	15%	8%
<b>Round 7 new</b>	12,987	11,608	4,726	1,136	944	4,105	5,049	667	41%	10%	8%	35%	6%
<b>Round 7 all</b>	32,823	30,461	17,541	2,402	1,503	6,898	8,401	2,087	58%	8%	5%	23%	7%
<b>Round 8 old</b>	20,316	19,114	10,882	4,724	246	2,378	2,624	884	54%	23%	1%	12%	4%
<b>Round 8 new</b>	16,377	15,852	4,268	9,176	221	1,872	2,093	315	26%	56%	1%	11%	2%
<b>Round 8 all</b>	36,693	34,966	15,150	13,900	467	4,250	4,717	1,199	41%	38%	1%	12%	3%

- 
1. A small number of responding households may be removed from the final analysis files while processing.
  2. For wave two, the achieved Wave 1 sample was issued, plus all of the non-contacts. Due to the change from a July to an April start of fieldwork, the Round 6 response rates include the last 3 months of Wave 5 (April, May and June 2016).

## **Keeping in Touch**

WAS is a longitudinal survey that follows all adults interviewed in Wave 1 (original sample members, or OSMs). The survey is biennial, i.e. two years in-between each interview. WAS, like other longitudinal surveys, experiences attrition, which may occur for inevitable reasons such as death, or for reasons that can be minimised such as failure of tracing, failure of contact, or refusal.<sup>1</sup>

The longitudinal design of WAS requires following OSMs over time to measure changes in wealth. It is evident that tracing and following sample members becomes difficult when circumstances of sample members, in particular their location, change over time<sup>2</sup>. To minimise attrition caused by the loss of sample members due to the failure of tracking, WAS has implemented measures in the survey design to maximise the likelihood of contact being made with the sample member at the next wave.

Firstly, the WAS questionnaire asks respondents at the interview to confirm their address details as well as further contact details such as phone numbers, email address, and contact details of two nominated persons (not resident at the same address) that are authorised to provide ONS with the respondent's new address in case the respondent has moved and cannot be traced. Secondly, a few weeks after the interview all respondents receive a 'Change of Address' card together with the posted incentive (alternatively this will be sent by email), which aims to encourage respondents to inform the ONS if their contact details change. Thirdly, a brief telephone interview is conducted prior to the next wave's interview. This telephone interview is referred to as the 'Keep in Touch Exercise', or KITE. During this interview information about household members as well as their address and contact details are confirmed or updated. It provides the opportunity to identify movers from the household, and their new contact details; as well to identify joiners to the household.

## **Questionnaire Content**

### **Overview**

The Wealth and Assets Survey (WAS) collects data on a wide range of assets and liabilities that private individuals and households in Great Britain have. The primary aim of the survey is to derive overall estimates of wealth and monitor how these change over time. WAS broadly splits wealth into four categories:

- Financial wealth
- Pensions wealth
- Physical wealth
- Property wealth

The questionnaire is designed to collect relevant information across these four domains of wealth, to provide aggregated measures of wealth, but also to afford significant potential for analysis within these four domains. The questionnaire is therefore both broad and detailed in coverage, with a wide range of stakeholders interested in the data WAS provides.

---

<sup>1</sup> Portanti, M.: "Attrition on Longitudinal Survey – Literature Review", ONS Working Paper, Social Survey Division, November 2009, pg. 2

<sup>2</sup> Plewis, I., 2007. Non-Response in a Birth Cohort Study: The Case of the Millenium Cohort Study. *International Journal of Social Research Methodology*, 10(5), p. 3

Laurie, H., Smith, R. & Scott, L., 1999. Strategies for Reducing Nonresponse in a Longitudinal Panel Survey. *Journal of Official Statistics*, 15, p. 269

The Wave 1 questionnaire content was determined by the requirements of the WAS consortium of government departments at that time; namely the Department for Business Innovation and Skills (BIS); Department for Work and Pensions (DWP); HM Revenues and Customs (HMRC); HM Treasury (HMT); the Office for National Statistics (ONS); the Department for Communities and Local Government (DCLG) and the Cabinet Office (CO). The primary focus of the questionnaire was to provide estimates of wealth; however some additional information was also collected on non-wealth topics such as socio-demographic characteristics, income and financial acuity. This allowed for aggregate and component analysis of wealth with other factors.

For Round 8 the average time taken to complete the questionnaire was 54 minutes.

### ***Questionnaire changes***

WAS is a longitudinal survey and therefore in order to measure change over time the questionnaire needs to be as stable as possible; so as to reduce discontinuities in the outputs. However, there is scope to make changes to the questionnaire between waves to adopt harmonised question standards and/or meet emerging information requirements.

Changes between waves are made with consortium agreement. Sponsoring departments provide information requirements and specify any requested changes. These changes are discussed by the WAS Technical Group (TG), with recommendations for questionnaire changes submitted to the WAS Steering Group (SG) for approval. The WAS SG is formed from senior representatives of the consortium departments.

Recommended questionnaire changes were subject to cognitive question testing and quantitative piloting. Cognitive question testing had the following objectives:

- ascertain whether the proposed questioning will address the information needs identified by key users and stakeholders, from the respondents' perspective
- establish what respondents understand the questions to mean and the terminology used
- understand how respondents formulate their answers and by so doing ensure that the questions are interpreted as key users and stakeholders intended
- ensure that response options are comprehensive
- ensure that respondents are willing to provide answers
- ensure that respondents are able to provide answers
- ensure that the order in which the questions are asked does not affect the answers given
- address issues relating to the collection of proxy data (if proxy information can be collected).

Quantitative piloting aimed to provide a test run of the new questionnaire, to identify any issues with the questionnaire before the next wave's data collection starts. Interviewer de-briefs were held following the pilot to seek feedback on the questionnaire and any areas for improvement. The pilot also provided the opportunity to produce survey metrics such as interview length (broken down by topic area) and indicative response

Reacting at pace as the pandemic started, and the need to shorten the length of the questionnaire as a switch from face-to-face to telephone interviewing was rapidly introduced, it wasn't possible cognitively test or quantitatively pilot in advance of the questionnaire changes introduced for Round 8. Detail on questionnaire changes prompted by the pandemic can be found in the Annex to this guide.

### ***Programming and testing***

The Wealth and Assets Survey data has collected using CAPI and CATI approaches. The software loaded into interviewer's laptops is called Blaise. All ONS social surveys use Blaise for interviewing as ONS feel that it has the flexibility and technical capability to cope best with the complexity of social research surveys. Blaise's powerful programming language offers numerous

features, and its data entry program supports a variety of survey processing needs<sup>3</sup>.

A number of features of Blaise are particularly advantageous for this survey:

- Blaise CAPI and CATI scripts have an in-built hierarchical block structure that effectively makes all questionnaires modular. The ability to handle the associated routing of a modular questionnaire is core to Blaise's architecture. In addition to its hierarchical block structure, Blaise also allows the creation of 'blocks' which can be accessed in parallel, allowing interviewers to switch out of 1 set of hierarchical blocks to another set. This provides valuable flexibility as it, for instance, allows an interviewer to pause an interview with 1 household member, initiate an interview with another household member (e.g. a household reference person), and then resume the interview with the original household member at a convenient time in the future
- Blaise meets the requirement of being able to split the sample geographically or by sample identifiers. Separate questions can be allocated to these different sections of the sample or to randomly selected sub-samples of different sizes
- handling complex routing (including loops and repeated events), applying automatic logic and consistency checks in real time during the interview, and using text fills where required, are all core to Blaise's architecture. They are functions that we make extensive use of on the Wealth and Assets Survey
- Blaise allows interviewers to exit and restart interviews at any point which allows interviews to be suspended and resumed

The Wealth and Assets Survey questionnaire records the length of time spent on different questions during interviews, by placing 'time stamps' at the start and end of different questions. We can use the session log file (called the audit trail in Blaise) to time individual questions. This method affords us the ability to monitor how different questions contribute to the overall length of the questionnaire, which is essential when conducting questionnaire content reviews.

Other features of Blaise which make it excellent for undertaking the Wealth and Assets Survey include:

- the ability for interviewers to back track in instances where later sections of an interview highlight an error made earlier
- flexibility over styles, fonts, font sizes and colours. Blaise allows these to be specified for all text or for individual words/questions. This helps ensure the screen seen by the interviewer is as well designed as possible, with effective interviewer prompts. This in turn helps promote interviewer-respondent rapport, thereby contributing to better data quality
- the ability to interact with a 'question by question' (QbyQ) help facility. This provides interviewers with real-time access to guidance on specific questions during the interview. This is an electronic programme that operates in conjunction with Blaise.

The Wealth and Assets Survey questionnaire is tested extensively prior to being scattered to field interviewers. Currently, staff in the research team independently test the questionnaire; along with staff in ONS Survey Support operations team. Questionnaire testing is done every month prior to the questionnaire scatter for the next fieldwork period.

## Occupational Coding Update

Like most surveys WAS collects and codes information on the occupation of each respondent using coding frames embedded within the questionnaire. Coding to the [Standard Occupational Classification \(SOC\)](#), the surveys of ONS aim to co-ordinate the introduction of updated classifications to minimise discontinuities within time series, while maintaining the consistency of

---

<sup>3</sup> <http://www.blaise.com/capabilities>

collection and coding by interviewers across surveys each month. Updates to the SOC used are typically introduced from April for the household financial social surveys of ONS.

Collecting and analysing data over a 2 year Round this can lead to an updated SOC being introduced midway through a Round. The household financial social surveys of ONS started to use [SOC 2020](#) from April 2021. For WAS Round 8 this resulted in Year 1 of the Round being coded to SOC 2010 with Year 2 coded to SOC 2020. To bring consistency to the occupational codes Year 1 data were mapped from SOC 2010 to SOC 2020 for Round 8.

Implemented simultaneously across ONS household financial surveys, the updated SOC 2020 coding frames embedded within the questionnaires were introduced with some information missing. Causing marginal net changes at the SOC Major Group level, the impact of the error upon analyses completed at this level of aggregation to occupation will be minimal. Further information can be found within [The impact of miscoding of occupational data in Office for National Statistics social surveys, UK](#).

Previously included to the more detailed 4 digit level of the classification within the ONS Secure Research Service (SRS) and Government Licence (GL) versions of the microdata, the SOC 2020 variables included have been limited to the 2 digit level for Round 8. More affected by the error beneath the SOC Major Group level, SOC 2020 variables presented on the SRS and GL datasets match those included on the End User Licence dataset.

With the National Statistics Socio-Economic Classification (NS-SEC) derived from a combination of age, educational status, occupation & employment status, within the questionnaire NS-SEC was re-derived outside of the questionnaire for Year 1 using the 2020-21 data mapped to SOC 2020.

## Editing

An extensive range of computer edits are applied to both the household and individual questionnaires during data entry in the field and to the aggregate data file in the office. Round 7 checks followed a similar pattern to those previously completed in Round 6.

These edits checked that:

- logical sequences in the questionnaire had been followed
- all applicable questions had been answered
- specific values lay within valid ranges
- there were no contradictory responses
- relationships between items were within acceptable limits

Edits are also designed to identify cases for which values, although not necessarily erroneous, were sufficiently unusual or close to specified limits as to warrant further examination.

Once an interview had taken place, the WAS data are transmitted back to ONS and aggregated to monthly files. Further editing occurred at this stage and included:

- recoding text entries if an appropriate response category was available
- investigating interviewer notes and utilising the information where applicable
- confirming that overridden edit warnings had been done correctly
- broad data consistency checks.

The next stage involved checking that the routing of the questionnaire output is correct, using a process referred to as 'base checks'. SPSS programmes are run to emulate the routing performed in Blaise. This process is used to identify where Blaise appeared to incorrectly route respondents. This can either be corrected for by recoding data, or, where cases haven't been routed as they should have been; imputation requirements are specified. Where errors in routing are discovered, the Blaise questionnaire was corrected to enhance the quality of future data collection. The sooner base checks are performed; the sooner the Blaise questionnaire can be corrected; thus leading to

lower levels of data imputation.

Due to the longitudinal component of the survey design, part of the achieved sample size in Round 8 is linkable to previous waves of data. Therefore, it was important to introduce longitudinal edit checks to the existing editing and validation processes.

The edit and validation checks were run in two stages, whereby first cross-sectional checks were carried out on the seventh round to validate or edit outliers. Checks for financial and pension wealth data were exclusively done on individual level because the data is collected for each individual in the household, while checks for the property and physical wealth data are carried out at the household level. The investigation of outliers largely focused on the top and bottom ten per cent of the distribution of each wealth component, although for some variables this proportion was reduced if the number of cases highlighted for investigation was particularly high. When outliers were investigated in the pensions or the financial section, various variables within the same wealth component section or even different sections of the questionnaire were included to establish whether particularly large outliers could be explained by the circumstances of respondents. The majority of investigated cases proved to be genuine and only a small number of cases had to be edited, data was only edited if sufficient information was recorded by interviewers to establish the correct response, and the majority of investigated cases proved to be genuine with only a small number of cases being edited.

The second stage of checks was conducted after the linkage exercise was completed. At this stage the change of wealth components between the two waves was calculated and subsequently outliers of change were highlighted. To investigate these longitudinal outliers, the circumstances of relevant respondents in both current and previous waves had to be considered to decide whether the value in either Round 7 or Round 8 was correct. As with the cross-sectional checks, only a small number of longitudinal corrections were made for each wealth component variable where enough information was available.

## Imputation

### *General Methodology*

In a similar way to all social surveys, data from the Wealth and Assets Survey (WAS) contains missing values. Item non-response occurs when a respondent does not know or refuses to answer a survey question. Unit non-response refers to cases where an individual in a responding household refuses to be interviewed or contact cannot be made. Item and unit non-response can be problematic in that many standard analytical techniques are not designed to account for missing data. More significantly, missing data can lead to bias, error, and inconsistencies in estimates and publication figures. Imputation is a statistical process that serves to counter item non-response by estimating the statistical properties of the missing data. These estimates are used to replace missing data with valid, plausible values.

Information about discrete assets or liabilities recorded by the Wealth and Assets Survey are collected through a relatively consistent question structure. Typically, an affirmative response to routing questions designed to determine; *do you have asset/liability x?* is followed by a question to specify the value; *what is the amount/income/expenditure of asset/liability x?* In cases where an exact amount is not known, participants are asked to provide a banded estimate from a range of bound values such as £0 to £100, £101 to £500, and so on.

The key analytical aim of the WAS is to provide longitudinal estimates of change over time as well as cross-sectional/single year estimates. To meet this aim, the imputation strategy is designed to estimate the longitudinal and cross-sectional properties of the missing data depending on the availability of data from previous waves. The strategy is designed to preserve relationships in the data defined by **implicit laws or regulations governing the absolute value of an asset or liability**. Important or significant relationships between variables in the end-to-end question set

and between persons in a household are also taken into account.

In general, the WAS imputation strategy is based on non-parametric Nearest-Neighbour donor-based imputation methodology (Bankier, Lachance, & Poirier, 1999<sup>4</sup>; Durrent, 2005<sup>5</sup>; de Waal, Pannekoek, & Scholtus, 2011<sup>6</sup>). In this framework, missing data is replaced with plausible values drawn from other records in the data belonging to respondents with similar characteristics. Donor based methods serve to avoid the distributional assumptions associated with parametric methods. Consequently, they are robust; imputed values are always consistent with values actually observed in the data; and imputation can be applied under strict edit constraints ensuring relationships between variables are appropriately maintained. Significantly, if applied correctly, donor-based methods will preserve the conditional statistical distributions in the observed data and/or adjust them in the presence of a non-response bias (Rubin, 1987<sup>7</sup>; Chen & Shoa, 2000<sup>8</sup>, Durrent, 2005).

### **Donor Selection**

The key to a successful application of Nearest-Neighbour imputation is the selection of a pool of suitable 'potential' donors. In general, selection is based on information specified by other 'auxiliary' variables in the data. The set of auxiliary variables is typically referred to as the imputation model. The imputation model ensures that imputed values are drawn from a representative distribution of plausible values from respondents with similar characteristics. For all imputed variables in the Wealth and Assets Survey, appropriate auxiliary variables were identified through traditional regression-based modelling supplemented by guidance from experts familiar, not only with a particular subject domain, but also with the analytical program designed to provide outputs that meet customer needs.

For a discrete imputable record, the pool of potential donors is determined by calculating the 'distance' between a record that needs repair and other fully observed respondent records, keeping only those that match with minimum distance. Where appropriate, auxiliary variables in the imputation model are given a higher weight to account for cases where some auxiliary information is more important. In general, one of two distance functions were used to calculate the distance between the potential donor and the recipient record, depending on the characteristics of each particular auxiliary variable.

---

<sup>4</sup> Bankier, M., Lachance, M., & Poirer, P. (1999). A Generic Implementation of the New Imputation Methodology. Statistics Canada methodology paper.

<sup>5</sup> Durrant, Gabriele B. (2005) Imputation Methods in the Social Sciences: A Methodological Review. National Centre for Research Methods, School of Social Sciences, University of Southampton.

<sup>6</sup> De Waal, R., Pannekoek, J., & Scholtus, S. (2011). Handbook of Statistical Data Editing and Imputation. New Jersey: John Wiley & Sons.

<sup>7</sup> Rubin, D. B. (1987). Multiple Imputation for Nonresponse in Surveys. New York: John Wiley & Sons.

<sup>8</sup> Chen, J., & Shao, J. (2000). Nearest Neighbor Imputation for Survey Data. *Journal of Official Statistics*, 16(2), 113-131.

$x_f$  = the recipient record with  $n$  auxiliary variables  
 $x_d$  = the potential donor record with  $n$  auxiliary variables

$$D_{fd} = \sum_{i=1}^n \omega_i D_i$$

$\omega_i$  = the weight for the  $i^{\text{th}}$  variable  
 $D_i$  = the individual distance for the  $i^{\text{th}}$  variable

For categorical data with no ordinal relationship between categories:-

$$(2) \quad D_i = \begin{cases} 0 & \text{where } x_f = x_d \\ 1 & \text{where } x_f \neq x_d \end{cases}$$

For categorical or continuous data with an ordinal and/or ratio relationship between categories or values:-

$$(2) \quad D_i = \begin{cases} 1 & \text{if } |x_f - x_d| \geq y \\ 1 - \left(1 - \frac{|x_f - x_d|}{y}\right) & \text{otherwise} \end{cases}$$

$y$  = desired minimum( $|x_f - x_d|$ ) at which point and beyond  $D_i = 1$

The final imputed value is selected from the donor pool based on the probability distribution associated with the range of plausible values.

Table 15 shows a typical example of an auxiliary variable set used to impute a respondent's private pension. All Wealth and Asset variables were treated in a similar way.

**Table 14 - Imputation Class and Matching Variables used for imputing values for Private Pensions**

Imputation Class		Matching Variable		
Variable	Classification	Variable	ω	Classification
Banded Estimate	1: Less than £2,500	Previous value of private pension	0.50	N/A
	2: £2,500 > £4,999	Banded Net Salary (An.)	0.25	12 bands
	3: £5,000 > £9,999	Employment Status	0.05	1: Employee 2: Self-Employed
	4: £10,000 > £19,999	Single year of age	0.25	>=0
	5: £20,000 > £49,999	Sex	0.05	1: Male 2: Female
	6: £50,000 > £99,999	NS-SEC	0.05	1: Professional 2: Intermediate 3: Routine 4: Never worked 5: Unclassified
	7: £100,000 or more	Employment Sector	0.05	1: Private 2: Public 3: Other
		Household Type	0.05	1: Single Person 2: Couple with no children 3: Couple with children 4: Lone parent 5: More than one family
		Government Office Region	0.25	12 locations

To impute missing values for the estimated amount in a private pension, donors were selected from an imputation class derived from the Banded Estimates. The Banded Estimate was either given by the respondent when they were not able to answer the exact amount or imputed beforehand where not available<sup>9</sup>. The Banded Estimate provided an important constraint on donor selection based on observed data.

The matching variable set consisted of variables related to the observed data identified through modelling and domain- expert review. Where the previous pension value is observed, this is the highest weighted matching variables Processing Strategy.

The Wealth and Assets Survey data were processed in four sections: Income, Property & Physical, Pensions & Financial. For all variables, imputation followed a basic processing strategy. First, missing routing variables for a particular asset/liability were imputed against an appropriate set of auxiliary variables. Following that, where the routing indicated a missing value for the amount associated with a particular asset/liability, the value was imputed against its own set of auxiliary variables. The simplification of the imputation strategy meant that cross-sectional and longitudinal groups (i.e., those with previous observations and those without) could be processed together. Processing focused on imputing a discrete category or value drawn from the range and distribution of categories/values observed directly in the data of records reaching the final potential donor pool. Any relevant longitudinal observations were used in the imputation, as an auxiliary variable, to help ensure interdependencies and rates of change in the data between previous and current rounds were preserved.

### **Quality Assurance and Evaluation**

For all Wealth and Asset Survey variables, the imputed data was examined and tested before being formally accepted. The overarching aim of the evaluation was to ensure that the distributional properties of the observed data had not been distorted inappropriately by the imputation process. Fundamentally, evaluation was based on comparing the observed data prior to imputation with the fully imputed data. In all cases, any notable departures from the observed data based on statistical measures such as shifts in central tendency or variance and/or the introduction of unexpected changes in the shape of the distribution had to be justified. This preliminary evaluation was supplemented by a more detailed review of the utility of the data by topic experts familiar, not only with the analytical aims of the survey, but also with expected data trends and characteristics inferred from other reliable external data sources.

### **Percentage of Data Imputed for Round 8**

For Round 8, **665 variables underwent nearest neighbour imputation. The percentage of data that was imputed for each variable varied widely between less than 0.5% and 90%. This is an increase in both the number of variables being imputed and the percentage of data imputed for each variable from Round 7, where the percentage of data imputed for each variable varied between less than 0.5% and 62%. For the majority of variables, however, less than 10% of data was imputed. The number of variables with varying levels of data imputed in Rounds 7-8 are summarised in Table 15.**

**Table 15 – Imputation Rates for Variables in Round 7 and Round 8**

Percentage of data imputed	Number of variables in Round 8	Number of variables in Round 7
Below 10%	414	328
10-20%	84	61
20-40%	79	52
Over 40%	88	26
Total	665	467

<sup>9</sup> Banded Estimates imputations would use the same matching variable set as the continuous variable.

# Weighting<sup>10</sup>

## Overview

Survey weights are a prerequisite to using WAS data for both cross-sectional and longitudinal analyses of wealth in Great Britain. WAS is designed to follow the same people over time; this is achieved through consecutive waves of interviews. This longitudinal perspective of the survey allows for estimation of gross change over time.

The cross-sectional perspective of the survey is another important feature as it allows for estimation of wealth at certain time points. Top-up panels have been introduced from Wave 3 onwards to firstly increase the sample size and to secondly update the sample.

This chapter will discuss the survey weighting methods applied to calculate the Round 8 (R8; April 2020 – March 2022) longitudinal and cross-sectional weights. Methods used to account for attrition, non-response and the complex multi-panel design are discussed. Some properties of the final weights are presented.

An important change from the established methodology used in Round 7 and earlier concerned the inclusion of tenure as an additional control variable in the final calibration step of the weighting pipeline. This was introduced to counteract the increased selection bias in the achieved sample due to the Covid-19 pandemic (see Kastberg and Siegler 2022)<sup>11</sup> and followed the example of the Labour Force Survey (Watson 2020)<sup>12</sup> and other ONS social surveys.

Note that the switch from waves (two-year reporting periods starting *July* of even years) to rounds (start in *April*) from Round 6 onwards had previously necessitated some changes to the established weighting methodology for WAS waves; these are discussed in Zobay et al (2020) and the Round 6 and 7 user guides.

## Different Types of Weights

As the survey develops there are numerous longitudinal weights that could be calculated from the many different combinations of the waves. From Wave 3 onwards, four types of weights were produced; for Round 8, these are as follows:

- longitudinal weights for the survivors from Round 2 to Round 8 (R2-R8)
- longitudinal weights from Round 7 to Round 8 (R7-R8)
- cross-sectional weights for Round 8 (R8)
- cross-sectional non-proxy weights for Round 8.

The survivor weights are computed for individuals who responded in all rounds from Round 2 (April 2008 – March 2010) to R8. Note that for WAS waves, survivor weights were calculated for survivors from Wave 1 (July 2006 – June 2008). With the change from Waves to Rounds, it was decided to use Round 2 as the basis for the weights rather than Round 1 as the latter covers only

---

<sup>10</sup> For more detailed descriptions of the weighting process, see Davies R (2016), *The weighting methodology for Wave Four of the Wealth and Assets Survey*, ONS Survey Methodology Bulletin **75**, 42-56, <https://www.ons.gov.uk/file?uri=/methodology/methodologicalpublications/generalmethodology/surveymethodologybulletin/surveymethodologybulletinno.75spring2016.pdf>; Zobay O, Law E, and Merad S (2020), *Changes to the Wealth and Assets Survey weighting system because of a shift in the reporting period*, ONS methodology paper,

<https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/earningsandworkinghours/methodologies/changes-to-the-wealth-and-assets-survey-weighting-system-because-of-a-shift-in-the-reporting-period>

<sup>11</sup> Kastberg S and Siegler V (2022), *Impact of COVID-19 on ONS social survey data collection*, ONS methodology paper,

<https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/methodologies/impactofcovid19ononsocialsurveydatacollection>

<sup>12</sup> Watson B (2020), *Coronavirus and its impact on the Labour Force Survey*, ONS article,

<https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/articles/coronavirusanditsimpactonthelabourforcesurvey/2020-10-13>

seven quarters of sampling.

The longitudinal weights for the latest two consecutive rounds are calculated for all individuals who responded in both R7 and R8. The cross-sectional weights serve to make the full set of R8 responders representative of the target population, i.e., residents of private households in Great Britain at the time of R8. In contrast, the non-proxy weights are calculated for the subsample of adult non-proxy respondents and children (including those aged 16-18 in full-time education). These weights can be used to analyse questionnaire items where there are no responses by proxy or where their validity may be questionable, e.g., for questions on attitudes and opinions.

### **Longitudinal R7-R8 Weights**

The longitudinal weighting strategy is based on a principle of updating the initial selection probabilities by means of the ongoing loss-to-follow-up adjustments that remain for the evolving respondent subset over time. This is achieved through developing the longitudinal base weight (see, e.g., Verma et al. 2007)<sup>13</sup>. This principle enables the weights to refer back to the desired population as closely as possible with the current sample design and respondent follow-up procedures. A summary of the computation of the R7-R8 weight is given below, followed by a more detailed description of some of the steps:

- separate the dataset for the R7 cross-sectional weights into the original wave- and round-based panels (i.e., W1, W3, W4, W5 Q1-Q7, R6, R7).
- for each panel, compute and apply attrition adjustments for drop-out from unknown eligibility and non-response between R7 and R8. This yields the longitudinal base weights.
- combine the panels according to their effective sample sizes using the base weights computed in step 2.
- form quarterly datasets and calibrate each to the midpoint of R7 (March 2019) using the standard calibration groups (i.e. age-sex groups and Government Office Regions).
- rescale each of the eight quarterly sets by 1/8 and combine them to form the R7-R8 longitudinal weights.

#### *Attrition*

As a first step to attrition modelling, every R7 responder is classified according to their R8 response status as either R8 responder (R), ineligible case (IE), non-responder (NR) or case with unknown eligibility (UE). R8 ineligible cases are those who do not fulfil the R8 sample inclusion criteria, for example because of death or emigration. Non-responders are known to be still in the R8 population but were not productive in R8, most often because of refusal. For some cases, R8 eligibility cannot be established, for example, if the current location of the household is unknown.

If there were no NR and UE cases, the original R7 population would be faithfully represented by the R8 R and IE cases and their weights, and no reweighting would be necessary. To deal with attrition, a two-step process is applied. First, UE attrition is handled by deriving an adjustment factor from a suitably defined logistic regression model with eligibility as outcome (i.e., unknown versus known eligibility) and fitted using all R7 responders. As all these cases have R7 data, a rich set of person- and household-level predictors is available. Models are fitted for each panel separately. The eligibility adjustments  $w_{8,i}^E$  are calculated as

$$w_{8,i}^E = \frac{1}{\hat{\phi}_{8,i}^E}, i \in s_8^E$$

where  $\hat{\phi}_{8,i}^E$  is the predicted probability that the eligibility status of case  $i$  at R8 is known, as estimated from the logistic regression model, and  $s_8^E$  is the set of people who have a known

---

<sup>13</sup> Verma V, Betti G, Ghellini G (2007), *Cross-sectional and longitudinal weighting in a rotational household panel: applications to EU-SILC*, Statistics in Transition New Series 8, 5-50.

eligibility status at R8. These individuals are weighted up to compensate for the loss of cases whose eligibility status has become unknown between R7 and R8.

The second attrition model estimates the probability  $\hat{\phi}_{8,i}^R$  of response, using a logistic regression fitted on the R8 responders and non-responders, again by panel, and including R7 individual and household level characteristics as predictor variables. The non-response adjustment is given by

$$w_{8,i}^R = \frac{1}{\hat{\phi}_{8,i}^R}, i \in s_8^R$$

with  $s_8^R$  the set of individuals within a responding household at R8. These cases are weighted up to compensate for loss to follow-up because of non-response.

For R8 responders, the longitudinal base weight  $w_{8,i}^{long}$  is the product of the R7 cross-sectional weight  $w_{7,i}^{XS}$  and the two attrition adjustments. For IE cases, the base weight is the product of  $w_{7,i}^{XS}$  and the unknown-eligibility adjustment.

$$w_{8,i}^{long} = \begin{cases} w_{7,i}^{XS} w_{8,i}^E w_{7,i}^R & i \in s_8^R \\ w_{7,i}^{XS} w_{8,i}^E & i \in s_8^{IE} \end{cases}$$

The base weights are trimmed at the 99<sup>th</sup> percentile to reduce the effects of outliers and scaled to the R6 population total.

### Combining panels

Regression models are derived for each panel separately. After applying the attrition adjustments, the panels are combined in proportion to their respective effective sample sizes (step 3 above). The effective sample size of panel  $p$  is defined as  $n_{eff,p} = n_p / (1 + CV_p^2)$  with  $n_p$  the actual sample size and  $CV_p$  the coefficient of variation of the panel's weights (i.e., standard deviation divided by mean). When combining by effective sample size, weights in panel  $p$  are multiplied by a factor  $n_{eff,p} / N_{eff}$  with  $N_{eff}$  the sum of the effective sample sizes of all panels.

### Calibration

The combination of eligible responders and ineligible outflows, together with their attrition-adjusted base weights, can be expected to have a good degree of representativity of the population to which it relates back to. For the R7-R8 dataset, this will be the target population at the midpoint of R7.

As a further improvement, a final calibration step is applied to the base weights. The purpose of calibration is to make the weights of certain subgroups of the sample sum to known population totals. Calibration is performed in such a way that the differences between the pre-calibration weights and the calibrated weights are minimised. As mentioned in Step 4 above, calibration is done on quarterly datasets after panels have been combined. Calibration constraints are applied for age-sex bands and regions whereas tenure is not controlled for in the derivation of the R7-R8 longitudinal weights.

### R2-R8 survivor weights

Survivor weights are obtained in an iterative process starting from the R2 cross-sectional weights.

- **R2 weights:** R2 cross-sectional weights are computed for all R2 responders using a procedure analogous to the one for R5 weights outlined in Zobay et al (2020) and the WAS R6 and R7 user guides.
- **R2-R3 weights:** The computation of R2-R3 weights follows along the lines of the R7-R8 calculation but includes only a single panel. The R2 cross-sectional weights of R3 responders and ineligible cases are first adjusted for attrition and then calibrated to the R2

population totals.

- *R2-R4 weights*: As a first step, cases who responded in both R2 and R3 are classified according to their R4 response status. The R2-R3 weights of the R4 responders and ineligible cases are then adjusted for R4 attrition. Their new weights are combined with the unadjusted R2-R3 weights of the R3 ineligible cases, and the complete set is calibrated to the R2 population totals
- *R2-R5 and subsequent weights*: The R2-R4 weights of R5 responders and ineligible cases are adjusted for attrition between R4 and R5 and then combined with the R2-R4 weights of the R2 and R3 ineligible cases before calibration to R2 population totals. This yields the R2-R5 weights. R2-R6 and subsequent weights are computed accordingly.

In practice, since the R2-R7 survivor weights were already available from the previous round, the actual calculations that needed to be carried out in R8 only involved the update step from the R2-R7 to the R2-R8 weights.

### **Cross-sectional weights**

An R8 pseudo cross-sectional weight has been created; the designation “pseudo” stems from the fact that the set of weights contains R8 respondents from all panels. Therefore, the samples that were selected in, for example, R1 or R4 may not be representative of the population at the time of R8.

The calculation of the cross-sectional weights proceeds in two main steps:

- for each of the W1, W3, W4, W5 (Q1-7), R6, R7 and R8 panels, compute a set of pre-calibration weights.
- combine the seven panels according to effective sample size and calibrate by quarter.

### **Pre-calibration weights for returning panels (W1, W3, W4, W5 Q1-7, R6, R7)**

In the process of computing pre-calibration weights for returning panels, one needs to distinguish between different types of respondents:

- continuing cases (R7 respondents)
- joiners to households (first responding in R8 and not part of the original sample)
- births between R7 and R8
- re-entrants (cases who last responded before R7)

Note that entry original sample members (EOSMs), i.e., members of the original sample or earlier boost panels who first responded in the present round, were also part of the samples up to R7 but were no longer included in R8.

The calculation proceeds as follows.

- Cases continuing from R7 are assigned their pre-calibration longitudinal R7-R8 weights  $w_{8,i}^{long}$
- In a continuing household  $h$  with joiners, a weight-sharing method is applied. The weights of all continuing household members  $i$  are summed ( $w_h = \sum_{i \in h} w_{8,i}^{long}$ ), and this sum is divided by the number  $n'_h$  of people in the household including joiners but excluding any births. Each household member is then assigned this average weight  $w_h/n'_h$ .
- The WAS weight share method was constructed following Kalton and Brick (1995)<sup>14</sup>. This standard approach is based on the continuing household members' weights not including the joiners and sharing these weights between all associated R8 household members. A key challenge for the weight share method is being able to distinguish between those

---

<sup>14</sup> Kalton G, Brick JM (1995), *Weighting schemes for Household Panel Surveys*, Survey Methodology 21, 33-44, <https://www150.statcan.gc.ca/n1/en/catalogue/12-001-X199500114412>.

joiners who are new population entrants and those who were in the original population but not originally in the sample. Unfortunately, it is not possible to make this distinction with WAS data and consequently, all joiners are treated as if they were in the population at the time the sample was drawn.

- Any births to a continuing household between R7 and R8 are assigned their mother's weight.
- After applying these steps, the weights of all continuing households in a panel are calibrated to the R8 midpoint population totals (i.e., March 2021). A calibration control for tenure is not applied at this stage.
- Re-entrants are assigned a weight that is the product of their last cross-sectional weight and an attrition factor. This factor is computed from a round-based regression model which describes the attrition rate, within the re-entrant's panel, between the round that the re-entrant last responded in and the current round (more specifically, the regression estimates the probability of responding in R8 using all cases in a panel who (a) responded in the re-entrant's last round and (b) are classified as either responders or non-responders in R8).
- Re-entrants are added into their respective panels  $p$  after applying a shrinking factor  $S_p$  to the calibrated weights of the continuing households. This factor is given by

$$S_p = \frac{\text{Pop}_{R8} - \sum_{i \in \text{Re-entrants}(p)} w_i}{\sum_{i \in R7-R8(p)} w_i + \sum_{i \in \text{Joiners}(p)} w_i + \sum_{i \in \text{Births}(p)} w_i}$$

with  $\text{Pop}_{R8}$  the R8 population total and  $w_i$  the weights that have been assigned to the cases in the previous steps. As the continuing households have already been calibrated to the population total the shrinkage factor ensures that the weights of the complete panel still sum to  $\text{Pop}_{R8}$ .

### **Pre-calibration weights for the R8 panel**

Cases for the new panel were only sampled during the first year of R8. Their pre-calibration weights were computed as follows:

- assign each case their original design weight (see below)
- apply a non-response factor computed from a logistic regression model for response propensity. The model considers all households sampled during the first year of R8 and estimates response probability using output area classification, region and the WAS wealth stratum as predictors
- calibrate the panel to the R7 midpoint population totals

### **Design weights**

In R7 and R8, primary sampling units (PSUs) for the WAS boost panels were selected from the PSUs of the Survey of Living Conditions (SLC) that were drawn for Wave 1 of the previous year. In R7, WAS households could thus be visited at the same time as SLC W2 households in a given PSU area which provided some savings on cost. However, this advantage ceased with the suspension of face-to-face interviewing after the start of the pandemic.

Design weights are obtained as the reciprocals of the initial sample selection probabilities which were derived as follows. After deciding on the sample size for the new WAS panel, the numbers of PSUs that were needed in each NUTS2 region were determined, based on the sample composition in previous rounds. WAS PSUs were then drawn by simple random sampling from the SLC PSUs in that region. If more PSUs were required in a region than were available from the SLC, PSUs from the Living Cost and Food survey (LCF) were used as a backup. The probability of selecting a WAS PSU can be shown to be given by  $p_s = n_{PSU,R} * S_p / S_R$  with  $n_{PSU,R}$  the number of sampled PSUs in the NUTS2 region  $R$  of the PSU  $P$ ,  $S_p$  the size of PSU (number of addresses on the frame) and  $S_R$  the size of the NUTS2 region.

Based on a matching against tax data, HMRC ranked all selected addresses and returned a stratification into the top 1%, 2-10% or remaining 90% presumptive “wealth groups”. To sample households within a PSU, the two higher wealth strata were oversampled by factors of 5 and 3, respectively. The sampling probability  $p_h$  of a household in wealth stratum  $m$  is thus obtained as

$$p_{h,m} = \frac{f_m * n_{\text{PSU}}}{M_{90\%} + 3M_{2-10\%} + 5M_{1\%}}$$

with  $n_{\text{PSU}}$  the total number of households sampled in the PSU,  $f_m$  the oversampling factor of wealth stratum  $m$  and  $M_m$  the total number of addresses in the stratum. The final design weight of a household is given by  $d_h = 1/(p_s * p_{h,m})$ . Note that the stratification into three groups was only introduced in R6; for previous waves two strata were used with the top 10% wealth stratum oversampled by factor of 3.

### **Household-level cross-sectional weights**

After computing the pre-calibration weights as described above, panels are combined by effective sample size and the resulting dataset is split again by quarter. Household-level cross-sectional weights  $W_{\text{HH},h}$  are then obtained by performing a specific final calibration step on the quarterly datasets. The calibrated weights have to fulfil constraints of the type

$$\text{Calibration group total} = \sum_{\text{All households } h} m_h * W_{\text{HH},h}$$

in each calibration group. Here, the summation is over all households  $h$  in a given quarter, and  $m_h$  denotes the number of household members in the calibration group in question. The input (pre-calibration) weights are the means of the person-level pre-calibration weights within households. In R8, WAS uses age-sex groups, Government Office Regions and tenure categories “owning outright”, “owning with mortgage” and “renting” as constraints in this final calibration step.

Following the integrated weighting method proposed by Lemaître and Dufour (1987)<sup>15</sup>, all analyses of WAS data are carried out with the household-level weights, with individuals being assigned household weights in person-level calculations.

The non-proxy weights are computed from the cross-sectional weights by removing all adult proxy cases and recalibrating the weights of the whole remaining sample in a single step without split into quarters.

### **Descriptive statistics for R8 weights**

The table below presents basic descriptive statistics for the various types of R8 weights. For the R8 cross-sectional household weights expanded to person level, R7-R8 longitudinal and R2-R8 survivor weights, the sum of the weights equals the R8, R7 and R2 population totals, respectively, that were used for calibration. The sum of the R8 household-level weights provides an estimate of the total number of households in R8. It is lower by about 1.1% compared to the household count that we have estimated from other sources.

For the R7-R8 and R2-R8 weights, ineligible cases need to be retained to ensure consistency in the calibration. After excluding these cases, the sum of the weights estimates the size of population remaining since R7 and R2, respectively.

<sup>15</sup> Lemaître G, Dufour J (1987), *An integrated method for weighting persons and families*, Survey Methodology Bulletin 13, 199-207, <https://www150.statcan.gc.ca/n1/en/catalogue/12-001-X198700214607>.

**Table 16 – R8 Weights**

Type of Weight	N	Sum	Mean	Std Dev	Coefficient of Variation	Minimum	Maximum
R8 household-level cross-sectional	15,130	26,990,502	1,784	2,209	124%	11.8	29,610
R8 HH weights at person level	32,293	63,927,960	1,980	2,446	124%	11.8	29,610
R7-R8 longitudinal (resp. and inelig.)	21,192	63,726,785	3,007	2,993	100%	30.1	23,457
R7-R8 longitudinal excl. ineligibles	20,592	62,750,598	3,047	3,013	99%	30.1	23,457
R2-R8 survivor (resp. and inelig.)	8,756	59,094,732	6,749	7,713	114%	116.2	67,439
R2-R8 survivor excl. ineligibles	5,691	48,342,935	8,495	8,251	97%	259.8	63,450

**Cross-sectional weights for WAS W1-R7**

The table below summarises cross-sectional weights from WAS Waves 1 to Round 8. Due to updates, there can be slight differences to data published in previous user guides.

**Table 17. Cross-sectional weights Waves 1 to Round 7**

Type of weight	N	Sum	Mean	Std Dev	Coefficient of Variation	Minimum	Maximum
R8 XS	32,293	63,927,960	1979.623	2446.42	123.6%	11.84	29610.3
R7 XS	38,930	63,726,785	1636.96	1341.84	82.0%	14.35	14952.3
R6 XS	40,540	62,933,121	1552.37	1184.74	76.3%	18.01	9262.5
W5 XS	42,896	62,134,570	1448.49	1056.47	72.9%	21.35	12387.9
W4 XS	46,451	61,169,463	1316.86	880.69	66.9%	45.79	7618.4
W3 XS	49,440	59,658,894	1206.57	877.26	72.7%	68.90	9999.0
W2 XS	46,342	59,183,270	1277.10	730.94	57.2%	105.82	3700.0
W1 XS	71,268	58,012,026	814.00	283.66	34.8%	132.51	2245.1

WAS weights are (model-assisted) design-based weights. Users can re-scale and normalise, if they so wish, in order to get the weights to sum to the sample size. However, the majority of popular statistical packages available on the market now account more accurately for the weights, so we recommend using survey-based procedures, where possible, when using weights in the analysis of survey data.

## **Data Quality**

All reasonable attempts have been made to ensure that the data are as accurate as possible. However, there are two potential sources of error which may affect the reliability of estimates and for which no adequate adjustments can be made. These are known as sampling and non-sampling errors and should be kept in mind when interpreting the WAS results.

### **Sampling error**

Sampling error refers to the difference between the results obtained from the sample population and the results that would be obtained if the entire population were fully enumerated. The estimates may therefore differ from the figures that would have been produced if information had been collected for all households or individuals in Great Britain.

One measure of sampling variability is the standard error which shows the extent to which the estimates should be expected to vary over repeated random sampling. In order to estimate standard errors correctly, the complexity of the survey design needs to be accounted for, as does the calibration of the weight to population totals (see [Weighting](#)). WAS has a complex design in that it employs a two-stage, stratified sample of addresses with oversampling of the wealthier addresses at the second stage and implicit stratification in the selection of PSUs. The actual computation of the standard errors for WAS uses the jackknife linearization variance estimator.<sup>16</sup>

For statistical disclosure reasons, information on the sample design, such as clustering and stratification, cannot be provided with the WAS datasets available at the UK Data Service. Data users therefore cannot reproduce the full standard error calculations. However, it is possible to compute useful approximations to the full results by:

- ignoring clustering
- using Government Office Regions for stratification
- setting the pre-calibration weights equal to the calibrated ones.

Calibration group indicators can be derived from DVAge6R8 and SexR8 for age-sex groups, GORR8 for regions, and ten1R8 for tenure groups with categories combined as 1, (2,3), (4,5). The resulting estimates are typically within 10% of the full calculation.

An R script detailing the approximate calculation is available on request from [wealth.and.assets.survey@ons.gov.uk](mailto:wealth.and.assets.survey@ons.gov.uk).

Further background can be found within [ONS methodology working paper series no. 9 - Guide to calculating standard errors for ONS Social Surveys](#).

Note that some initial estimates of standard errors for key variables are available in the supporting tables to the report referred to above, but imputation effects need to be taken account of, so these should be treated as preliminary: more accurate estimates would be likely to be larger.

### **Non-sampling error**

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-

---

<sup>16</sup> Yung W, Rao JNK (1996), *Jackknife Linearization Variance Estimators Under Stratified Multi-Stage Sampling*, Survey Methodology 22, 23-31, <https://www150.statcan.gc.ca/n1/pub/12-001-x/1996001/article/14388-eng.pdf>

sampling error are:

- response errors such as misleading questions, interviewer bias or respondent misreporting
- bias due to 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 impact through careful design and testing of the questionnaire, training of interviewers and extensive editing and quality control procedures at all stages of data processing. The ways in which these potential sources of error were minimised in WAS are discussed below.

Response errors generally arise from deficiencies in questionnaire design and methodology or in interviewing technique as well as through inaccurate reporting by the respondent. Errors may be introduced by misleading or ambiguous questions, inadequate or inconsistent definitions or terminology, and by poor overall survey design. In order to minimise the impact of these errors, the questionnaire, accompanying documentation and processes were thoroughly tested before being finalised for use in the first wave of WAS.

To improve the comparability of WAS statistics, harmonised concepts and definitions were also used where available. Harmonised questions were designed to provide common wordings and classifications to facilitate the analysis of data from different sources and have been well tested on a variety of collection vehicles.

WAS is a relatively long and complex survey and reporting errors may also have been introduced due to interviewer and/or respondent fatigue. While efforts were made to minimise errors arising from deliberate misreporting by respondents some instances will have inevitably occurred.

Lack of uniformity in interviewing standards can also result in non-sampling error, as can the impression made upon respondents by personal characteristics of individual interviewers such as age, sex, appearance and manner. In ONS thorough training programmes, the provision of detailed supporting documentation, and regular supervision and checks of interviewers' work are used to encourage consistent interviewing practices and maintain a high level of accuracy.

One of the main sources of non-sampling error is non-response, which occurs when people who were selected in the survey cannot or will not provide information or cannot be contacted by interviewers. Non-response can be total or partial and can affect the reliability of results and introduce a bias.

The magnitude of any bias depends upon the level of non-response and the extent of the difference between the characteristics of those people who responded to the survey and those who did not. It is not possible to accurately quantify the nature and extent of the differences between respondents and non-respondents. However, the level of non-response bias is mitigated through careful survey design and compensation during the weighting process, the latter having been discussed earlier. To further reduce the level and impact of item non-response resulting from missing values for key items in the questionnaire, ONS undertook imputation (see [Imputation](#)).

Non-sampling errors may also occur between the initial data collection and final compilation of statistics. These may be due to a failure to detect errors during editing or may be introduced in the course of deriving variables, manipulating data or producing the weights. To minimise the likelihood of these errors occurring, a number of quality assurance processes were employed which are outlined elsewhere in this guide.

### ***External source validation***

In the final stages of validating the WAS data, comparative checks were undertaken to ensure that the survey estimates conformed to known or expected patterns and were broadly consistent with data from other external sources.

In advance of publication interim data sets were also shared with DWP, HMRC & Scottish Government to permit additional quality assurance to be completed by external stakeholders.

The following guidelines were applied by ONS when undertaking the external source validation process:

- identify alternate sources of comparable data
- produce frequencies and cross tabulations to compare proportions in the WAS dataset to those from external sources
- if differences were found, assess whether these were significant
- where significant differences were found ensure that reference periods, populations, geography, samples, modes of collection, questions, concepts and derivations were comparable.

Results from these analyses indicated that estimates from the Wealth and Assets Survey were broadly in line with results from other administrative and survey sources.

## **Wealth Estimates**

The wealth estimates in this report are derived by adding up the value of different types of asset owned by households and subtracting any liabilities. Total wealth with pension wealth is the sum of four components:

- net property wealth
- physical wealth
- net financial wealth
- private pension wealth.

Total wealth without pension wealth is the sum of the first three of these components. The components are, in turn, made up of smaller building blocks:

- net property wealth is the sum of all property values minus the value of all mortgages and amounts owed as a result of equity release
- physical wealth is the sum of the values of household contents, collectibles and valuables, and vehicles (including personalised number plates)
- net financial wealth is the sum of the values of formal and informal financial assets, plus the value of certain assets held in the names of children, plus the value of endowments purchased to repay mortgages, less the value of non-mortgage debt.

Some points to note:

- informal financial assets exclude very small values (less than £250)
- money held in Trusts, other than Child Trust Funds, is not included
- financial liabilities are the sum of current account overdrafts plus amounts owed on credit cards, store cards, mail order, hire purchase and loans plus amounts owed in arrears
- private pension wealth is the sum of the value of current occupational pension wealth, retained rights in occupational pensions, current personal pension wealth, retained rights in personal pensions, Additional Voluntary Contributions (AVCs), value of pensions expected from former spouse or partner and value of pensions in payment. Note that, while net property wealth, physical wealth and net financial wealth are calculated simply by adding up the value of assets (minus liabilities, if applicable) for every household in the dataset, private pension wealth is more complicated because modelling is needed to calculate the value of current occupational pension wealth, retained rights in occupational pensions etc for each household. As with all models, the results depend on the assumptions made.

## **Private Pension Wealth Measures**

Nine separate components of private pension wealth were calculated based on the WAS survey

responses. There were four categories of pension to which respondents were making (or could have made) contributions to at the time of the survey:

- defined benefit (DB)
- additional voluntary contributions (AVCs) to DB schemes
- employer-provided defined contribution (DC)
- personal pensions.

The distinction between employer-provided DC pensions and personal pensions is as reported by the respondent. So, for example, if an individual had a Stakeholder Pension facilitated by their employer and chose to report that as an 'employer-provided/occupational scheme', this is counted as an employer-provided DC pension. Conversely, if an individual reported this simply as a 'Stakeholder Pension', it would be included in personal pensions.

In addition to these four categories of current pension scheme, wealth from five other types of pension was calculated:

- pensions already in receipt
- retained rights in DB-type schemes
- retained rights in DC-type schemes
- pension funds from which the individual is taking income drawdown
- pensions expected in future from a former spouse.

How the wealth for each of these components was calculated is described in detail in the following sections. Please note that the financial assumptions used to value defined benefit (DB), pensions in payment and pensions expected in future from a former spouse were modified at Round 8. More detail covering the rationale for the updated approach can be found in [Estimating defined benefit pension wealth in Great Britain](#) (Office for National Statistics (ONS), released 6 December 2024, ONS website, article)

#### *Current defined benefit occupational pension scheme wealth*

Individuals could report up to two current defined benefit pensions. The wealth in each of these schemes was calculated separately (as described below) and then summed to derive total wealth in current defined benefit (DB) occupational schemes.

Wealth in these schemes was defined as:

Where:

$$W_i = \frac{A_R Y_i^P + L_i}{(1+r)^{R-a}}$$

$A_R$  is the age- and sex-specific Superannuation Contributions Adjusted for Past Experience (SCAPE)-based annuity factor which account for a respondent reported age they will receive their pension<sup>[OBJ]</sup>, adjusted for the number of years they have until this retirement age,  $R$ , based on (joint life) annuity rates sourced from the Government Actuary's Department (GAD)<sup>[OBJ]</sup>. The discount rate<sup>[OBJ]</sup>, which is set at 2.4 per cent. The rate is taken as reported at the end of the fieldwork period.

Where:

- $Y_i^P$  is annual pension income, defined as  $Y_i^P = \alpha_i n_i s_i$
- $\alpha_i$  is the accrual fraction in the individual's scheme
- $n_i$  is the individual's tenure in the scheme
- $s_i$  is the individual's gross pay at the time of interview
- $L_i$  is the lump sum that the individual expects to receive at retirement
- $r$  is the real investment return (SCAPE Discount Rate),
- $R$  is the normal pension age in the pension scheme

- a is the individual's age at interview

*Definition of wealth from Additional Voluntary Contributions (AVCs)*

Individuals who reported being members of an occupational DB scheme were asked whether they had made any AVCs and, if so, what the value at the time of interview of their AVC fund was. Current AVC wealth is, therefore, simply defined as the fund value reported by the respondent at the time of the interview.

*Definition of current defined contribution occupational pension scheme wealth*

Individuals could report up to two current defined contribution pensions. The wealth in each of these schemes was calculated separately (as described below) and then summed to derive total wealth in current defined contribution (DC) occupational schemes. This procedure was also followed for those who reported that their employer-provided scheme was a hybrid scheme or that they did not know the type of scheme.

Individuals were asked to report the value of their fund at the time of the interview and were encouraged to consult recent statements where available. Current occupational DC pension wealth is, therefore, simply defined as the fund value reported by the respondent at the time of the interview.

*Definition of current personal pension wealth*

Individuals could report up to two current personal pensions in Wave 1, up to three current personal pensions from Wave 2 onwards; current being defined as schemes to which the individual was (or could have been) contributing at the time of interview. The wealth in each of these schemes was calculated separately (as described below) and then summed to derive total wealth in personal pensions.

Individuals were asked to report the value of their fund at the time of the interview and were encouraged to consult recent statements where available. Current personal pension wealth is, therefore, simply defined as the fund value reported by the respondent at the time of the interview.

*Retained rights in defined benefit occupational pension scheme*

Individuals could report up to three pensions (in Wave 1) in which rights have been retained. From Wave 2 onwards individuals could report up to six pensions in which rights have been retained. These could be either DB or DC schemes. The wealth in each DB retained scheme was calculated separately (in much the same way as for current DB schemes described above) and then summed to derive total wealth held as retained rights in defined benefit (DB) occupational schemes.

Wealth in these schemes was defined as:

Where:

$$W_i = \frac{A_R Y_i^P + L_i}{(1+r)^{R-a}}$$

$A_R$  is the age- and sex-specific annuity factor at retirement age matched based on their current age and the number of years to retirement age R (see above)

$Y_i^P$  is expected annual pension

$L_i$  is the lump sum that the individual expects to receive at retirement

r is the real investment return (SCAPE discount rate, see above).

R is assumed to be 66, or the individual's current age if he/she was already aged over 66

a is the individual's age at interview

*Retained rights in defined contribution occupational pension scheme*

The wealth in each DC retained scheme was calculated separately (in much the same way as for current DC schemes described above) and then summed to derive total wealth held as retained rights in DC

schemes. Specifically, individuals were asked to report the value (at the time of interview) of their retained DC fund.

*Retained rights in group personal, group stakeholder, private personal, private stakeholder, self-invested personal pension or retirement annuity contract*

The wealth in each of these schemes types was calculated separately. In these cases, individuals were asked to report what the fund value for their scheme was at the time of interview.

*Rights retained in schemes from which individuals are drawing down*

At the start of the round 7 questionnaire wording changes were made to questions that relate to the entry into private preserved pension wealth. These changes relate to how preserved pension wealth is divided amongst three main sections, drawdown, defined benefit and defined contribution. Consequently from round 7 onwards preserved pension wealth is separated into either defined benefit or defined contribution and whether the pension pot has been partially accessed, (eg. from taking a lump sum payment, but a portion of the pot remains for access at a later date) or whether the pension pot has not been accessed at all. Preserved pension entitlements in pensions for drawdown are therefore included within either defined benefit or defined contribution where the pot has been partially accessed and not shown as a separate category. Please reference the pension derived variable specifications for details on the derivation of these estimates.

*Pensions expected in future from former spouse/partner*

Individuals were asked to report in total how much they expected to receive in the future from private pensions from a former spouse or partner. Respondents were given the choice to report this either as a lump sum wealth figure, or as an expected annual income. Two slightly different approaches were followed, depending on how the respondent answered.

For those who reported a total lump sum value, this figure was taken as the relevant wealth measure and discounted back to the time of the interview. For those who reported an expected future annual income, wealth was calculated in much the same way as for DB schemes described above:

Where:

$$W = \frac{A_R Y^P}{(1+r)^{R-a}}$$

- $A_R$  based on their current age and the number of years to retirement age, R (see above)
- $Y^P$  is expected annual pension
- r is the real investment return (SCAPE-based rate as set at the end of the interview)
- R is assumed to be 66, or the individual's current age if he/she was already aged over 66

a is the individual's age at interview. Please note for simplicity annuity rates used to model DB wealth were also used for pensions expected in future from a former spouse/partner. These rates include a joint life assumption, which would be unlikely to be present for this type of pension. This will lead to a slight over-valuation of this type of pension, however, given the low prevalence of this type of pension the impact is likely to be small on overall pensions and total wealth measures.

*Definition of wealth from pensions in payment (employer and personal pensions)*

The SCAPE based annuity factor, A, is also used to convert a pensions in payment income stream into a wealth value. A joint life annuity factor is applied to reported pension in payment incomes from employer or personal pension schemes. Wealth from these types of pensions in payment was therefore defined as:

$$W = A_a Y^P$$

Where:

- $A_R$  is the age- and sex-specific joint life annuity factor based on respondent's current age.  $a$
- $Y_P$  is reported current annual private pension income

*Definition of wealth from pensions in payment (from spouse's previous employer)*

The SCAPE based annuity factor,  $A$ , is also used to convert a pensions in payment income stream into a wealth value. A single life annuity factor is applied to the pension in payment income from spouse's previous employer. Wealth from these pensions in payment was therefore defined as:

$$W = A_a Y^P$$

Where:

- $A_R$  is the age- and sex-specific single life annuity factor based on respondent's current age,  $a$
- $Y_P$  is reported current annual private pension income

## Contact details

For further information, or to provide feedback on the Wealth and Assets Survey documentation, please email [Wealth.and.Assets.Survey@ons.gov.uk](mailto:Wealth.and.Assets.Survey@ons.gov.uk).

# **Annex: Impact of COVID**

## ***Background***

Data collection for WAS, and other social surveys undertaken by ONS, was suspended on the evening of Monday 17<sup>th</sup> March 2020 in response to the COVID-19 pandemic and Government issued social distancing guidance.

With fieldwork suspended midway through the final month of Round 7 (April 2018-March 2020), the impact upon WAS was greatest for Round 8.

The initially planned strategic response of ONS to the pandemic can be found on the [ONS website](#).

Seeking to ensure the country had the best possible information, while safeguarding the wellbeing of respondents and interviews, end-to-end data collection and processing methodologies were revised to permit WAS and other social surveys to continue.

## ***Incremental Changes***

Responding to the rapidly changing circumstances of the pandemic, changes were incrementally introduced, and withdrawn, over the 2 years of the Round 8 data collection cycle.

## ***Initial Approach***

Continuing the survey in telephone mode using face-to-face interviewers prompted change across the end-to-end data collection process to:

- modify the questionnaire design to make more usable over the phone
- revise the sample design
- stand-up new methods of making contact with Wave 1 households
- update advance materials and interviewer briefings
- update interviewer training.

## ***Questionnaire Changes***

Questionnaire design changes to make it more usable over the phone saw:

- questions that would have used showcards reworded
- content removed.

## ***Modal Questionnaire Changes***

The pre-pandemic face-to-face questionnaire originally prepared for use from April 2020 had many questions that made use of showcards.

To reduce interviewer and respondent burden more than 200 questions were reworded for April.

In addition to these amendments three standard prompts were introduced to advise on how these questions should be asked:

- PROMPT AS NECESSARY Allow respondent time to answer, then prompt with responses if required
- RUNNING PROMPT Pause after reading out all coding options with no pause between each option
- INDIVIDUAL PROMPT Pause after each coding option to give respondent opportunity to answer.

## ***Question Content***

With the objective of reducing the length of the WAS questionnaire from 75-80 minutes to 40-45 minutes, while seeking to retain questions on wealth and core variables required for the calculation

of income, the removal of content was staggered.

Most pandemic prompted changes were made in April and May with further pandemic and maintenance driven changes to questionnaire content made later in the Round.

From April 2020 all, or subsets, of the following question blocks were removed:

- demography:
  - country of birth
  - health
  - illness
  - religion
- asset transfers:
  - inheritances
  - gifts
  - loans
- subjective questions covering:
  - attitudes to risk
  - attitudes to saving for retirement
  - debt burden
  - financial acuity
  - financial situation/expectations
  - household bills
  - household budget/current financial planning
  - savings attitude and behaviour
  - wellbeing

From May 2020 all, or subsets, of the following question blocks were removed:

- housing tenure & type of accommodation
- equity release
- business assets
- time in current job
- deductions and refunds on employment income for main job
- usual pay, why not usual and deductions from usual pay
- company vehicles
- retirement plans
- use of past pensions
- State Pensions
- auto enrolment
- Pension Wise
- pension sources
- financial situation\expectations
- trusts
- beneficiaries
- status of parents

May also saw the introduction of a block of questions to assess the impact upon income of the pandemic.

From September 2020:

- the Worth and Happy Wellbeing questions were removed
- additional interviewer guidance was added to one of the mail order questions
- changes were made to the questions covering the payment of incentives

From October questions on the taking of lump sum payments from employer and personal pensions were replaced by a series of questions on Pensions Freedoms. Further amendments to the Pensions Freedoms questions were made in January and February 2021.

From January 2021 changes made to value of past pensions question to revert to wording of Round 6.

From February 2021:

- further interviewer guidance was added to value of outstanding mortgage question
- lifetime inheritance questions drafted for the Round 8 questionnaire, originally prepared, were re-instated.

From April 2021 changes were made to:

- update the disability benefit questions to directly reference Child Disability Payments in Scotland
- update the Social Fund questions to directly reference payments made from the Scottish Welfare Fund
- add further options to the block of questions to assess the impact upon income of the pandemic first introduced in May 2020
- add a block of questions to assess the impact upon savings of the pandemic.

From May 2021 further amendments were made to the Pensions Freedoms questions introduced from October 2020.

## ***Sample Design***

From June 2020 changes were made to the issued Wave 1 sample with quota sizes doubled from 13 to 26 to compensate for the increased difficulty interviewers faced in seeking to make contact with households with the limited contact details provided by the Royal Mail Postcode Address File.

Issued Wave 1 quota sizes were then progressively reduced in size over the first year of the Round as new approaches to making contacts were introduced.

WAS Wave 1 quota sizes:

- 13, April & May
- 26, June to September
- 24, October to January
- 20, February & March

Following a reduction in funding for the survey for the second year of the Round, Wave 1 data collection for WAS suspended until the end of the Round.

From June 2020 unconditional incentives offered to Wave 1 households were increased from £5 to £10; with no measurable impact on response rates unconditional incentives reverted to £5 from October.

To mitigate for the increased difficulties interviewers faced in making contact with sampled households over the phone, field sampling procedures previously applied when concealed multi-household addresses were discovered were relaxed to permit the first identified address to be selected when the interviewer was unable to follow the KISH grid procedure over the phone.

## ***New Methods of Making Contact***

Initial changes to how sampled Wave 1 households were first contacted included:

- the introduction of compliments slips with business-as-usual materials to inform respondents of changes to data collection
- extending the existing use of telematching

- inviting sampled households to provide their telephone number via an [online portal](#), the ONS Survey Enquiry Line or directly to interviewers sending personalised letters.

In July 2020 ONS began a series of Knock to Nudge (KtN) trials, in which interviewers would visit households to “nudge” those selected to participate when they had not:

- provide their contact details via the online portal
- had their telephone details captured via tele-matching
- contacted ONS or an interviewer directly.

KtN provided an opportunity for potential respondents to schedule a telephone appointment with the interviewer at the doorstep, and for the interviewer to build rapport. For non-contacts, a call today card, explaining that an interviewer had visited the address, was posted through the door.

Following the trials KtN was initially scaled up for use with the Living Costs and Food survey and Survey on Living Conditions from October 2020; from January 2021 this approach was also adopted for WAS.

### *Advance Materials and Interviewer Briefing*

To support field interviewers in the transition from face-to-face to telephone interviewing changes were also made to advance letters, the materials interviewers personally used to engage with respondents and the briefings interviewers received.

Briefings provided to interviewers were extended to cover the changes made. In summary:

- generic briefings covering the switch from face-to-face to telephone interviewing were cascaded through Survey Champions
- WAS specific briefings on COVID-19, and previously agreed questionnaire changes, were also cascaded through Survey Champions
- generic and survey specific briefings were loaded to the ONS Learning Hub, accessible to all interviewers.

Alongside the generic briefing guidance on how to obtain and conduct a telephone interview, WAS interviewers were also asked to:

- maintain as much person level interviewing as possible i.e. minimise the use of proxies
- consider splitting the interview into two or more calls if they felt it would help to secure a full or partial interview
- consider splitting by respondent or pairs of respondents rather than subject matter as the household section represents a relatively short proportion of each WAS interview compared to the individual interview.

### **Impact Upon Analysis**

Findings from Round 8 of the data need to be treated with a degree of caution due to changes in data collection approach in response to the coronavirus (COVID-19) pandemic, as outlined in [Impact of COVID-19 on ONS social survey data collection](#), along with the shifting patterns in economic and social behaviour during the period. In particular, steps were taken to account for substantial bias in tenure introduced within the WAS achieved sample because of the change in survey mode of collection, with an observed increase in the proportion of households who owned their property outright, and a decrease in the proportion of those who rented their property. As such, latest tenure population totals were added as calibration control within the survey weights.

London is subject to additional non-response bias, despite having additional tenure control. This is because there is a higher proportion of renters in London (over 50%) than in other regions (for example, England is 38%). Also, in London there is a greater degree of under coverage of renters in the survey sample. As a result, there is increased uncertainty attached to wealth estimates for London, and we advise users to apply caution when drawing direct comparisons with other

regions.

Additionally, for Round 8 income estimates do not account for COVID-19 financial support schemes and reflect value reported at time of interview, rather than usual pay. Users should apply a degree of caution when interpreting due to additional uncertainty.