

ENGLISH HOUSING SURVEY

Household Dataset User Guide

2009/10

Contents

The English Housing Survey (EHS)

1. Introduction to the EHS Survey

- Reporting
- EHS datasets
- Overview of the EHS Household dataset
- EHS data base
- Other sources of information

2. Using the data base - first steps

- How the Household data has been produced
- Sample structure
- Datasets at the UK Data Archive (UKDA)
- Levels of data
- Types of data

3. Key Concepts and Definitions¹

- Household
- Household Reference Person
- Un-clustered Sampling

4. Locating information in the database

- General Overview
- Interview (household) survey variables
- Rotating Modules
- Multi-coded variables

5. How to match files

6. How to apply grossing factors to obtain national estimates

7. How to deal with missing data

- Missing Data Conventions
- Dealing with missing data in analysis
- Checking if results are significant

¹ A full glossary of terms used in the EHS can be found on the DCLG web site – <http://www.communities.gov.uk/housing/housingresearch/housingsurveys/englishhousingsurvey/ehsusers/>

Appendix A: UK Data Archive folder structure
Appendix B: Content and level of Interview Survey Files
Appendix C: Content of Derived Files

The English Housing Survey (EHS)

1. Introduction to the EHS

1. The EHS is a national survey commissioned by the Department for Communities and Local Government that collects information about people's housing circumstances and the condition and energy efficiency of housing in England. The survey is run on a continuous basis as part of the wider Office for National Statistics (ONS) Integrated Household Survey (IHS).
2. The survey has a complex multi-stage methodology consisting of 3 main elements: an initial interview survey of around 17,000 households with a follow up physical inspection and a desk based market valuation of a sub-sample of approximately 8,000 of these dwellings, including vacant dwellings. The interview survey sample forms part of the Integrated Household Survey (IHS), and the core questions from the IHS form part of the EHS questionnaire. More information about the IHS is available from its webpage:

<http://www.statistics.gov.uk/CCI/nugget.asp?ID=936&Pos=1&ColRank=1&Rank=224>

3. The link below provides a range of information that will support anyone wishing to make use of the survey data to conduct their own analysis or wishing to understand in detail how the survey is run and managed. All users are encouraged to familiarise themselves with the technical background before undertaking any detailed work using the EHS results.

<http://www.communities.gov.uk/housing/housingresearch/housingsurveys/englishhousingsurvey/ehsusers/>

4. The English Housing Survey methodology comprises the following
 - An initial interview with approximately 17,000 households a year. This is referred to as the 'full interview sample'. This dataset is made available on a single financial year basis as the EHS Household data set.
 - A follow up physical inspection of around 8,000 of these respondents' homes to assess the condition and energy performance of the property. An independent market valuation is also undertaken. This is referred to as the 'dwelling sample'. This dataset is made available on a rolling 2-year basis. (The 'dwelling sample' comprises all cases where an occupied dwelling has an interview, physical and market value survey completed plus vacant dwellings with a physical and market value survey only).

Reporting

5. ***EHS Headline Report.*** This short report is published annually and presents preliminary headline findings including key indicators related to departmental housing policies in areas such as trends in tenure and household composition, overcrowding, housing costs, and the condition and energy performance of the stock. Analysis is at a national level from both the full and dwelling samples. This is followed by two separate annual reports.

Annual Household Report

6. Available in the summer/autumn each year, starting with the 2008/09 report published in Autumn 2010.
7. The EHS report is based on the full interview sample only and presents comprehensive analysis of housing trends across each of the sectors and for different household groups including changing tenure patterns, over crowding and under occupation, second homes, rents and mortgages. Both national and regional level findings are presented where possible. New topics will also be introduced on a rolling basis reflecting changing policy interests.
8. The reports are available to download from the web site. The reports include an overview of the survey methodology with guidance about sampling errors and a glossary of terms.

<http://www.communities.gov.uk/housing/housingresearch/housingsurveys/englishhousingsurvey/ehspublications>

Annual Housing Stock Report

9. Also available in the summer/autumn each year, (starting with the '2008' report published autumn 2010) based on the rolling two year dwelling sample i.e. cases where a physical inspection has been undertaken. The results presented for '2009' cover the fieldwork period April 2008 to March 2010.
10. The overarching focus will remain on housing conditions and energy performance but with other topics introduced each year to reflect emerging policy priorities.
11. Again changes to the interview survey content may mean not all topics from former EHCS reports can be covered in full but the overall level of continuity with the former EHCS is high.

The 2009 dwelling stock report can be downloaded from:

<http://www.communities.gov.uk/housing/housingresearch/housingsurveys/englishhousingsurvey/ehspublications>

EHS Data Sets

12. The public datasets that will be available to external users through the UK Data Archive reflect the above EHS reporting strategy. Two separate data sets are therefore available as follows:

a) **EHS Household Data Set** – as used for the EHS Household report. The data set comprises the full interview data (plus associated derived variables) for all cases where an interview has been completed – approximately 17k cases pa. Data sets are provided for single financial years together with annual weights. This data set should be used for any analysis where only information from the household interview is required.

b) **EHS Housing Stock Data Set** – as used for the EHS Housing Stock report. The data set is available for all cases where a physical survey has been completed and for occupied cases comprises data from the household interview as well as data from the physical survey and market value survey. For vacant properties only data for the physical survey and market value survey is provided. The data is made available for a two year rolling sample i.e. approximately 16k cases together with the appropriate 2-year weights. For example, the EHS Housing Stock Data Set is for '2009' covering the period April 2008 to March 2010. The Housing Stock Data Set should be used for any analysis requiring information relating to the physical characteristics and energy efficiency of the housing stock.

13. **This User Guide relates solely to the EHS Household Data Set.** Users who wish to access the EHS Housing Stock Data Set should refer to the separate EHS Housing Stock Data User Guide that accompanies that data set. DCLG have deposited the data in this way to provide greater continuity for previous users of the SEH and EHCS and so that discrete data sets can be made available in the most straightforward way.

14. New data files have been created, some have been dropped and a number of new variables have been introduced. This applies particularly to the primary Interview files and users will need to be particularly careful when using these files. Further information is available in the Technical Advice Notes on the EHS website.

<http://www.communities.gov.uk/housing/housingresearch/housingsurveys/englishhousingsurvey/ehsusers/technicalbackground>

Overview of the EHS Household Data Set

15. The EHS Household data is available in SPSS format through the UKDA as a download. To download or access the data you will need to register

with the archive, though it is possible to download descriptions of data and a general overview of the survey without registering (see link below). The data files and documentation are held in separate folder structures, following the archive guidelines. See Appendix A for an overview.

<http://www.esds.ac.uk/Lucene/Search.aspx>

EHS Database

The information provided on the EHS Household Database, (starting with the 2008/09 data set) comprises

- a) The full set of primary and derived files including weights
- b) Documentation and reports
- c) Data dictionaries for all files

16. **Primary Data** - The EHS Household data set consists of primary data files containing basic survey data collected by the interviewers.

17. **Derived Data** - a small number of files containing 'derived' variables are also included on the database holding variables created by calculating with or recoding the basic survey data. These include for example household composition and the bedroom standard. Weighting factors are also provided. These files will be the most widely used for analysis. (See Appendix C).

There are two derived files available for the EHS household data set, which contain the most commonly used derived variables. These are

- Generalfsxx.sav
- Interviewfsxx.sav

Where xx is the survey year. These include the key indicators summarised below:

Generalfsxx.sav	Tenure, nature of area grossing factors;
Interviewfsxx.sav	Household composition, length of residency/ownership; income; rent, mortgages

There is a further detailed derived file which includes information on equivalised income. Further information on content of the two main derived variable files is in Appendix C.

Users are encouraged wherever possible to make use of these two derived files for their analysis.

18. **Grossing weights** – these have been produced for all household sample cases in order to produce accurate national estimates. See Section 6 below for advice on applying these weights. The weights are held on the derived file `generalfsxx.sav`.
19. **Documentation and Reports** - A number of documents are included on the EHS Household database to help users understand how the database is organised and where to find the data they need. These are found in the folder [\\mrdoc](#). See Appendix A.
20. **Interview survey documentation** - this document shows all the questions asked of households, the variable names and available response categories and showcards. Some new questions have been introduced and others have been dropped. The interview questionnaire provided indicates the new interview content. Certain topics are also only collected on a rotating basis (section 4).
21. **Derived variable documentation** - See Appendix A. These are excel spreadsheets which give detailed information on the two main derived variable files, `Interviewfsxx.sav` and `Generalfsxx.sav`
22. **Data dictionaries** - These are detailed descriptions of variables held in a file and generated by SPSS and can be found in `\\mrdoc\allissue\xxxx`, where `xxxx` is the name of the folder holding a data dictionary for the corresponding set of files.
23. **EHS Reports** - annual reports from EHS are available in the EHS dataset at the UKDA and also available on the DCLG web site.

<http://www.communities.gov.uk/housing/housingresearch/housingsurveys/englishhousingsurvey/ehspublications/>

Other sources of information

24. Further information about the way in which the survey is organised is available on the EHS web site.

<http://www.communities.gov.uk/housing/housingresearch/housingsurveys/englishhousingsurvey/>

Below are links to specific documents that may be of interest.

25. **EHS Technical Advice notes** –These provide details of how the fieldwork is organised, how the sample is drawn and the results grossed each year, estimates of sampling and measurement error together with a full glossary of terms and details of how derived measures such as energy efficiency and income are created.

Detailed information on how derived variables are produced and how they can be used are given in the Advice notes as follows:

- Survey Methodology and Data Collection
- Sample structure and weighting
- Data quality
- Creating derived measures for EHS analysis – household related measures
- EHS Glossary

<http://www.communities.gov.uk/housing/housingresearch/housingsurveys/englishhousingsurvey/ehsusers>

26. **EHS Supporting Tables** – an updated set of tables are produced each year and made available on the web site.

<http://www.communities.gov.uk/housing/housingresearch/housingsurveys/englishhousingsurvey/ehstables/>

2. Using the data base - first steps

27. To use the data set correctly you must first ensure that you familiarise yourself with and be aware of the following:

- How the dataset has been produced
- What information is published
- The key concepts and definitions used in the survey
- How to locate the information you want in the database (find both survey and computed data by variable and file names)
- How to match files
- How to apply grossing factors to obtain national estimates
- How to deal with missing data
- How to deal with sampling and survey error

Each of these is described briefly in the sections below but please note that this guide is not intended to be a comprehensive source of information. Reference is made to the more detailed explanations and guidance contained in other files, publications and documents which are either at the UKDA or at the DCLG web site.

The guide assumes knowledge and experience of using SPSS.

How the Household dataset has been produced

28. Primary data is split into a number of topic based files. There are two derived files which contain key variables used in analysis. (See Appendices B and C). All files can be matched as required using the case identifier aacode.

Sample structure

Annual Sample Size	32,100
Frequency of fieldwork	2 months per quarter
Number of months of interview fieldwork	8
Coverage	England only
Clustering	No
Design Description	Single stage sample of addresses
Achieved response rate	59%
Household interviews achieved	17,042
Paired cases achieved	8,178

29. The sample for the EHS is drawn at random from a list of all addresses held by the Royal Mail (known as the Postcode Address File - PAF). The Household Data Set is composed of the 17,042 addresses where a full household interview has been achieved. This design means that addresses are sampled directly from the entire Postcode Address File (PAF) rather than being selected for example, from a sub-group of postcode sectors. An un-clustered sample helps to produce more precise results without increasing the sample size. This is because people with the same characteristics are often geographically clustered. Therefore, by increasing the 'spread' of addresses sampled, this clustering effect is reduced.

30. The principal sampling methodology of the EHS :

- The EHS uses an un-clustered sample. This enables a smaller sample to be used with no loss of precision, i.e. without sampling errors being increased. The more scattered sample does, however, have some implications for fieldwork organisation.
- The slightly smaller sample achieved in the EHS will give more robust estimates for many measures from the household sample.
- The EHS selects one dwelling per address and one household per dwelling and interviews only the household reference person (HRP) of that household.

Full details of sample methodology can be found below

<http://www.communities.gov.uk/publications/housing/ehstechnicaladvice/notes>

Datasets at the UKDA

31. The EHS data has been accorded National Statistics (NS) status. The EHS Household data set comprises data collected over a single financial year (April 2009-March 2010).

The Household Sample is the set of cases from the Interview survey only.

The number of (un-weighted) cases in the Household Data set is as follows:

Household Dataset	Un-weighted Cases, households	Weighted cases, Households (thousands)
EHS 2008/09	17,691	21,530
EHS 2009/10	17,042	21,554

Levels of data

32. Most of the data is at household level. However, some of the primary data files contain data at sub-household level. These are indicated for the Interview Survey at Appendix B. All the files containing derived variables are at household level.

Types of data

33. **Primary Survey data-** primary data collected in the field is often very specific and detailed and in many cases is only collected in order to provide the building blocks for computing more useful pieces of information - in particular the household composition and income measures used in the analysis of the survey. There will also be a small amount of missing data in the primary data collected in the field.

In many cases omissions or inconsistencies in the base variables will have been identified in the course of producing derived variables for final analysis. ***Users are recommended wherever possible to make use of the final derived variables rather than the raw data from the primary files as these are likely to contain more complete information on key topics such as rent, mortgage payments and income.***

34. **Derived/computed data-** While the information that is collected in the field provides all the base variables needed, a considerable amount of further processing and modelling is needed to fully validate the results and produce secondary variables needed for detailed analysis. Producing derived variables often involves reconciliation of conflicting information from different parts of the survey (e.g. tenure) and/or the imputation of

missing data where this has been possible from other data collected in the survey.

35. It is important to note that changes introduced to some of the EHS modules will have a significant impact on any time series analysis attempted and users should familiarise themselves with relevant parts of the Technical Report before drawing conclusions about changes over time. Revised variable names have been used wherever there has been a major break in methodology year on year to minimise the risk of data being mis-interpreted.
36. **Missing data** - as with all surveys there will be some level of missing data on the primary data files where information was not available or for example, a respondent refused to answer specific questions. Most of the derived variables for the dwelling sample however have no missing data, as missing values have been imputed based on cases with similar characteristics. Where imputation has taken place there are accompanying flag variables to indicate the level of imputation. Imputation is not always achievable so for some of the key derived variables there will be missing data which you may wish to distribute proportionately using the tool we have provided – see Section 7. The method of data imputation is described more fully in the EHS Technical Advice Notes.

<http://www.communities.gov.uk/housing/housingresearch/housingsurveys/englishhousingsurvey/ehsusers>

Guidance on dealing with missing data in analysis is provided in Section 7 below.

37. Alongside the annual data files, the UK Data Archive holds supporting documentation such as questionnaires, specifications for derived variables and show cards.

It should be noted that several variables have been top coded on this dataset. Data has also been cleaned of all identifiable variables to maintain the confidentiality of respondents.

3. Key Concepts and Definitions²

38. The Household dataset relates to households. It is therefore important to understand how this and related terms are defined and applied in the survey.

Household

² A full glossary of terms used in the EHS can be found in the EHS Technical Advice notes and in the Reports

A Household is defined as one person or a group of people, who have the accommodation as their only or main residence and (for a group) either share at least one meal a day or share the living accommodation, that is, a living or sitting room.

Household Reference Person

The Household Reference Person is defined as the “householder” (that is the person in whose name the accommodation is owned or rented). This term replaced the concept of “head of household” which was used in SEH datasets prior to April 2001. For joint householders (joint owner or joint tenants), the household reference person is whoever has the highest income. If incomes are the same, the older person is defined as the household reference person. Thus the household reference person definition, unlike the old head of household definition, no longer gives automatic priority to male partners.

4. Locating information in the database

General Overview

39. Appendix A provides a diagram showing the overall structure of the EHS Household database held at the UKDA. Essentially files are held under three main folders comprising

- Interview
- Derived
- Documentation

Interview (household) survey variables

40. The topics covered in the Interview Survey questionnaire are illustrated in Figure 1 below. The core questions from the Integrated Household Survey (IHS) form part of the EHS questionnaire. It should be noted that the majority of questions on the IHS core cover sections which existed under the SEH. This includes modules on

- Household Demographics
- Household Reference Person and Relationships in household
- Accommodation
- Tenure
- National Identity/ Ethnicity
- Time at current address
- Health and Disability
- Job Details

In bringing together the EHS interview modules, the underlying principles were to retain all the key topics from the former SEH and EHCS but where necessary to trim back the number of questions being asked. In some areas whole topics have been transferred across with fairly minimal change. For former SEH users this applies in particular to the modules on

- Housing history
- Housing aspirations
- Tenancy Deposits
- Home and Neighbourhood
- Leases
- Second homes
- Mortgages and Rents
- Economic status

41. Other topics have been introduced from the former EHCS – these include

- Damp in the home
- Work done to the home
- Adaptations for disabled people
- Council tax and utility bills

42. Other sections have been extensively modified – in particular improvements have been introduced to how information on mortgages is now collected.

Please note that wherever appropriate variable names have remained unchanged from the former SEH or EHCS i.e. when a question has been taken directly from one of the former surveys with the same routing, question wording and response categories the variable name will be unchanged. Where there has been a small change e.g. a change to a response category or a slight change to the routing then a number will have been added to the variable name to indicate this is a variant of the original variable e.g. Vnllrd2 – (From whom did the household rent the property?), previously this was Vnllrd.

It is also important to note that the survey no longer collects information on individual tenancy groups as under the former SEH.

Rotating Modules

43. Because of pressure on the size of the interview survey, the EHS also uses the concept of 'rotating modules' to bring questions in and out of the survey to meet users' demands and policy needs. The pattern of rotation will vary with some topics such as Fire being included normally every other year starting in 08/09 while others will be retained for a two year in/two years out basis depending on policy interest. Full details are given on the Interview survey questionnaire.

44. The precise set of files and data provided each year will therefore vary reflecting the changing content of the interview survey

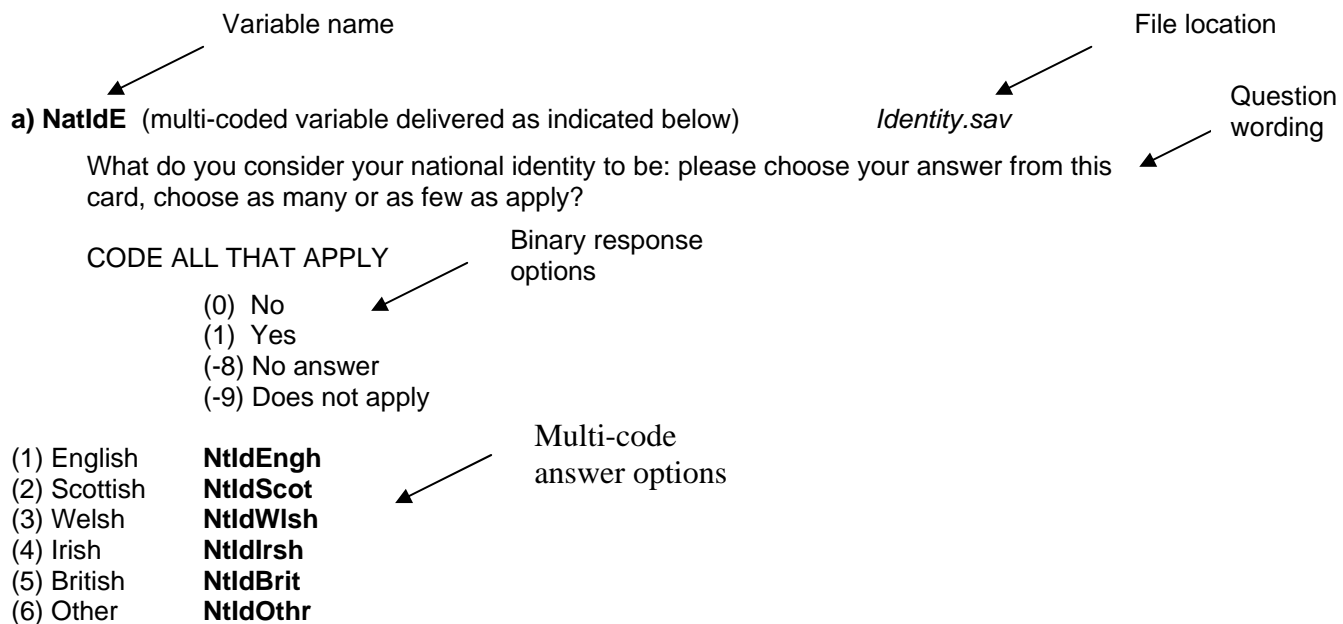
Details of the annual changes are available on the web site.

<http://www.communities.gov.uk/publications/housing/ehssurveychanges>

Multi-coded variables

45. The EHS contains a number of multi-coded questions. These are questions where households could give more than one answer; for example 'national identity', 'adaptations for disability' or 'source of funding for buying a property'.

In the EHS household dataset, all multi-coded questions have been stored as a series of binary variables, with a separate variable for each answer option. It is expected that users may find this approach facilitates easier analysis. See examples below from the questionnaire documentation:



b) **HAS443b** (multi-coded variable delivered as indicated below) *Adapt.sav*

You have told me that there are some adaptations that you need that have not been made. Can you tell me why these modifications have not been made?

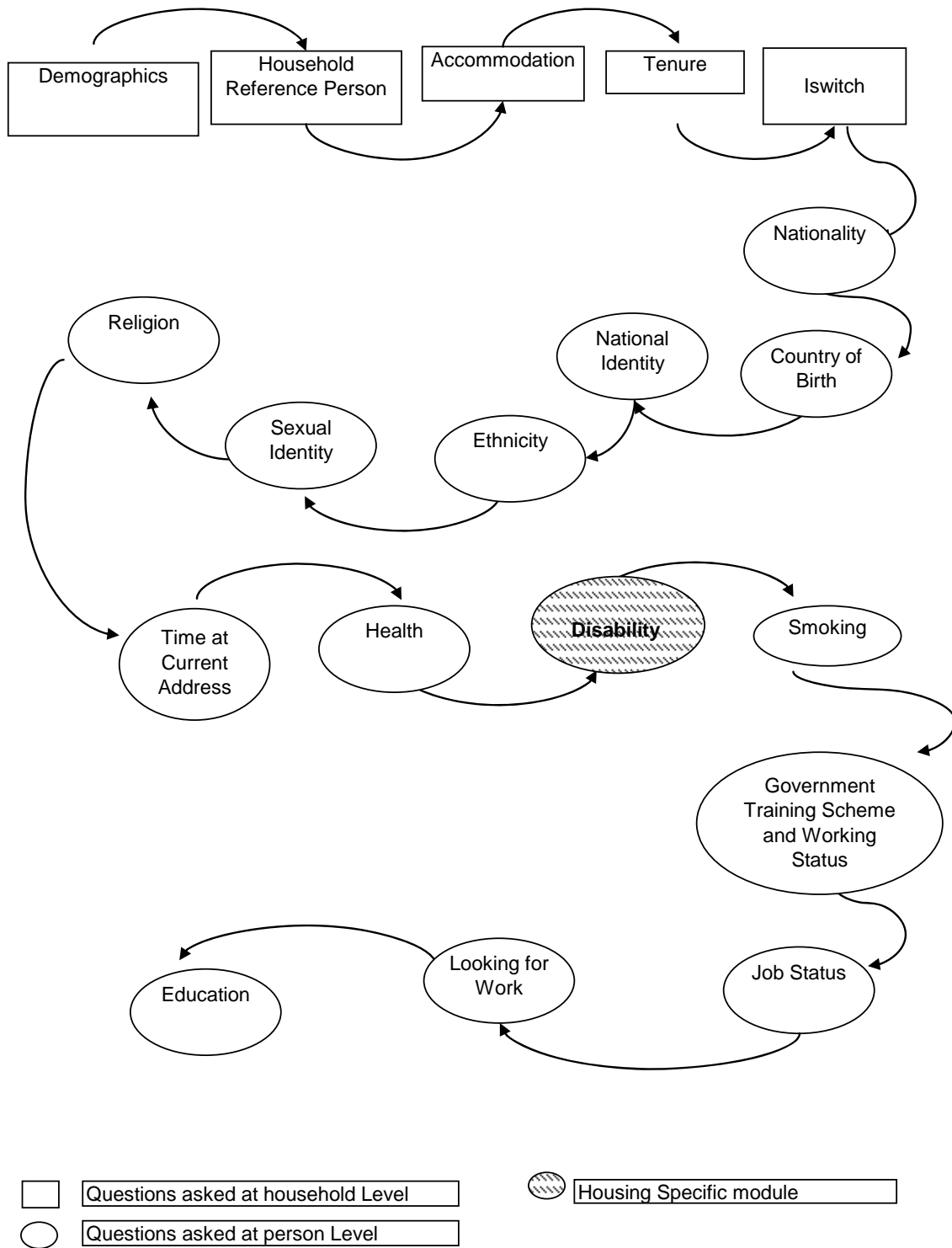
CODE ALL THAT APPLY

- (0) No
- (1) Yes
- (-8) No answer
- (-9) Does not apply

- | | |
|----------------------------------------------------------------|-----------------|
| (1) Expect the modifications will be made, not enough time yet | HASbExpt |
| (2) Modifications not worth doing | HASbNWth |
| (3) Wouldn't know how to get something done | HASbNtKw |
| (4) Would cost more than can afford | HASbCost |
| (5) Doesn't trust builders | HASbTrst |
| (6) Can't/wouldn't get a grant | HASbGrnt |
| (7) Landlord won't allow it | HASbLdAw |
| (8) Landlord won't pay | HASbLdPy |
| (9) Other reason for no modification | HASbOthr |

FIG. 1

Questionnaire Structure - EHS 2009/10: Core Questions



5. How to match files

Matching files that exist at the household level

46. The key variable for matching at the household level within or across any of the survey data is **aacode** (an eight character string variable). You must specify this as the key field when matching. All files on the database have been sorted by this variable so should match. 'Sort file' or 'split file' commands, however, re-sort the file. Matching with standard database files will fail if the data is not first re-sorted by aacode.

Matching lower level files (i.e. below household level) up to the grossing file

47. Files at a lower level of organisation may have more than one case per household. Some of them have exactly the same number of cases per address e.g. the damp.sav file has 6 cases for each address each representing a room for where the question applies. With other files, the number of cases varies e.g. for people.sav the number of cases is the number of people in the household so one person households will just have one case, 6 person households will have 6 cases etc. As there may be more than one case per address, these cannot simply be matched into the household level files to give results at the household level. There are 2 ways to approach these files:

- Match the household file in as 'tables'. Here the information at household level will be copied for each person etc. within that aacode. The results will then represent the number of people, number of types of window etc. rather than the number of households. This can be useful for very specific analyses, e.g. on numbers of people sharing amenities.
- Aggregate the lower level file up to address code level and then match it in with the household file. (Most key variables for analysis have been derived at the household level so this will not normally be necessary).

6. How to apply grossing factors to obtain national estimates

48. The EHS Household dataset is based on a stratified sample with over-sampling of the rented tenures to achieve large enough sample sizes to produce reliable results for the social and private rented sectors.

Grossing factors have been calculated to

- Compensate for the design of the sample i.e. the over sampling of some dwellings and under sampling of others; and
- Take account of non-response bias: the survey response rates achieved for different groups of households and dwellings;

Within SPSS, results must be based on data weighted by the relevant grossing factor to produce national estimates.

49. The EHS Household Data set comes with its own grossing factors covering the household sample of cases (see above). The grossing factor can be found in the file **generalfsxx.sav** in the 'derived' folder. The relevant variables are:

- Household grossing **aagfhxx**

Where xx indicates the year during which fieldwork was conducted. For example, **Aagfh09** is the household grossing factor for the 2009-2010 household data i.e. for the EHS 2009 household data set.

- **Aagfhxx** should be used for any analysis in which the aim is to provide estimates of households that is based on the interview survey data (e.g. percentage of households below the bedroom standard).

The grossing factor can only be used on each full year household data sets. It cannot be used on the dwelling stock data even when this has been split into separate years.

7. How to deal with missing data

Missing Data Conventions

50. The files contain no system missing values. Where values are unknown or the question was not asked, specific codes are used and the values are set to user defined missing. The conventions adopted are:

Does not apply -9 for single digit fields, -99 for 2-digit fields etc.

Unknown -8 for single digit fields, -88 for 2-digit fields etc.

Not asked for a particular Quarter -2 for single digit fields, -22 for 2 digit fields etc

No partner -7 for single digit fields, -77 for 2-digit fields etc.

51. Users need to be aware of the implications of having missing values switched on or off within SPSS. It is useful for example to leave missing values switched on (the SPSS standard) if users wish to establish the level

of missing data in a frequency or cross tab. The SPSS output Case Processing Summary will show the level of missing data. This can be used to decide on whether the level of missing data is significant and whether it is appropriate to re-distribute missing cases.

52. It is particularly important to leave missing values switched on when dealing with a continuous variable e.g. costs to make decent. Including -99, -88 etc as a normal value (i.e. not as a missing value) would distort any outputs.

When you create any new variables, you should declare missing values as below:

```
Missing values var1 (-9).
```

This code sets up '-9' as a user defined missing so will be excluded from any cross tabs or statistics of var1.

Dealing with missing data in analysis

Analysis that produces national estimates

53. Analysis that uses derived variables with no missing cases will have no problem with missing data. However, when undertaking analysis using other variables that have missing or unknown codes, these cases can be redistributed to produce national estimates.
54. Some procedures like 'Cross tabulation' exclude missing cases from the table. To include these as the basis for any redistribution it is necessary to switch off the user defined missing values for each variable used. This is done by

Missing values Var1 (). (Missing data is included in outputs)

55. Missing values can then be switched back on using:

Missing values var1 (-9). Etc

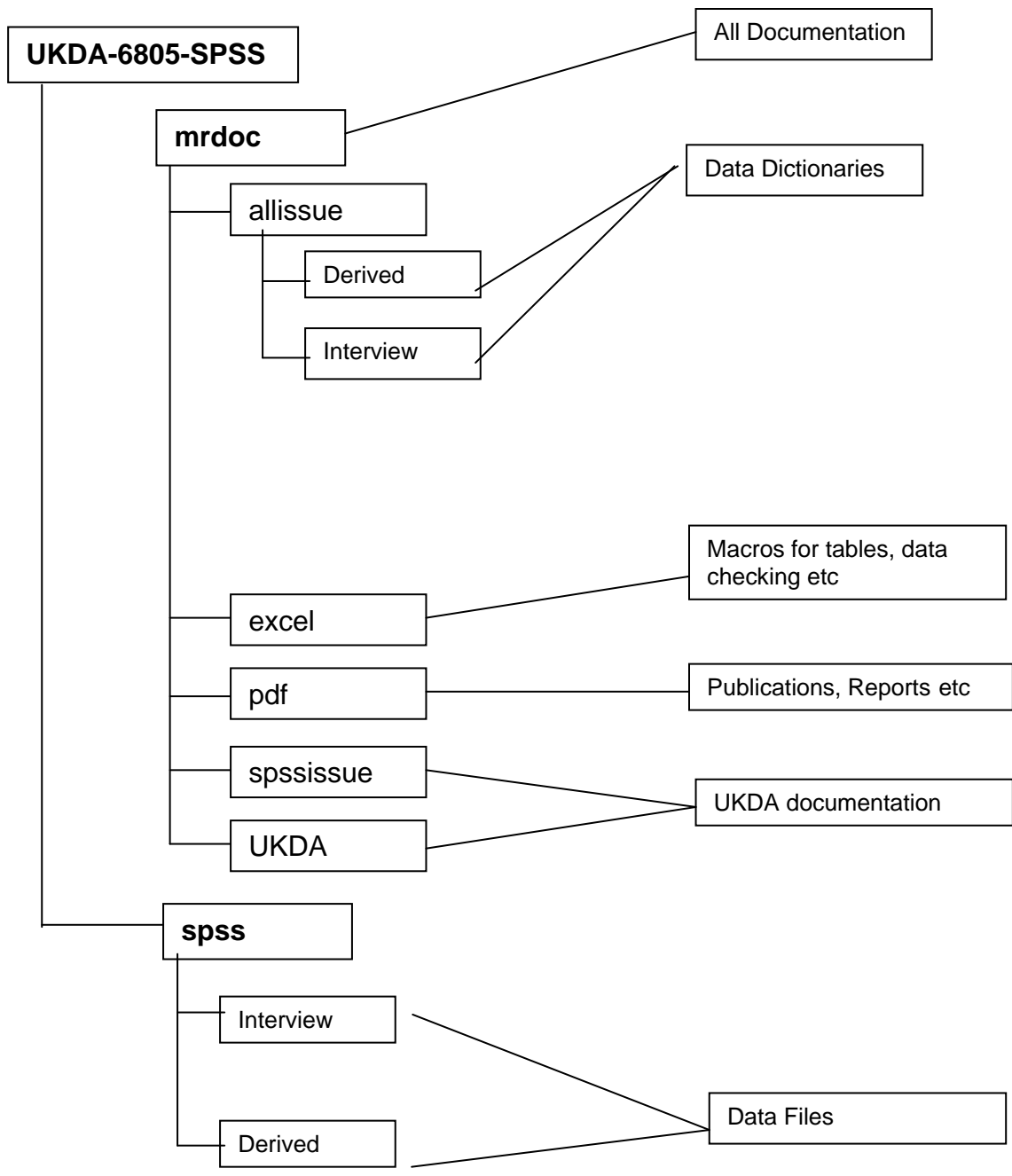
56. Where you have missing values within the table, unless there is a clear reason to do otherwise, these need to be allocated pro rata. An Excel spreadsheet ***distributor.xls*** is included with this database to help to ensure that this is done in a consistent fashion. (\\Data\Tools).

Checking if results are significant

57. An additional Excel spreadsheet, ***significant change tester.xls***, is included on the EHS folder \\mrdoc\excel that provides a quick method of checking whether changes in a particular indicator over time are

statistically significant. There is an accompanying word document in the same location, which explains how the Excel macro works.

Appendix A: UK Data Archive folder structure



Appendix B: Content and level of Interview Survey Files

File name	level	Key identifier	Contents
<i>Adapt.sav</i>	household	aacode	Information on disability aids used by most disabled person and degree of difficulty experienced due to disability
<i>Adaptation.sav</i>	adaptation	aacode	Information on internal and external adaptations for those with disability
<i>Attitudes.sav</i>	household	aacode	Information on attitudes about the home, neighbourhood and moving
<i>Contact.sav</i>	household	aacode	Key information from the doorstep form
<i>Damp.sav</i>	household	aacode	Type of damp and type of problem
<i>Disability.sav</i>	person	persno	Information on types of disability
<i>Dwelling.sav</i>	household	aacode	Type of accommodation Age of acc, housing history Ownership type, buying aspirations
<i>Employment.sav</i>	person	persno	Employment information and ethnicity
<i>FirstImp.sav</i>	household	aacode	Key information on the first impressions of the dwelling and neighbourhood as recorded by interviewer
<i>Hhdtype.sav</i>	household	aacode	Type of damp and type of problem
<i>Identity.sav</i>	person	Person	Nationality, Country of Birth, National Identity, Ethnicity, Religion
<i>Income.sav</i>	household	aacode	Information on amounts and sources of income including benefits
<i>Owner.sav</i>	household	aacode	Ownership details, mortgage
<i>People.sav</i>	person	persno	Key info for each person (age, sex etc.) and relationships with other people in the household
<i>Renter.sav</i>	household	aacode	Renting, Social renters, housing benefit Tenancy agreements and deposits,
<i>Repairs.sav</i>	household	aacode	Information on repairs, builders and part of the set of questions on damp
<i>Rooms.sav</i>	household	aacode	Key information about the number of rooms and whether shared
<i>Secondhomes.sav</i>	household	aacode	Second homes
<i>Vacant.sav</i>	household	aacode	Information relating to vacant properties from the doorstep form
<i>Waitlist.sav</i>	household	aacode	Waiting lists for social housing module
<i>Workdone.sav</i>	item	aacode	Key information about work done to the home

Appendix C: Content of derived variables

Filename	Variable	Label
Generalfsxx.sav	aacode	EHS case number
	aagfhxx	Household weight
	Paired	Whether paired sample case
	fqtr	Fieldwork quarter
	fmonth	Fieldwork month
	fyear	Fieldwork year
	tenure8x	Tenure
	tenure4x	Tenure
	tenure2x	Tenure
	gorEHCS	Government office region- EHCS
	gorEHS	Government office region- EHS
	region3x	Overall region of England
	rumorph	Rural urban morphology (COA)
	nrf88	NRF districts (original funding)
	nrf86	NRF districts (2006 revised funding)
	nrf91	NRF districts (original or revised)
	imd1004	2004 decile ranking of areas (lower layer SOA)
	imd1007	IMD07 decile ranking of areas
	pthfind	Market renewal pathfinder areas
Interviewfsxx.sav		
	aacode	EHS case number
	hhcomp	Household composition
	hhcomp1	Household composition
	hhtype11	Household type - full 11 categories
	hhtype7	Household type - 7 categories
	hhtype6	Household type - 6 categories
	hhsizex	Number of persons in the household
	famnumx	Number of family units in hhold
	loncoup	Single householder or with partner
	NLpar	Number of lone parent families in household
	ncouple	Number of couples in household
	Nsing	Number of one-person family units in household
	NDEPCHILD	Number of dependent children in household
	NXDEPCH	Number of non-dependent children in household
	otherfam	Additional families present in household
	othfam1p	Type of additional families in household
	NoUnits1	Banded number of family units in household
	Cohabhrp	If HRP is cohabiting
	Cohabprt	If HRP partner is cohabiting
	Studprt	If HRP's Partner is a full time student
	SFT	Number of FT workers in HHLD
	agehrpx	Age of HRP - continuous
	agehrp2x	Age of household reference person - 2 band
	agehrp4x	Age of household reference person - 4 band

	agehrp6x	Age of household reference person - 6 band
	ager	Report age categories
	agepartx	Age of partner - continuous
	pyngx	Age of youngest person in household
	pyngbx	Age band of youngest person in household
	agen16	Number of persons under 16 in household
	GrossA	Age Group & Sex of youngest person in household
	ageoldx	Age of oldest person in household
	ageoldbx	Age of oldest person in household - banded
	olderx	No of people aged 60 plus who are HRP or partner
	emphrpx	Employment status (primary) of HRP
	emphrp3x	Working status of HRP (primary) - 3 categories
	empprt	Employment status (primary) of Partner
	empprt3x	Working status of partner (primary) - 3 categories
	nssech	NS-SEC Socio-economic Classification - HRP
	nssecp	NS-SEC Socio-economic Classification - HRP's partner
	hhempx	Employment status of HRP and partner combined
	NEmp	Number of employed persons in household
	NUemp	Number of unemployed persons in household
	NInac	Number of economically inactive persons in household
	NStud	Number of students in household
	studhrp	If HRP is a full time student
	workless	Household with no one of working age employed - ILO defn
	ethhrp8x	Ethnic origin of HRP - 8 categories
	ethhrp4x	Ethnic origin of HRP - 4 categories
	ethhrp2x	Ethnic origin of HRP - 2 categories
	ethprt	Ethnic group of hrp's partner
	ethhrp8N	Ethnic origin of HRP - 8 categories (inc missing val)
	ethhrp4N	Ethnic origin of HRP - 4 categories (inc missing val)
	ethhrp2N	Ethnic origin of HRP - 2 categories (inc missing val)
	ethprtN	Ethnic group of hrp's partner (inc missing val)
	sexhrp	Sex of household reference person
	sexprt	Sex of hrp's partner
	FreeLeas	Freehold or leasehold
	Owntype	Type of ownership
	hatentp	Type of housing association tenancy
	bedstdx	Bedroom standard
	NBedsX	Total no of bedrooms household actually has
	bedrqx	No. of bedrooms required by the household
	nshare	If shares any part of accomodation
	nkit	If shares a kitchen
	totkits	Number of unshared kitchens at least 6.5 feet wide

	nbath	If shares a bathroom
	nother	If shares other room
	nrooms	Number of rooms available to household
	nrooms1	Number of rooms available to household grouped
	sharer	If shares accommodation with other household
	accomhh	Type of accommodation for household
	accomhh1	Type of accommodation for household & if not self-contained
	lenres	Length of residence (years)
	lenresb	Length of residence
	lenown	Length of ownership (years)
	lenownb	Length of ownership to date of survey
	Buypresh	Year HRP bought present accommodation
	ftbuyer	If first-time buyer
	Prevten	Previous Tenure
	tenex	Extended tenure of household
	tenure1	Tenure Group 1
	tenure2	Tenure Group 2
	tenure3	Tenure Group 3
	tenure4	Tenure Group 4
	hhitsick	Anyone in hhold have < illness or disability?
	hpregdis	HRP or Partner registered disabled?
	mortwkx	Weekly mortgage payments
	rentwkx	Total weekly rent payable (rent plus HB)
	rentflg	Rent/HB changed/imputed
	rentExS	Total weekly rent excluding the cost of services
	rentExSflg	Rent excluding services changed/imputed
	hhincx	EHS Basic Income (annual net household income (HRP + Partner) including savings)
	hhinc5x	All households - income in 5 bands
	hhincflg	Imputations used to create net total hhold income
	amthbenx	Weekly HB
	housbenx	Household receives any housing benefit?
	hhbensx	Household on means tested bens or tax credits with income under £15460?
	JOINTINCx	Annual gross income of the HRP and partner
	HYEARGRx	Household gross annual income (inc. income from all adult household members)
	hhvulx	Household vulnerable - on means tested or WF benefits?
	equityh	Equity in home (hybrid version based on mkt value and hhold estimate)
	equityh5	equity in home (hybrid version based on mkt value and hhold estimate)
	equityr	Equity in home (based on respondent valuation only)
	equityr5	equity in home (based on respondent valuation only)
	BHCinceq	BHC equivalised weekly income (modified OECD scale)
	BHCinceqv5	BHC equivalised income quintiles (weighted by

		peoplegross)
	BHCinceqv60h	below 60% of median income (weighted by peoplegross)
	AHCinceq	AHC equivalised weekly income (modified OECD scale)
	AHCinceqv5	AHC equivalised income quintiles (weighted by peoplegross)
	AHCinceqv60h	below 60% of median income (weighted by peoplegross)