Standard Report on Methods and Quality for Household Budget Survey

Standard Report
on
Methods and Quality (v1)

for

<Household Budget Survey>

This documentation applies to the reporting period:

<2015-2016>

Last edited: <26.06.2017>

CENTRAL STATISTICS OFFICE

Skehard Road, Cork

021 4535000

www.cso.ie

Table of Contents

6 Additional documentation and publications

1 Overview

The Household Budget Survey (HBS) is a survey of a representative random sample of all private households in the State. HBS surveys have been carried out periodically in Ireland since 1951. The main purpose of the HBS is to determine in detail the pattern of household expenditure in order to update the weighting basis of the Consumer Price Index. The maintenance of a detailed diary of household expenditure over a two-week period by the surveyed households is thus the main distinguishing feature of the HBS. Detailed information is also collected on all sources of household income and on a range of household facilities.

The reliability of the results depends to a large extent on the accuracy, reliability and completeness of the information provided by the respondents. Estimates for certain types of expenditure (e.g. alcohol and tobacco) and some categories of income tend to be underestimated in surveys of this nature and thus caution should be exercised in making comparisons between different categories of households. Notwithstanding these shortcomings, the HBS provides a wealth of valuable information on the income and expenditure of Irish households.

In 2015-2016, the data collection process was carried out by a team of five Temporary Full-Time Field Co-ordinators and 55 Temporary Part-Time Field Interviewers.

Each household that participated in the HBS completed a detailed household questionnaire which included questions on tenure status, household appliances, household facilities and housing costs (e.g. mortgage, rent). Each household member aged 16 years and over completed a personal questionnaire which included questions on income, education, work status and other demographic related questions. Data capture for both personal and household questionnaires was by means of CAPI (Computer Assisted Personal Interviewing). All household members aged 16 and over were also required to fill in a paper diary over a two week period, detailing all expenditure over this period. During the two week diary phase of a HBS, an interviewer returned to the participating household in order to check that household members were filling out their diaries in the correct manner.

In 2015-2016, a total of 18,499 sample addresses were visited by interviewers during the course of the survey with a final achieved sample of 6,839.

2 General Information

2.1 Statistical Category

Primary Statistical survey. However, a combination of survey and administrative data is used to calculate household income.

2.2 Area of Activity

Housing and households

2.3 Organisational Unit Responsible, Persons to Contact

The Household Budget Survey falls under the remit of the Social and Demographic Statistics Directorate, headed by Richard McMahon, Assistant Director General. Paul M. Crowley is the Senior Statistician with responsibility for Social Analysis. If you have any queries on this report please contact:

Claire Burke	Tel: +353 21 4535046	Email: claire.burke@cso.ie
Tricia Brew	Tel: +353 21 4535148	Email: patricia.brew@cso.ie

2.4 Objectives and Purpose; History

The first HBS was conducted in 1951 and the survey has been carried out periodically since then. The main purpose of the HBS is to determine in detail the pattern of household expenditure in order to update the weighting basis of the Consumer Price Index. The HBS is intended to give a very comprehensive picture of household expenditure over the whole range of consumer goods and services. Detailed information is also collected on all sources of household income and on a range of household facilities.

2.5 Periodicity

Currently, the HBS is carried out every five years. Ten large-scale surveys have been undertaken in respect of the periods 1951-1952, 1965-66, 1973, 1980, 1987, 1994-1995, 1999-2000, 2004-2005, 2009-2010 and 2015-2016. The 1951-1952 and the 1965-1966 surveys were, however, restricted to Urban areas. All subsequent surveys covered both Urban and Rural areas.

2.6 Client

The release is produced to meet National requirements. There is currently no Eurostat regulation governing this data.

2.7 Users

- Other CSO sections - Consumer Price Index, National Accounts
- Eurostat
- Government departments
- Other research centres
- General Public

2.8 Legal basis

 Activities of the CSO are governed by the Statistics Act 1993.

3 Statistical Concepts, Methods

3.1 Subject of the Statistics

Household expenditure and income

3.2 Units of Observation/Collection Units/Units of Presentation

The survey relates solely to private households in the State.

The survey population and collection units are private households in the States. A household is defined as a single person or group of people who regularly reside together in the same accommodation and who share the same catering arrangements. The household members are not necessarily related by blood or marriage. Institutional households, (e.g. nursing homes, barracks, boarding schools, hotels etc.) were not covered by the survey.

Information is collected on each household and each usual resident within that surveyed household.

The reference person is the person in whose name the accommodation was owned or rented. Where the mortgage/rent is jointly paid, the respondent with the highest income is taken as the reference person. In cases where household members receive an equal salary, the eldest member is taken as the reference person.

The main units of presentation are:
Household composition
Household economic status
Characteristics of the household reference person
Household income
Household expenditure
Household accommodation
Household facilities and appliances
Household characteristics, eg location, NUTS3, tenure, size etc

3.3 Data Sources

HBS survey data is collected from individuals in households. Administrative data is provided by Revenue and the Department of Social Protection

3.4 Reporting Unit/Respondents

Each sample household was required to:
    (a) provide particulars of household membership, accommodation, facilities, amenities and regular household expenses (eg rent, electricity, telephone etc.)
    (b) have all constituent household members aged sixteen years and over provide details of their income and regular personal expenditure, and also maintain detailed diary records of their day to day personal expenditure for fourteen consecutive days.

Households were accepted as co-operating only if all the required data was provided and each household member aged sixteen years and over kept an expenditure diary. The refusal of one member to co-operate resulted in the whole household being eliminated.
Where a particular individual was not available for interview, information was provided by another member of the household, familiar with this individual, in most circumstances via a proxy interview. A proxy interview refers to data which is collected from another member of the household due to the unavailability of the specific respondent at the time of the interview. However, all household members aged sixteen years and over were required to maintain an expenditure diary for a period of 2 weeks.

3.5 Type of Survey/Process

Household sample survey

3.6 Characteristics of the Sample/Process

3.6.1 Population and Sampling Frame

The reference population is all private households in Ireland as per Census 2011.

3.6.2 Sampling Design

In 2015-2016, a new sampling methodology was introduced to improve the robustness of the HBS Sample. The following is a brief overview of the revised HBS sample methodology:

- The HBS sample was a multi-stage cluster sample resulting in all households in Ireland having an equal probability of selection.

- The sample was stratified by NUTS4 and quintiles derived from the Pobal HP (Haase and Pratschke) Deprivation Index.

- A sample of 1,430 blocks (i.e. Household Survey Collection Unit Small Areas, Census 2011) from the total population of 17,320 blocks was selected.

- Blocks were selected using probability proportional to size (PPS), where the size of the block is determined by the number of occupied households on Census night 2011.

- The second sampling stage involved the selection of 13 households within each block using simple random sampling without replacement (SRS) for inclusion in the survey sample. All occupied households on Census night 2011 within each block are eligible for selection in the SILC sample.

3.7 Survey Technique/Data Transfer

HBS interviewers visit each household and ask the usual residents aged 16 years and over to voluntarily participate in the survey. Questionnaire data are collected on encrypted laptop computers using CAPI (Computer Assisted Personal Interviewing). The survey interviewer transmitted the data collected from households to head office using a secure encrypted data tunnel. Each household member aged 16 years and over also maintained detailed diary records of their day to day personal expenditure for fourteen consecutive days. On completion, the interviewer collected the diaries and posted them back to HQ in registered security-tagged mailing pouches.

3.8 Questionnaire (including explanations)

Each sample household was required to:

(a) provide particulars of household membership, accommodation, facilities, amenities and regular household expenses (eg rent, electricity, telephone etc.)

(b) have all constituent household members aged sixteen years and over provide details of their income and regular personal expenditure, and also maintain detailed diary records of their day to day personal expenditure for fourteen consecutive days.

Households involved in farming or market gardening were asked a special farm income questionnaire.

A copy of the paper diary is available here:
http://www.cso.ie/en/media/csoie/releasespublications/documents/ep/householdbudgetsurvey/2015-2016/appendices/HBS_Diary_2015-2016_for_web.pdf

3.9 Participation in the Survey

Participation in the survey is voluntary. Co-operating households benefited from an inducement in appreciation of their participation. If each person aged sixteen years and over in a sample household satisfactorily co-operated in the survey, they were given a gratuity payment of €30.

3.10 Characteristics of the Survey/Process and its Results

The main data detailed in the publication refer to the average weekly household expenditure classified by various household characteristics. Data is also published on household income, household composition and household appliances and facilities, classified by the same household characteristics.

Detailed data is collected through the interview and diary keeping period to allow outputs to be produced according to all the relevant classifications

3.11 Classifications used

Urban and rural households

Urban households are defined as being located in cities and towns, including their suburbs, distinguished for Census of Population purposes. The boundaries of towns (with population of 1,000 or more) and their classification by population size were based on the 2011 Census of Population.

Rural households are all other households located outside the boundaries of cities and towns (including suburban areas).

Gross household income deciles

The income decile groups are ten equal-sized groups of households, each group containing 10% of households. The "1st decile" group contains the 10% of households with the lowest gross household income, the "2nd decile" group contains the 10% of households with the next lowest gross household income etc. The "10th decile" group contains the 10% of households with the highest gross household income.

Regions

Regional data is coded to NUTS3 as described in Section 3.12

Social Group

Sample households have been classified in this report to the social group of the reference person of the household. Five Social Groups are distinguished. These are defined in terms of the following combinations of the eleven Socio-Economic Groups used in the 2011 Census of Population:

Social Group   **Socio-Economic Group**

1   Higher professional
               Lower professional
               Employers and managers

| 2 | Own account workers |
| | Non-manual workers |

| 3 | Manual skilled workers |

| 4 | Semi-skilled manual workers |
| | Unskilled manual workers |

| 5 | Farmers, agricultural workers |
| | All other gainfully occupied and unknown |

Deprivation index

The Pobal Haase-Pratschke Deprivation Index is used to analyse Household Budget Survey questionnaire responses experienced by households. The Index uses Census data to measure levels of disadvantage or affluence in a particular geographical area. More detailed information on the index can be found here: https://www.pobal.ie/pages/new-measures.aspx. The results are presented by quintiles, five equal-seized groups of households, with the first quintile representing the least deprived/most affluent area and the fifth quintile represent the most disadvantaged areas.

Household composition

This is a classification of households based on the number of adults and children in the household. The categories are:

- 1 adult

- 2 adults

- 1 adult with children

- 2 adults with 1-3 children

- 3+ adults

- Other households with children

3.12 Regional Breakdown of Results

The regional classifications in this release are based on the NUTS (Nomenclature of Territorial Units) classification used by Eurostat. The NUTS3 regions correspond to the eight Regional Authorities established under the Local Government Act, 1991 (Regional Authorities) (Establishment) Order, 1993, which came into operation on 1 January 1994. The NUTS2 regions, which were proposed by Government and agreed by Eurostat in 1999, are groupings of the NUTS3 regions. The composition of the regions is set out below.

Border, Midland and Western NUTS2 Region		**Southern and Eastern NUTS2 Region**	
Border	Cavan	**Dublin**	Dublin
	Donegal		Dun Laoghaire-Rathdown

	Leitrim		Fingal	
	Louth		South Dublin	
	Monaghan			
	Sligo	**Mid-East**	Kildare	
			Meath	
Midlands	Laois		Wicklow	
	Longford			
	Offaly	**Mid-West**	Clare	
	Westmeath		Limerick City	
				Limerick County
West	Galway City			
	Galway County	**South-East**	Carlow	
	Mayo		Kilkenny	
	Roscommon		South Tipperary	
			Waterford City	
			Waterford County	
			Wexford	
		South-West	Cork City	
			Cork County	
			Kerry	

4 Production of the Statistics, Data Processing, Quality Assurance

4.1 Data Capture

Questionnaire data is collected in the field by a team of interviewers using laptop computers (CAPI using a Blaise application) and data is then transmitted to the main processing unit in the CSO.

Expenditure diaries were paper based. Once completed by HBS household members, they were collected by interviewers and returned to the CSO in Cork. A processing team in the CSO scanned diary and receipt items included in the diaries (over 950,000 individual items) into a data processing system.

A copy of the paper diary is available here:
http://www.cso.ie/en/media/csoie/releasespublications/documents/ep/householdbudgetsurvey/2015-2016/appendices/HBS_Diary_2015-2016_for_web.pdf

4.2 Coding

Occupation and Industry text strings are captured and coded in the field to the relevant classifications (see Section 3.11) by interviewers using the Blaise application

Each expenditure item was assigned a COICOP (Classification of Individual Consumption According to Purpose) code. This coding was done in the CSO. Data range checks were also carried out to ensure cleanliness of data provided.

The COICOP classification is available here:
http://ec.europa.eu/eurostat/ramon/nomenclatures/index.cfm?TargetUrl=LST_NOM_DTL&StrNom=HICP_2000&StrLanguageCode=EN&IntPcKey=37591913&StrLayoutCode=HIERARCHIC

4.3 Data Editing

The majority of questions only allow answers to be entered to a limited set of predefined categories and therefore the number of edits required is limited. Questionnaire routing is used to ensure questions are only asked to relevant respondents e.g. unemployment questions are only asked to those who are unemployed. In addition, invalid responses are prevented at the point of capture where appropriate (for example a flag is raised to an interviewer if the date of birth captured which imply a derived age of more than 110) and this ensures that implausible data is prevented from being captured. Data range checks were also carried out on expenditure items to ensure cleanliness of data provided.

4.4 Imputation (for Non-Response or Incomplete Data Sets)

Item non-response was generally very low.

Expenditure:
If an individual/household declared a certain type of expenditure but could not recall the monetary amount involved, the average of other individuals/households with the same characteristics was applied to this instance.

Income:
If an individual could not recall their income, and no administrative data was available, the average hourly rate of other persons with similar jobs was applied to this individual.

4.5 Grossing and Weighting

In order to provide national results, the survey results were weighted to agree with the number of private households from the 2011 Census of Population. The process used was as follows:

- ☐ Firstly design weights were calculated for all units selected in the initial sample and are computed as the inverse of the selection probability of the unit. The purpose of design weights is to eliminate the bias induced by unequal selection probabilities.
- ☐ Next these design weights were then adjusted for non-response. This eliminated the bias induced by discrepancies caused by non-response between the initial sample and the achieved sample. Particularly critical when the non-responding households are different from the responding ones in respect to some survey variables as this may create substantial bias in the estimates. Design weights are adjusted for non-response by dividing the design weights of each responding unit in the final/achieved sample by the (weighted) response probability of the corresponding group or strata.
- ☐ To obtain the final weights for the results, after the previous steps were carried out, the distribution of households by deprivation, NUTS3 region, tenure household size, household type, was calibrated to the population of households in 2015 (as derived from the Quarterly National Household Survey). The CALMAR2-macro, developed by INSEE, was used for this purpose and both household and individual external information was used in a single-instance calibration at household level.

4.6 Computation of Outputs, Estimation Methods Used

Output results are aggregated to produce the various totals published. In general, all aggregations produced are done by way of various SAS procedures. The aggregate results produced for any given set of classifications will be the sum of the individual grossing factors of the valid responses which belong to that set of classifications and no estimates are made unless the data itself has been captured within the survey.

4.7 Other Quality Assurance Techniques Used

A series of field audits were carried out each to verify the quality of the individual data being collected. Interviewers receive regular feedback on the quality of the data they are producing and any unusual trends are highlighted and followed-up. Interviewers undergo training when first employed and regular meetings are held with field coordinators to ensure standards are maintained. Further training is also provided to interviewers where appropriate.

5 Quality

5.1 Relevance

The main purpose of the HBS is to determine in detail the pattern of household expenditure in order to update the weighting basis of the Consumer Price Index.
Other CSO sections which make use of HBS data include, but not limited to, National accounts and Tourism

External users who make significant use of HBS data include, but is not limited to:
- ESRI
- Central Bank
- SEAI
- Insolvency Service of Ireland
- Department of Finance
- Department of Social Protection

5.2 Accuracy and Reliability

5.2.1. Sampling Effect & representivity

The HBS sample is designed to be a randomly selected cluster sample with each household in the target population having an equal and known probability of selection. Non-response has the potential to introduce bias into the sample. HBS sample implementation procedures were designed to minimise non-response. The sample was designed for a field force of 55 interviewers. Adequate monitoring and management of the field-force availability was critical in assuring a high quality representative sample.

An on-going issue with household samples is the availability of field interviewers. When any of the interviewers were not available due to holidays, sickness or retirement, the interviewers were replaced by alternative interviewers (back-ups) whenever possible. Only 91 households were dropped from the original sample. Households initially flagged as uncontactable or soft-refusals were revisited at a later date. This had the effect of reducing non-response.

Design weights were adjusted for non-response. This eliminated the bias induced by discrepancies caused by non-response between the initial sample and the achieved sample.

5.2.2. Non-Sampling Effects

In addition to known sampling errors, any survey will be subject to other non-sampling errors (for example measurement errors arising from questions not capturing the desired information accurately). Non-sampling error is far more difficult to measure than sampling error and no formal estimate of non-sampling error is available in the HBS. Information on the interviews was

collected and analysed to help minimise non-sampling effects (including, for example, when interviews were conducted and their duration). This information was compared across the interview team to ensure no unusual variation in interviewer performance exists. Co-ordinators, as an additional check on the quality of the interviewer's work, called back to a number of each interviewer's households to check the quality of the collected data.

5.2.2.1 Quality of the Data Sources used (other than survey register)

The HBS 2015-2016 sample was based on the population of private households as per Census 2011.

The availability of administrative data from the Revenue and the Department of Social Protection has greatly improved the accuracy of HBS income data. Reliability of the overall income levels and overall social welfare payments for each individual on the dataset has also greatly improved. The variable that allows all this data to be linked is the PPSN number. Anomalies may still arise in these data sources and these are identified and resolved using a comprehensive micro-editing system. Plausibility checks are carried out on the survivor's pension payments, the old age pension and the disability payments where any anomalies uncovered are addressed. A small percentage of individuals are missing from these registers. In these instances, survey response data is used.

5.2.2.2 Register Coverage

The entire stock of private households at the time of the most recent Census of Population in the country represents the full sampling frame for the HBS. The sample was based on the 2011 Census. To reduce response burden households included in samples for other surveys eg QNHS, SILC are excluded.

5.2.2.3 Non-response (Unit and Item)

Currently there is no non-response adjustment used in the HBS. The table below gives a breakdown of the response and non-response to the HBS for the current publication.

Household status	
Target households	18,590
Households dropped *	91
Revised household target	18,499
Vacant	1,401
Outright refusals	6,371
Circumstantial refusals	669
Temporarily absent	2,082
Other reason no interview (including incomplete)	1,137
Actual households interviewed	6,839

* Households were dropped due to staffing shortages and unforeseen local circumstances.

5.2.2.4 Measurement Errors

No formal evaluation of sources of error is available, although measures were in place to minimise error (details below).
• The quality of the data collected was improved using field staff training and debriefings throughout the survey period
• Respondent effects - most of the requested information was readily available to respondents.
• Comprehension errors - most of the terms used by the survey are readily understood, although some issues occasionally arise.

5.2.2.5 Processing Errors

a. Data capture errors: These errors are minimised by logic checks and limits on values that can be keyed for each question in the electronic questionnaire at the data collection point.

b. Coding error: Checks were in place to minimise this risk, particularly with respect to COICOP coding. The coding is initially performed in the data processing section with checks performed by managers.

5.2.2.6 Model-related Effects

Not applicable

5.3 Timeliness and Punctuality

5.3.1 Provisional Results

No provisional outputs are published.

5.3.2 Final Results

HBS 2015-2016 results have been published 15 months after the completion of fieldwork.

The HBS publications can be found on the CSO website:
http://www.cso.ie/en/statistics/housingandhouseholds/

5.4 Coherence

Coherence checks were carried out with auxiliary sources, eg National Accounts, SILC

5.5 Comparability

 To ensure comparability with other official statistics standard classifications are used for HBS expenditure estimates (i.e. COICOP). The classification used is set by Eurostat.

5.6 Accessibility and Clarity

5.6.1 Assistance to Users, Special Analyses

All publications are available on the CSO website. Information on methodology is also available on the website. The background notes on the publication provide some detail on the survey.

Ad-hoc analysis can also be produced on request.

Anonymised microdata is made available to researchers via the Irish Social Science Data Archive (ISSDA). Such data is accessible by researchers applying directly to the ISSDA. For further information see: http://www.ucd.ie/issda/

Access to a Research Microdata Files (RMFs) can be requested from the CSO under the CSO's microdata access policy. Extensive use is made of this facility by the research community. For further information see:
http://www.cso.ie/en/aboutus/dissemination/accesstomicrodatarulespoliciesandprocedures/researchaccesstomicrodatafiles/

5.6.2 Revisions

Not applicable

5.6.3 Publications

5.6.3.1 Releases, Regular Publications

The HBS is carried out every 5 years. Results are published as promptly as possible.

5.6.3.2 Statistical Reports

Not applicable

5.6.3.3 Internet

All HBS publications are available on the CSO website in publication format.
In addition data is made available via the CSO's main databank dissemination tool and is also hosted on the CSO website in Excel format:

Releases and publications

http://www.cso.ie/en/releasesandpublications/ep/p-hbs/hbs20152016/

Databank dissemination

http://www.cso.ie/px/pxeirestat/Database/eirestat/Household%20Budget%20Survey%202015%20to%202016/Household%20Budget%20Survey%202015%20to%202016_statbank.asp?SP=Household%20Budget%20Survey%202015%20to%202016&Planguage=0

5.6.4 Confidentiality

The confidentiality of all information provided to the CSO by individual respondents is guaranteed by law under the 1993 Statistics Acts. All CSO office and field personnel become "Officers of Statistics" on appointment and are liable to penalties under this Act if they divulge confidential information to any outside person or body. Extreme precautions are taken to ensure that there are no violations of this principle throughout the survey process. The laptops on which the data

was collected are encrypted and contain several layers of password protection. Data are only published in aggregate form and care is taken to ensure that the data are aggregated to avoid the indirect identification of respondents. Confidentiality is also ensured within the anonymised microdata by using coded variables instead of original values for key characteristics. For example, age groupings are provided instead of single year of age

6 Additional documentation and publications

The main CSO Housing and Households homepage can be found at the following link:
http://www.cso.ie/en/statistics/housingandhouseholds/

Statcentral – Ireland's portal to official statistics:
http://www.cso.ie/px/pxeirestat/statire/SelectTable/Omrade0.asp?Planguage=0