

General features of household surveys at Statistics Finland

Veli-Matti Törmälehto
Amman, Jordan, 12 June 2023

Contents

- 01 Household Surveys at Statistics Finland
- 02 Sampling, frames and precision requirements
- 03 Modes of collection
- 04 Collection of consumption data in the HBS
- 05 Non-response, editing, imputation and weighting/estimation
- 06 Quality reporting and dissemination

Household surveys at Statistics Finland

According to Statistics Act, Statistics Finland must use administrative data to produce statistics if such data are available

Surveys are conducted mostly for European statistics under the [EU framework regulation](#) for

Integrated European Social Statistics

(TESS) relating to

Surveys and data collections - Statistics Finland

Select inquiry

Inquiry name

Respondent group

Time period

[Filter](#)

Results

Showing 1 - 9 of 11 results

Adult education survey

Due date for response: Ended

Consumer Confidence Survey

Ended

Household Budget Survey

Ended

Housing and income survey

Labour Force Survey

Renovation building of dwellings and detached houses

Due date for response: Ended

Statistics on goods transport by road

Survey on Income and living conditions

Ended

Time use survey

Due date for response: Ended

- *Labour Force Survey (LFS), monthly, panel component*
- *Survey on Income and Living Conditions (EU-SILC), annual, panel component*
- *Household Budget Survey (HBS) ~every 5 years, pure cross-section*



Sampling, frames and precision requirements

Probability samples are drawn from a frame which is updated four times per year from a **continuously updated population statistics database** of Statistics Finland.

Systematic random samples of individuals from a sampling frame sorted by domicile code (address)

- implicit stratification by region & inclusion probability proportional to household size
- explicit **stratification** based on domain specific needs (e.g. income class in EU-SILC)

Frame is up to date and has low under-coverage. It can be restricted to private households when needed and generally has very low over-coverage.

- Target population of HBS and EU-SILC is private households, excluding persons without a postal address, the institutional population (e.g. long-term residents of old people's homes, care institutions, prisons or hospitals), persons permanently resident abroad and persons temporarily resident in Finland.



Sampling, frames and precision requirements

EU IESS framework regulation ([annex II](#)) sets the **precision requirements** expressed in standard errors for selected indicators in each domain

Some examples of the indicators:

- Labour (LFS): Estimated (national) quarterly unemployment-to-population 15-74 ratio
- LFS monthly sample size 12 500 individuals

- Income and Living Conditions (EU-SILC): Ratio at-risk-of-poverty or social exclusion to population
- EU-SILC sample size ~9500 households

- Consumption Expenditure (HBS): Percentage of households whose expenditure on housing-related categories including water, electricity, gas and other fuels, are more than 50 % of the total expenditure
- HBS sample size ~3 000 households



Data collection modes and data sources

Primary modes of direct data collection are **Computer Assisted Telephone Interviewing (CATI)** and **online interviewing (CAWI)**, face-to-face interviewing (CAPI) is very limited

In addition, HBS uses **scanning of receipts and a mobile app** to collect data on consumption

Data collection is managed by [Interview and Survey Services of Statistics Finland](#), interviewers are permanent employees of Statistics Finland

- Maximum use of **administrative data** in surveys; demographic and income data in all surveys based on administrative data record linked to survey units using unique personal identification numbers
- Register of household-dwelling units (household based on co-residence) + panel setting in LFS and SILC → direct identifiers are already available for most survey respondents (personal identification numbers)

• For further background information, see this [publication](#) on the use of registers in EU-SILC



HBS Data collection of consumption data

1. The contact interview
2. Questionnaire
 - Web or CATI
3. Accounting period
 - HBS app or paper diary

Contact interview is the first entry point for each household and it is conducted by phone due to data security reasons. At this point a household roster is established (using a pre-filled roster in the electronic questionnaire based on administrative data).





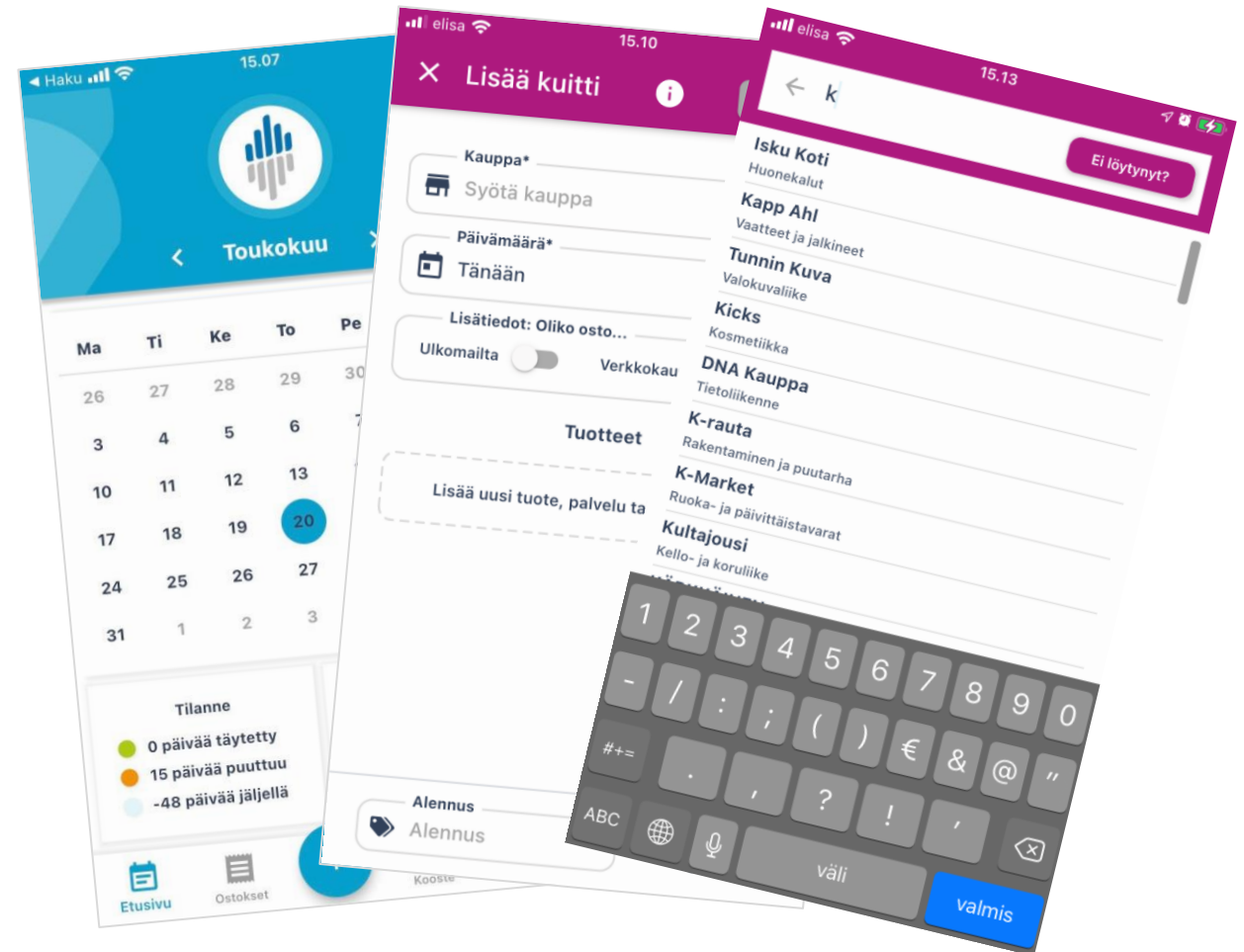
HBS Questionnaire

After the contact interview, the respondent can answer to the questionnaire:

1. online
2. in a phone interview
3. or in exceptional cases, in an on-site interview.

HBS accounting period

- The two-week period is conducted by collecting receipts and making manual entries for everyday purchases, groceries etc.
- Two alternative collection methods:
 1. HBS app for scanning a receipt with a smartphone (also includes a diary function for manual entries)
 2. Physical receipts are sent in an envelope by mail (includes a paper diary)
- Both options are included to prevent framing error
- More than 50 % of the participating respondents used the app during the data collection



Statistics Finland's HBS App

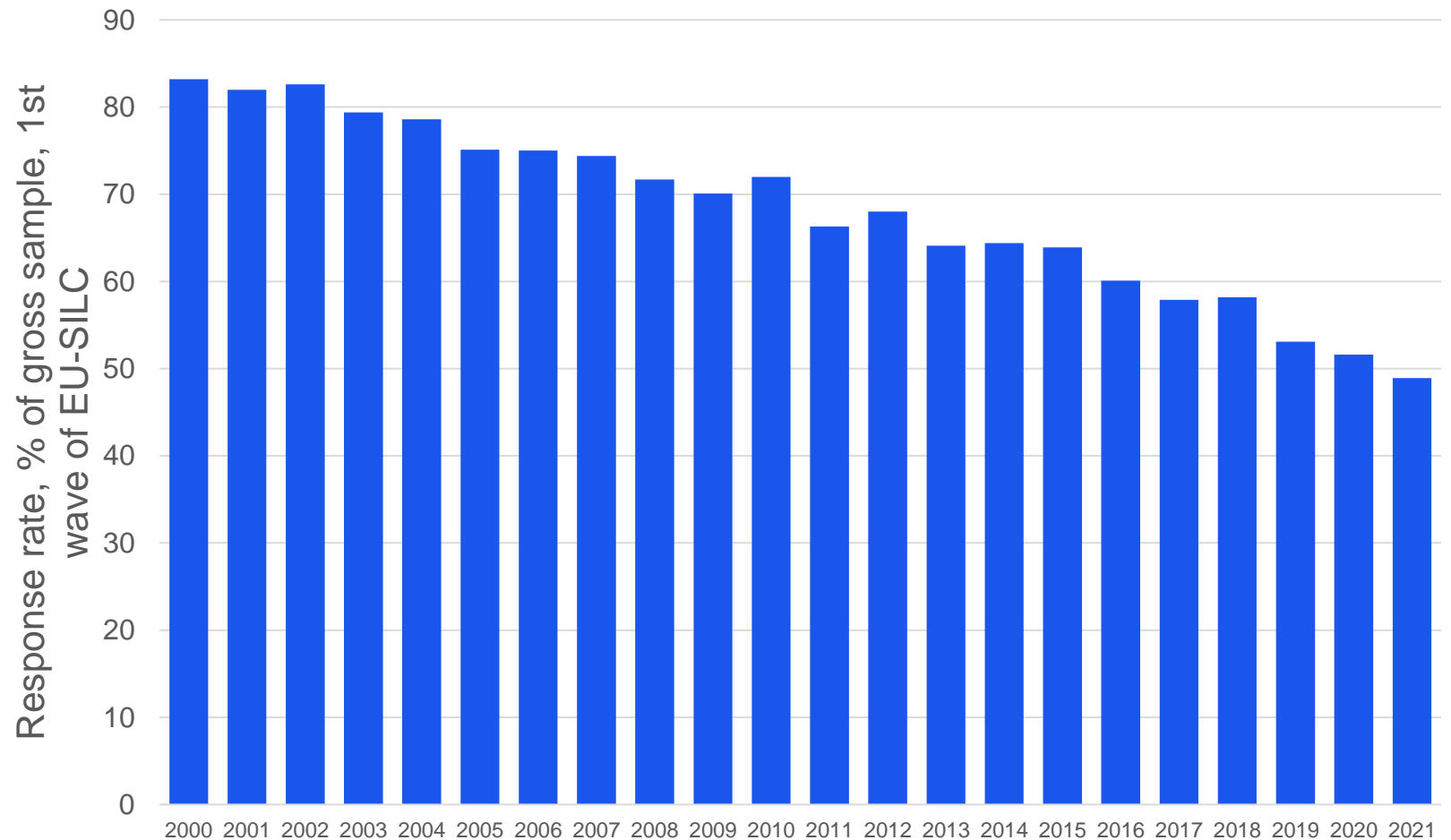
- The main functions of the app are:
 1. Receipt scanning utilizing the camera of a smart device
 2. A diary for manual entries (purchases without receipts etc.) – CPI based product lists guide the process.
- Mainly a receipt photo and purchase deposit tool.
 - In-house OCR and classification algorithm are used later in the process for app photos and scanned paper receipts alike.
- Each household has their own log in credentials.



Response rates

- Declining response rates a challenge in all surveys
- Response rate now slightly < 50 % in LFS and SILC; <40 % in the HBS
- Substitutes introduced in the HBS 2022 because of increasing unit non-response (controlled substitutes based on administrative data on age, gender, region, income decile, education, and household size and structure)

Response rate in the Survey on Income and Living Conditions (SILC) 2000-2021



Editing, imputation and weighting

Editing limited and mostly related to consistency checks with administrative data; logical checks already implemented in the electronic questionnaires to prevent inconsistencies

Item non-response on non-subjective variables are generally imputed (e.g. housing costs in SILC and HBS); item non-response rates tend to be low but there is variation between variables/topics

- Missing values in the data are flagged using flag variables/shadow variables and single imputation applied (multiple imputation has been used in wealth statistics and statistical matching)
- Preference to real donor/hot deck methods (to avoid regression towards the mean and implausible values)

All samples are probability samples → design **weights** as inverse of inclusion probabilities

- Non-response adjustments (in EU-SILC, inverse of response rates by strata)
- Calibration of weights to population margins; usually by age, gender, education, household size plus domain specific variables (e.g. income in EU-SILC)



Dissemination and quality reporting

- National official statistics, StatFin statistical database
- Micro data transmitted to Eurostat for European statistics and made available for research purposes
- Standardized quality documentation in all statistics (see e.g. [Labour Force Survey documentation](#))
- [Quality reports for Eurostat](#) based on Single Integrated Metadata Structure (SIMS) standard;
- Quality reports also publicly available from Eurostat (e.g. EU-SILC [comparative EU quality report and national reports](#))

Labour force survey: documentation of statistics

On this page

[Basic data of the statistics](#)

[Concepts](#)

[Classifications](#)

[Accuracy, reliability and timeliness](#)

[Comparability](#)

[Source data and data collections](#)

[Methods](#)

[Principles and outlines](#)

[Statistical experts](#)

[Revisions in statistics](#)

[Download the documentation](#)

[Comparison between the employment Ministry of Economic Affairs and Er](#)

Households' consumption: documentation of statistics

On this page

[Basic data of the statistics](#)

[Concepts](#)

[Accuracy, reliability and timeliness](#)

[Comparability](#)

[Source data and data collections](#)

[Methods](#)

[Principles and outlines](#)

[Statistical experts](#)

[Revisions in statistics](#)



Thank you!

