

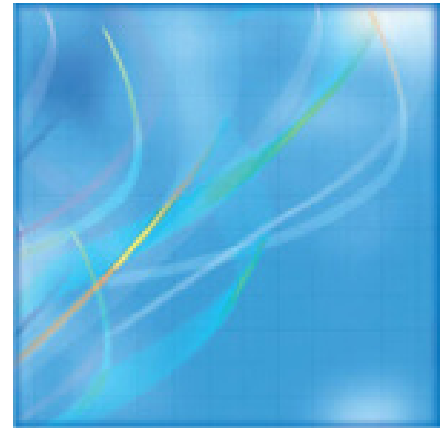
Catalogue no. 62F0026M
ISSN 1708-8879
ISBN 978-0-660-28910-6

Household Expenditures Research Paper Series

Survey of Household Spending Technical Note, 2017

Income Statistics Division

Release date: December 12, 2018



Statistics
Canada

Statistique
Canada

Canada

How to obtain more information

For information about this product or the wide range of services and data available from Statistics Canada, visit our website, www.statcan.gc.ca.

You can also contact us by

Email at STATCAN.infostats-infostats.STATCAN@canada.ca

Telephone, from Monday to Friday, 8:30 a.m. to 4:30 p.m., at the following numbers:

- Statistical Information Service 1-800-263-1136
- National telecommunications device for the hearing impaired 1-800-363-7629
- Fax line 1-514-283-9350

Depository Services Program

- Inquiries line 1-800-635-7943
- Fax line 1-800-565-7757

Standards of service to the public

Statistics Canada is committed to serving its clients in a prompt, reliable and courteous manner. To this end, Statistics Canada has developed standards of service that its employees observe. To obtain a copy of these service standards, please contact Statistics Canada toll-free at 1-800-263-1136. The service standards are also published on www.statcan.gc.ca under "Contact us" > "[Standards of service to the public](#)."

Note of appreciation

Canada owes the success of its statistical system to a long-standing partnership between Statistics Canada, the citizens of Canada, its businesses, governments and other institutions. Accurate and timely statistical information could not be produced without their continued co-operation and goodwill.

Published by authority of the Minister responsible for Statistics Canada

© Her Majesty the Queen in Right of Canada as represented by the Minister of Industry, 2018

All rights reserved. Use of this publication is governed by the Statistics Canada [Open Licence Agreement](#).

An HTML version is also available.

Cette publication est aussi disponible en français.

Table of contents

Executive Summary	4
Introduction	4
Data collection	4
Response rates	5
SHS 2017 response rates	5
SHS response rates over time	5
Response rates at the international level.....	6
Missing or invalid data	6
Minimizing errors and bias	6
Total non-response and calibration.....	7
Partial non-response	7
Influential values.....	7
Sampling variability	8
Data validation	8
Fitness for use of 2017 Survey of Household Spending data	9

Survey of Household Spending Technical Note, 2017

Executive Summary

The Survey of Household Spending (SHS), conducted by Statistics Canada, produces detailed spending information, as well as selected information on dwelling characteristics and household equipment.

SHS data are an important input into Statistics Canada's Consumer Price Index (CPI), and are used in the calculation of Gross Domestic Product (GDP). The SHS data are also used by many federal and provincial government departments to develop social and economic policies and programs. As well, the data are used by private sector, research organizations and academics to study various topics related to household consumption and well-being, such as housing, child care and health care affordability, energy consumption and spending patterns of various groups of the population.

In order to provide this information, households that are selected for the survey are asked to complete an interview. Further, half of those households are asked to complete an expenditure diary, reporting almost all expenditures made by the household members during the two-week period following the interview. In 2017, the response rate was 67% for the interview component and 41% for the diary.

Statistics Canada uses sound statistical methods and has thorough validation practices in the production and release of SHS estimates to reduce the risk of statistical bias. Through some statistical adjustments and techniques, Statistics Canada ensures the data are fit for use. Nevertheless, data users must take into consideration that estimates for detailed expenditures reported by a smaller portion of the sample, or for smaller geographical areas, may be subject to higher sampling error and higher risks of bias.

Introduction

The Survey of Household Spending (SHS) collects detailed information on household expenditures. It also collects information on household member demographic characteristics (e.g., age, sex, education level), as well as certain dwelling characteristics (e.g., type, age and tenure) and some information on household equipment (e.g., electronics and communication equipment).

The SHS disseminates detailed household expenditure estimates at the national, regional and provincial levels. At the national level, it provides household expenditures by household type, household tenure, age of the household reference person, and size of area of residence. Expenditures by household income quintile, as well as detailed food expenditures are disseminated at the national, regional and provincial levels. Because of small sample sizes, detailed expenditures are available for each of the territorial capitals only (i.e., with no further disaggregation by household characteristics).

This technical note provides information on the quality of the data from the 2017 SHS, and on the various adjustment methods and validation strategies used by Statistics Canada to ensure these data are fit for use.

Data collection

The 2017 SHS was a voluntary survey that collected information directly from respondents through two collection modes: a questionnaire administered through a personal interview, which uses recall periods based on the type of expenditure; and a diary of daily expenses that selected households completed over a two-week period.

The sample of the 2017 SHS consisted of 17,792 households in the 10 provinces and 929 households in the territorial capitals (Whitehorse, Yellowknife and Iqaluit). The data were collected on a continuous basis from January to December 2017, from a sample of households spread over 12 monthly collection cycles.

Following the interview, respondents selected to complete the expenditure diary (50% of the interview sample) were asked to record the expenditures of all household members for a two-week period starting the day after the interview. Households had the option of providing receipts for their purchases made during the two-week period to reduce the amount of information manually transcribed in the diary. For SHS 2017, expenditures estimated from the diary component of the survey accounted for 27% of total current consumption, with the remainder coming from the interview.

The collection model used for the survey is consistent with standard international models for household budget surveys. The vast majority of these programs use a combination of an interview and a diary. The different recall periods used in the interview and the use of a diary makes it easier for respondents to either recall or report expenditures. While this combination is deemed to be the most efficient method of collecting reliable expenditure information, it also places a high burden on respondents. In 2017, the median interview time was 63 minutes. The length of the interview depends on several factors such as the type or size of the household and the variety of expenditures involved (owners vs. renters, home repairs and renovations, car owners, etc.).

Response rates

SHS 2017 response rates

In 2017, the overall interview response rate was 66.9%, ranging from a high of 71.1% in Quebec to a low of 63.9% in Ontario. Half of the households selected for the interview were also selected to fill out a diary. Some of the households selected to complete the diary did not do so, or provided a diary that was deemed unusable. The diary response rate, among interview respondents who were selected to fill out a diary, was 62.9%. As a result, the final diary response rate (defined as the percentage of usable diaries relative to the number of households selected to fill out the diary) was 41.3% at the national level. It ranged from a high of 45.5% in Saskatchewan to a low of 36.8% in British Columbia.

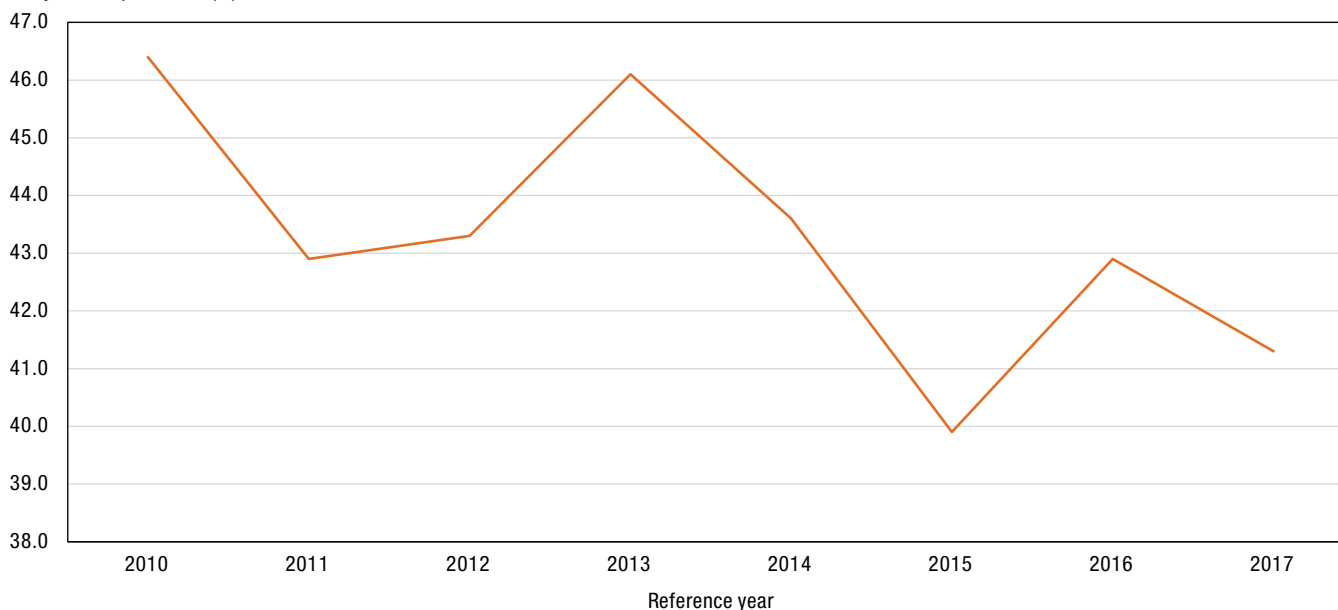
SHS response rates over time

Household survey response rates in Canada, as in other countries, have typically been declining over time. The SHS is not excluded from these trends. Although the response rate for the interview portion of the survey has been quite stable at around 65% since 2010, when the survey underwent a major redesign, it is somewhat lower than what it was in the late 1990s and early 2000s (70% to 75%). Maintaining a stable response rate in recent years has required an increased level of collection effort, with the number of household contacts required to complete an interview increasing during that period. The use of personal interviews for the interview component of the SHS, however, provides some level of mitigation against the declining response rates observed for the agency's telephone surveys, which can be affected by avoidance techniques (call blocking, call display) and harder-to-reach households (those with cell phones only).

With respect to the diary, although measures have been put in place to try to improve, or at least maintain response rates, household participation has been declining since 2010, as shown below. As mentioned earlier, expenditures estimated from the diary component account for roughly one quarter of total current consumption.

Chart 1
SHS diary final response rate, Canada (provinces only), 2010-2017

Diary final response rate (%)



Note: The diary final response rate is the percentage of usable diaries over the number of households selected for the diary.

Response rates at the international level

Expenditure surveys are commonly conducted by statistical agencies around the world. Although some differences may exist in their collection methodology (collection mode, length of interview, use of incentives), a comparison of these programs showed that among 35 countries, the SHS response rate was in the middle of the distribution. Many of these programs are experiencing a continuous decline in their response rates.

Missing or invalid data

Although the burden on SHS respondents is high, only a very small portion of the information reported in the questionnaires is missing or considered invalid. About 95% of households gave fewer than 10 missing or invalid responses out of almost 200 expenditure questions in the SHS 2017 questionnaire. Furthermore, a large portion of these missing data is due to the respondent not being able to provide the detailed expenditures that make up a bundled service, such as for bundled Internet, telephone, and television and satellite services.

Missing information is of greater concern for the expenses collected from the diary. Some responding households do not fully complete the diary over the two-week period, and respondents may also more easily forget certain types of expenses. As well, although a large proportion of households provide receipts for expenses such as groceries, when receipts are not provided, households often record only the total amount spent without transcribing the detailed items and costs in the diary. The incidence of this has increased over the past few years, requiring the total amount reported to be allocated to detailed expenditure items at the data processing step.

Minimizing errors and bias

Like all surveys, the SHS is subject to two types of errors: sampling and non-sampling errors. Sampling errors occur because inferences about the entire population are based on information obtained from only a sample of the population. Non-sampling errors include non-response errors, coverage errors, response errors and processing errors and may introduce bias in the estimates. More specifically, non-response can cause a bias in the estimates if the characteristics of non-respondents differ from those of respondents in a way that affects the expenditures studied.

Statistics Canada takes every effort to minimize bias in surveys. At the initial planning stages, qualitative testing techniques are used to ensure collection material is well understood by respondents and to assess burden. During collection, a follow-up is made with households that did not initially respond. Data collection reports are produced on an ongoing basis to be able to quickly react and take appropriate corrective action if necessary. An interview monitoring process is in place whereby recordings of a subsample of cases are reviewed in order to provide feedback to interviewers. Furthermore, interviewers follow up with diary respondents to remind them of items frequently forgotten in the diary.

In addition, sound statistical methods are used at the data processing and estimation steps to reduce the risk of non-response bias and/or the sampling errors.

Total non-response and calibration

To mitigate the potential non-response bias in the SHS, survey weights are adjusted so that the weights of the respondents represent the entire target population. This is done by grouping respondents and non-respondents who have similar characteristics together, and allocating the weights of the non-respondents to the respondents within each group. To create the groups, auxiliary information is required for both the respondents and the non-respondents. For this method to be effective in reducing the bias, the auxiliary information should be correlated with the probability of responding to the survey, as well as with the expenditure variables. Studies have shown that the auxiliary variables currently used to adjust the weights of the SHS respondents, such as household income, household size, and presence of a senior, are correlated with both the probability of response and the expenditures, and therefore contribute to reducing the bias due to the non-response.

Following non-response adjustments, the survey weights are further adjusted through calibration, which will ensure that certain estimates coincide with population totals obtained from external sources. Calibration contributes to reducing non-sampling errors such as non-response errors and coverage errors. When the calibration control totals are correlated with the variables of interest, the precision of the estimates can also be improved. In the case of the SHS, the survey weights in the provinces are calibrated on variables such as age groups, household size, and income groups, all of which are highly correlated with expenditures. This contributes to reducing the variance of the estimates, and potentially also reducing the non-sampling errors. In the three territorial capitals, the survey weights are calibrated on age groups and household size.

Partial non-response

To address missing or invalid information in interview questions, donor imputation using the nearest neighbour method is generally applied. That is, data from another respondent with similar characteristics (the donor) are used to impute. The characteristics used to identify the donor are selected such that they are correlated with the variables to be imputed. Household income, dwelling type, and the number of adults and children are commonly used characteristics. This contributes to reducing the bias due to partial non-response.

Donor imputation is also used when information is missing from the expenditure diary or when the level of detail provided by the respondent is not sufficient. Diary imputation is carried out at the reported item level, and the characteristics frequently used to identify the donor are the item's cost, available partial item code, household income and household size. Imputation is carried out by province and quarter to control for provincial differences and seasonality of expenditures.

Adjustment factors are also calculated to take the diary non-responded days into consideration, in order to avoid under-estimating the expenditures. For each diary, an adjustment factor is computed according to the number of non-responded days in the diary and the characteristics of the diary. Studies have shown that this adjustment factor contributes to reducing bias in the diary estimates.

Influential values

After edits and imputation, all the survey expenditure variables are annualized. That is, they are multiplied by an appropriate factor, based on the reference period, so that annual expenditure estimates can be produced. The annualization process can inflate expenditure values that are already large and amplify their impact on the estimates. It is generally advisable to identify these units and potentially reduce their impact on the estimates. The treatment of

influential values in the SHS relies on the contribution of a unit, which is defined as the product of the survey weight and the annualized expenditure. Units with an excessive contribution relative to the other units are adjusted. This correction process does not affect many units, but can significantly contribute to reducing the variance.

Sampling variability

The SHS estimates are based on a sample and are therefore subject to sampling error. One measure of sampling error is the coefficient of variation (CV). A high CV indicates high variability of the estimates. CVs are generally higher for expenditures reported by a smaller portion of the sample or when differences between households' spending are larger. Estimates for more detailed expenditure categories are thus more variable, in general, than summary-level expenditures. SHS estimates with a CV greater than or equal to 35% are deemed to be too unreliable and are therefore suppressed.

For national estimates, the majority of summary-level expenditures (e.g., food expenditures) have a CV below 4%. Nearly all intermediate-level expenditures (e.g., food purchased from stores) are considered reliable (they have a CV below 35%). The vast majority of detailed expenditures (e.g., eggs) are also considered reliable.

In general, CVs also increase as the geographic area becomes smaller. At the provincial level, except for Prince Edward Island, the total current consumption estimates all have a CV of 2% or less. Almost all summary-level expenditures have a CV below 13% and the vast majority of intermediate-level expenditures are deemed reliable. A majority of detailed expenditures are also considered reliable.

The main objective of the SHS sampling design is to find a balance between providing quality estimates at the national and provincial levels. For some smaller geographic areas such as Census Metropolitan Areas (CMA), a vast majority of summary-level expenditures have CVs below 35%, however, a substantial portion of detailed expenditures are deemed unreliable. The same holds true for the territorial capitals.

Lastly, while the SHS is a cross-sectional survey that is designed to provide estimates for one time period (the reference year), data users may be interested in comparing estimates through time. Users are advised that differences in estimates between two reference years are affected by the sampling variability of each of the two estimates compared. Therefore, the sampling variability associated with a difference between two estimates is generally higher than the variability of the estimate for a given reference year. For example, even though the majority of provincial-level detailed expenditure estimates have a CV below 35%, and are thus considered sufficiently reliable to be published, differences in the estimates between two years may not be statistically significant due to the variability in the estimate in each year.

Data validation

Validation in a statistical process is the set of activities which ensures that weighted estimates and aggregate statistics are reliable, sound and defensible. It includes the processes used to identify and correct inconsistencies in both micro-data and aggregate-level estimates through the use of diagnostic tools and subject matter expertise. Statistics Canada validated the SHS 2017 survey estimates, as required, and in accordance with the agency's standards on data validation. Prior to the survey release, a thorough data certification and analysis process implemented a series of detailed checks to ensure that the SHS produces quality estimates.

The certification process covers four main aspects:

- Expenditure certification – examination of year-to-year changes and historical trends for major expenditure categories at the national level as well as for the provinces and territorial capitals.
- Income certification – analysis of income statistics derived from tax and survey data at the national level as well as for the provinces and territorial capitals;
- Data confrontation – comparison of selected expenditure categories to the Consumer Price Index (CPI) and System of National Accounts (SNA) for Canada, the provinces, and the territorial capitals when applicable;
- Methodology and quality indicators – evaluation of data quality, through historical information on sample size, response rates, slippage rates, imputation rates, and CVs.

Fitness for use of 2017 Survey of Household Spending data

Collecting household expenditure data using a combination of an interview and a diary imposes a high burden on respondents. Although several international statistical agencies have researched ways to reduce this burden, no alternative collection approach has been identified that would allow for a comprehensive picture of all expenditures (from very infrequent to very frequent) from each selected household. Internationally, the diary is still considered the most appropriate tool to collect frequent expenditures such as food purchases. However, total and partial non-response is not negligible for this component, which is used to estimate about one quarter of the household expenditures.

Statistics Canada has applied and developed sound statistical methods at the processing and estimation steps to reduce the risk of non-response bias and sampling error. A very thorough validation process is also applied. Data users must nevertheless take into consideration that estimates for detailed expenditures reported by a smaller portion of the sample or for smaller geographical areas may be subject to higher sampling error and a higher risk of bias.