Canada 1998: Survey Information*

*The information presented here is based on the Robustness Assessment Report on Income Distribution Data for Canada by Philip Giles, Kevin Bishop, Christian Nadeau, Michel Dumoulin, Statistics Canada.

This document includes information about the Survey of Labor and Income Dynamics (SLID)

More documentation about the survey and its concepts in the following pdf-files:

Ca98Datacollection
Ca98Labour
Ca98Demographic

Table of contents:

A. General Characteristics

Name

Survey of Labor and Income Dynamics (1993 – ongoing) (SLID)

Purpose

SLID provides detailed annual data on personal and family incomes. In addition, SLID provides a longitudinal dimension to allow a better understanding of the linkage between labor market activity, income and family composition, and any changes which occur.

Administrative unit responsible

Income Statistics Division
Statistics Canada
5-C2 Jean Talon Building
Tunney's Pasture
Ottawa, Ontario CANADA
K1A 0T6

e-mail : giles@statcan.ca
Phone: 613-951-2891
Fax: 613-951-3253

B. Population, sample size and sampling methods

Population size

30,300,400 persons in 1998;

Population covered

Exclusion	Size of group excluded (individuals)	Likely effects on income analysis
Geographical areas: Persons residing outside the ten Canadian provinces or on an Indian Reserve	100,000 in territories 400,000 on Indian Reserves	Data users must recognize the exclusions, as income distributions are lower in these areas.
Groups defined by place of birth, citizenship, immigration status, nationality or ethnic origin: None	N/A	N/A
Homeless people Persons without a usual place of residence	Small	Minor
People in hospital, care or nursing homes: Excluded only if expected stay is more than 6 months	Small	Minor
People in hostels (students, nurses etc): Not excluded	N/A	N/A
Children's homes: Not excluded	N/A	N/A
Military, police, their families, civilians living in military installations:	Small	Minor

Only those living in barracks are excluded.		
Foreign armed forces, diplomats etc. – Excluded	Small	Minor
Prisoners – Excluded	Small	Minor
Others (*E.g defined by economic activity, age, income level, family size*) No other exclusions	N/A	N/A

The effects on the results for the non-coverage of certain population groups are considered to be non significant.

Sample Frame

The sample frame used is the Labor Force Survey area frame. The SLID survey uses the actual LFS samples: SLID uses households who have rotated out of the LFS to select a new longitudinal panel of respondents every three years.

The overall sampling fraction is about 0.2%. The sample is stratified by province, then sub-provincially by urban/rural and by major metropolitan areas
Overall, the stratification reduces the sampling error. The clustered sample design increases the sampling error.

Sample size and response rates

For SLID, the response rate is actually calculated as the number of longitudinal adults in responding households as a percentage of adults in the panel. A responding household is one in which at least one adult reported data. For the reference year 1997, the response rate for the first panel of SLID (after 5 years of survey) is 83% and the response rate for the second panel (after 2 years of survey) is 87%. These response rates do not incorporate the LFS response rate from which the panels were selected. For the same reference year, the number of persons aged 16 or more in panel 1 was 27,483 and, for panel 2 it was 30,030. The cross-sectional sample includes 31,914 responding households including 61,576 persons aged 16 or more.

Weighting procedure

Sampling weights are available for every person in the sample. Post-stratification is done to province / age / sex counts of individuals. In the future, household and family counts will be added to the post-stratification adjustment.

D. Data collection and acquisition

How are the data collected?

Questionnaires are mailed to respondents to provide them with the opportunity to complete them as their time allows. The current collection approach is to use decentralized computer-assisted telephone interviews, with the interviewers working from their homes. In most cases, one knowledgeable household member will provide information for himself/herself and proxy information for all other household members, if the questionnaire had not previously been completed.

How are the different sources matched?

For SLID income data, respondents are given the choice of responding to a traditional interview or of authorizing Statistics Canada to use the information that they have submitted for income tax purposes. Statistical matching is used to link the list of SLID respondents who have authorized access to income tax files to the administrative income tax files.

When are the data collected?

For SLID, a labor interview is conducted each January, and an income interview each May. Both these interviews use the previous calendar year as the reference period.
Collecting income data in the spring is preferred as income tax forms must be filed by April 30.

Income data, reference period

Income data are collected as an annual amount, corresponding to the previous calendar year.
Monthly receipt of Employment Insurance, Workers' Compensation and Social Assistance are collected for SLID as yes/no flags, with the total received only collected for the year as a whole.

Changes over time

With only a few minor exceptions, SLID has adopted concepts and definitions used by SCF. Comparison studies have indicated very little differences in income distributions between the two surveys. Changes to income variables over time have been necessitated primarily by changes in social programs, and thus in government transfers.

D. Definition of the survey units

Unit of observation

Data are collected independently for each individual. Household and family income data are obtained by aggregating income for all individuals in the unit.

Dwelling

In general terms a dwelling is defined as a set of living quarters. Two types of dwelling are identified: collective dwellings and private dwellings. The former pertains to dwellings, which are institutional, communal or commercial in nature. The latter pertains to a separate set of living quarters, which has private access.

Private dwelling

Private dwelling refers to a separate set of living quarters with a private entrance either from outside the building or from a common hall, lobby, vestibule or stairway inside the building. The entrance to the dwelling must be one that can be used without passing through the living quarters of some other person or group of persons.

Household

A household is generally defined as being composed of a person or group of persons who co-reside in, or occupy, a dwelling.

Economic family

Economic family is defined as a group of two or more persons who live in the same dwelling and are related to each other by blood, marriage, common-law or adoption.

How are individuals assigned to household or family?

Individuals are assigned to a household and family based on two data items. First, individuals are assigned to households based on their usual place of residence. Two individuals can only be in the same household or family if they have the same usual place of residence. The determination of a usual place of residence is not always straightforward, as the examples indicated in the question prove. The second data item required for family assignment is the household relationships.

Family relationships

SLID collects the relationship between every pair of persons in the household.

Sharing of income

SLID does not collect information, which allows the explicit determination of who shares economic resources.

Census family

Statistics Canada's standard definition has recently been modified to follow international standards established by the United Nations. SLID derives census families based on this new definition. Initial studies have indicated that this will not have a significant impact on income data at the census family level.

A census family is now defined as a now-married couple, a common-law couple or a lone-parent with a child or youth who is under the age of 25 and who does not have his or her own spouse or child living in the household. Now-married couples and common-law couples may or may not have such children and youth living with them. By definition, all persons who are members of a census family are also members of the same economic family.

Age of children: While there is no restriction on marital status for the "new" census definition, it does require that children can only be grouped with their parents if they are under age 25. There are no age restrictions for the other two definitions.

Children

Children are defined in many ways for different purposes. To be included in the survey, children must be living with their parents. SLID distinguishes on age: children living with their parents are only in the same census family if they are under age 25.

Income data is not collected for persons under a certain age. For SCF, the cutoff is 15 and for SLID it is 16.

Students and income

Student status is not a consideration in household and family assignment. Income data are collected for all persons aged above the appropriate cutoff. Household and family assignment are based on the criteria outlined in Question 7.8 above.

E. Quality of data

<u>How much of the income data was collected by proxy?</u>

The "true" proxy rate is unknown. Questionnaires are mailed to each respondent for completion in advance of the survey. At collection, a lone household member is typically the contact person for the interviewer. This contact would either relay the details from each completed questionnaire or would provide proxy responses. However, both cases are considered to be proxy responses.

For SLID reference year 1997, 73% of the people in responding households gave permission to access their tax data, and 14% responded to the income interview. The other 13% were imputed. Among the respondents to the income interview, 44% of the data were collected by proxy (according to the broad definition of proxy).

<u>Non-response</u>

It is suspected that people with high income are less inclined to answer. For SLID, a segmentation modeling approach is used to correct for non-response, using that data, which are known for the non-respondents.
The comparisons with external data sources are minimal and they did not indicate any obvious non-response bias.

For SLID, renters and young single males are those for which the non-response rate is the highest. The non-response adjustment model compensates for these.

<u>Income variables with a high risk for non-response bias.</u>

Investment income is subject to non-response bias. Recipients of large amounts of investment income, representing a substantial proportion of the total for investment income, tend to be a small group of individuals concentrated in the upper end of the distribution. Not only is their representation in the distribution underestimated due to under-representation in the sample, there is the possibility that they may tend to be more likely not to respond when sampled for the survey. For middle and lower income individuals, there is likelihood that a small amount of investment income (e.g. Small amounts of interest from bank accounts) goes unreported.

<u>Income items with high non-response rates</u>

For SLID, Child Tax Benefit (CTB) and Goods and Services Tax (GST) (33% imputation rate), and Old Age Security Pension, Guaranteed Income Supplement and Spouse's Allowance (OASGI) (21% imputation rate) are the highest. However, these rates include the total imputation of income variables for non-respondents in responding households.

<u>Categorical variables affected by non response</u>

For SLID, item non-response rates vary around 3 % for most categorical variables.

Imputation

A deterministic approach is used for government transfer programs while a hot-deck technique is used for labor variables.
The use of the previous year data to impute missing values of income is the method used whenever possible. Such a method may lead to an under-estimation of income mobility.

How is the net income reported?

For SLID, Tax Payable is imputed though a regression model for those not giving the permission to access their tax data and for those for which it was not possible to do so. They constitute 27% of the SLID sample. As a result, there is no significant impact on the resulting after-tax income distributions.

Has there been other editing besides imputation?

SLID: Apart from complete record imputation for 15 % of the sample due to income non-response there are no other edits affecting more than 5% of respondents. At the data collection stage, income range checks, total/summation checks and internal consistency checks are performed.

Which income items are most sensitive to imputation?

Based on informed judgment, the low-income rates for specific population subgroups are thought to be the most sensitive to editing and imputation procedures. For the elderly, the income distribution is clustered around the low-income cut-offs, due to the nature of the Old Age Security benefits. For this group, relatively small changes in total income resulting from edits/assignments can lead to relatively large differences in the low-income rates.

How much of the data on earned income was checked against employers' records?

No income data are obtained from the employer. However for SLID, if the respondent gives permission, all income data are taken from the administrative files containing individual income tax returns.