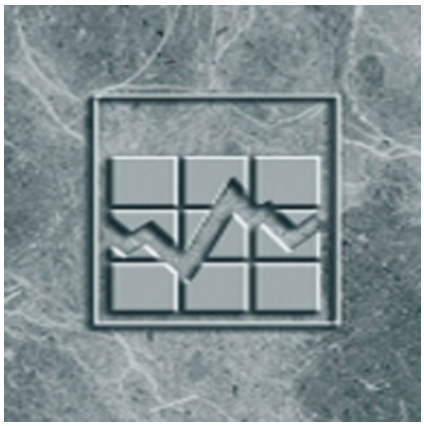
Catalogue no. 75F0002MIE — No. 009
ISSN: 1707-2840
ISBN: 0-662-40760-1

Research Paper

Income research paper series

Survey of Labour and Income Dynamics: 2003 Historical Revision



by Heather Lathe

Income Statistics Division
Jean Talon Building, Ottawa, K1A 0T6

Telephone: 613 951-7355

Statistics Canada Statistique Canada

Canada

How to obtain more information

Specific inquiries about this product and related statistics or services should be directed to: Income Statistics Division, Statistics Canada, Ottawa, Ontario, K1A 0T6 (telephone: (613) 951-7355; (888) 297-7355; income@statcan.ca).

For information on the wide range of data available from Statistics Canada, you can contact us by calling one of our toll-free numbers. You can also contact us by e-mail or by visiting our website.

National inquiries line	1 800 263-1136
National telecommunications device for the hearing impaired	1 800 363-7629
Depository Services Program inquiries	1 800 700-1033
Fax line for Depository Services Program	1 800 889-9734
E-mail inquiries	infostats@statcan.ca
Website	www.statcan.ca

Information to access the product

This product, catalogue no. 75F0002MIE, is available for free. To obtain a single issue, visit our website at www.statcan.ca and select Our Products and Services.

Standards of service to the public

Statistics Canada is committed to serving its clients in a prompt, reliable and courteous manner and in the official language of their choice. To this end, the Agency has developed standards of service that its employees observe in serving its clients. 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.ca under About Statistics Canada > Providing services to Canadians.

Statistics Canada
Income Statistics Division

Income research paper series

Survey of Labour and Income Dynamics: 2003 Historical Revision

Published by authority of the Minister responsible for Statistics Canada

July 2005

Catalogue no. 75F0002MIE, Vol. 9

Frequency: Occasional

ISSN: 1707-2840

ISBN: 0-662-40760-1

Ottawa

La version française de cette publication est disponible sur demande (nº 75F0002MIF au catalogue).

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 cooperation and goodwill.

Abstract

The release of the 2003 data from the Survey of Labour and Income Dynamics (SLID) was accompanied by a historical revision which accomplished three things. First, the survey weights were updated to take into account new population projections based on the 2001 Census of Population, instead of the 1996 Census. Second, a new procedure in the weight adjustments was introduced to take into account an external source of information on the overall distribution of income in the population, namely the T4 file of employer remittances to Canada Revenue Agency. Third, the low income estimates were revised due to new low income cut-offs (LICOs). This paper describes the second of these improvements—the new weighting procedure to reflect the distribution of income in the population with greater accuracy. Part 1 explains in non-technical terms how this new procedure came about and how it works. Part 2 provides some examples of the impacts on the results for previous years.

Acknowledgements: The author wishes to thank several colleagues at Statistics Canada for their contributions, comments, and assistance during the writing of this paper. In Income Statistics Division they are: Adam Wronski, Eric Olson, Stephanie Bonnell and Chris Li. In Social Survey Methods Division they are: Sylvie Laroche and Caroline Cauchon.

Table of contents

Summary

The following summary is also printed in *Survey of Labour and Income Dynamics: A Survey Overview, 2005*.

The release of the 2003 data is accompanied by a historical revision for 1990 to 2002 due to an update of the survey weights in both the Survey of Labour and Income Dynamics (SLID) and the Survey of Consumer Finances (SCF). The weighted data now take into account not only new population projections based on the 2001 Census of Population, but also valuable information on the overall distribution of wages and salaries—a major component of income—in Canada. A lot of research has been conducted on the new weighting methodology and it is felt that as a result of this, the ability of these surveys to accurately represent income levels in Canada has been given a new boost. The quality of the survey data on other topics has also been improved.

The impacts of the historical revision, which was applied to the data back to 1990, are quite similar for all years, so that the trends shown by the data remain very similar. But the actual levels of estimates do change in many situations. The estimates are in fact much closer to other reliable sources of information on similar topics.

The 2003 historical revision also incorporates revised 1992-based low income cut-offs (LICOs) resulting from a historical re-weighting of the 1992 Family Expenditure Survey.

Why we introduced new income adjustments in the weighting

Research has shown that historically, Statistics Canada's household income surveys have tended to under-represent people with very low levels of earnings or no earnings, as well as people with very high earnings, while over-representing people who are more in the middle of the earnings distribution. Average and aggregate earnings and incomes have been over-estimated as a result.

The System of National Accounts (SNA) corrected a comparable problem by applying direct adjustments to the estimates. In a household survey like SLID, such an approach is not possible; instead, the method for correcting the biases in estimation was to make differentiated adjustments to the weights attached to each of the respondents.

Statistics Canada did a lot of investigation over the past few years to develop the best possible options for improving the estimates through improved weighting techniques. We tested several options and evaluated their stability over several years of data, to ensure that no new biases would result from the corrections and to ensure that we chose the best possible option currently available. Finally, the evaluations included comparisons with other independent sources of information like the Census and the System of National Accounts. Apart from the time required to test and evaluate the changes, there is no particular reason why Statistics Canada introduced these changes now instead of next year or last year.

In addition to our usual adjustments to population estimates we concluded that the T4 administrative files (employer remittances of payroll taxes) were the best benchmark statistics available with which to adjust the weights in SLID and SCF. We believe that the quality of the T4 files is high and that historically, the quality of files is sufficient to allow this adjustment to be made back to 1990.

The improvements to survey weights during the 2000 and 2003 historical revisions have been part of a comprehensive project at Statistics Canada to improve the weighting strategies in the

main annual surveys on income, expenditures, and wealth. The project is now complete for the Survey of Labour and Income Dynamics and the Survey of Consumer Finances.

What the new weighting methods consist of

Survey weights are the numbers assigned to each of the sample respondents so that together they sum to the target population and sub-groups of that population. To do this as accurately as possible, the weights are often adjusted to be in line with other independent sources of information. The Census of Population, with its related population projections, is the source for the benchmark demographic statistics used to adjust the survey weights. The revised weights now take into account population projections from the 2001 Census.

The additional benchmark statistics that are now used to adjust survey weights come from the annual T4 file from Canada Revenue Agency (CRA), which is a compilation of employer remittances for the purposes of payroll taxes. SLID obtains tabulated statistics from the T4. Specifically, the statistics being used concern the total number of people in each province who had earnings from paid employment during the year and the amounts they made relative to each other, called the earnings distribution. The number and earnings of self-employed people are not included. It is important to add that aggregated data are used from this file and no matching of individual information between the T4 file and the SLID and SCF data files is done.

The annual T4 files cover all employees, regardless of whether they filed a T1 tax return. The wages and salaries of employees represent a very high proportion of all income received by individuals. Although many people in the population do not have this type of income, the T4 data allow us to correct biases not only for the wage-earning population but also for the rest of the population, because it provides an accurate way of determining the presence or absence of wages.

This is not the first time that a survey has used the T4 files from CRA in order to better represent the distribution of incomes in the population. The Survey of Household Spending implemented similar weighting methods using T4 data in its 1999 historical revision.

As evidence that the additional calibration of weights has helped to better reflect the distribution of incomes in the SLID, the estimates are now more similar to related estimates from other sources. These sources are not only the T4 file, but also the T1 file (personal income tax file), the Census, and the System of National Accounts.

What was the impact of the revision on the estimates?

The change from the 1996 Census to the 2001 Census for population estimates and projections affected mainly the period after 1996.

The historical revision had little impact on the trends shown by SLID since 1996 and SCF up to 1995. This is because the addition of weight adjustments based on income had a similar impact for all years. The weight revision was taken back to 1990; consequently, only between the years 1989 and 1990 is there potentially a disturbance in the data series that could affect the representation of trends.

The level estimates, however, did change in many cases. The precise impacts can be seen by comparing the data that were published before the historical revision with those published in May of 2005. Note however that previously published dollar figures must be compared with their

revised values using either current dollars or consistent constant dollars in both cases. Be aware also that rounding of dollar values in published tables may affect the precision of comparisons.

Finally, the SLID low income estimates changed as a result not only of the new weights in SLID, but also the new low income cut-offs with which family incomes in SLID are compared.

Examples of how the estimates changed

- In general, the number of people in the ten provinces with earnings from employment, also referred to as wages and salaries because it excludes self-employment, was adjusted downwards. However this is not true in all range categories of earnings: there are more people now in the lowest and highest earnings categories, but fewer in the middle of the distribution.
- Due to the predominant downward revisions in the number of persons by earnings classes and despite the upward revision in the number of high-wage earner adults, average and median wages and salaries usually decreased with the revision.
- Since wages and salaries constitute a large part of incomes, most estimates of market income, total income, and after-tax income were also adjusted downwards.
- As a consequence of lower wages and salaries, government transfers generally increased, including social assistance benefits, child benefits, and refundable sales tax credits such as the GST/HST Credit.
- The weight revisions due to the new income adjustments based on the T4 file affected estimates concerning children and adults up to age 64 relatively more than estimates concerning seniors, due to the fact that income from employment makes up a smaller portion of seniors' incomes.
- Consistent with the increases in the number of people with no earnings or very low earnings, the low income rates for persons and families (on an after-tax basis) generally underwent an upward revision in all years. Again the rates for senior-led families and individuals were less affected.
- Consistent with the increases in the number of people at the very low and very high ends of the earnings distribution, and the decrease in the number of people in the middle of the distribution, certain measures of income inequality (Gini coefficient, ratio of 5^{th} to 1^{st} quintile) generally increased.
- The revision also affected characteristics which are not directly related to incomes. For example, the following estimates are now closer to the estimates of the 2001 Census:
    - Level of educational attainment in Canada.
    - Rate of home ownership versus renting.

Introduction

The release of the 2003 data from the Survey of Labour and Income Dynamics (SLID) was accompanied by a historical revision which accomplished three things. First, the survey weights were updated to take into account new population projections based on the 2001 Census of Population, instead of the 1996 Census. Second, a new procedure in the weight adjustments was introduced to take into account a high-quality external source of information on the overall distribution of wages and salaries in the population—and by extension the distribution of incomes as well. Third, the low income estimates were revised due to new values of the 1992-based low income cut-offs (LICOs) following a historical re-weighting of the 1992 Family Expenditure Survey. The data in SLID were revised for 1993 to 2002, and the data from the Survey of Consumer Finances (SCF) were revised from 1990 to 1997. Published tables containing the SCF-SLID time series typically start with 1980. SLID provides data for years 1996 and later while SCF provides the historical time series on income content up to 1995.

This paper describes the second of these improvements—the new weighting procedure to reflect the distribution of income in the population with greater accuracy. Part 1 explains in non-technical terms how this new procedure came about and how it works. Part 2 provides some examples of the impacts on the results for previous years. (Part 1 also contains some charts illustrating the impacts, in the section called “Evaluation”.)

Although this paper focuses on the SCF and SLID, the Survey of Household Spending (SHS) has also used the T4 files from Canada Revenue Agency in order to better represent the distribution of incomes in the population. The SHS introduced similar weighting methods using T4 data in its 1999 historical revision, and will introduce similar demographic calibration in its 2004 revision. Part 1 is therefore relevant for that survey as well as SCF and SLID. The impacts of the 1999 historical revision of SHS are documented in a separate research paper.[1]

1. Willa Rea and Edith Greenlee, “The Effects of the Revised Estimation Methodology on Estimates from Household Expenditure Surveys”, (forthcoming).

Part 1
Income calibration based on administrative data

1.1 The purpose of weight calibration

Weight calibration provides a relatively inexpensive way to improve data quality in surveys, provided other higher quality data sources exist on similar characteristics. Furthermore, unlike most other types of improvements in a survey, it can be used to improve the survey results back in time as well as for the future.

Different methods of data benchmarking exist. Sometimes problems of over-estimation or under-estimation can be dealt with by applying direct adjustments to the estimates. But in a household survey like SLID, such an approach is not possible; instead, the method for correcting the biases in estimation is to make differentiated adjustments to the weights attached to each of the respondents.

There are different procedures for weighting in different surveys, but usually the three basic steps are: calculation of basic weights; non-response adjustment; and weight calibration. In order to help explain how calibration works and its role in heightening data quality, a brief description of the general weighting procedures is presented here. A thorough description of the weighting procedures in SLID can be found in the research paper “Longitudinal and Cross-Sectional Weighting of the Survey of Labour and Income Dynamics 1997 Reference year” (Lévesque and Franklin; an update of that paper is forthcoming.)

The basic weights in the surveys of Income Statistics Division are derived from the Labour Force Survey (LFS) weights, because these surveys obtain their samples from the LFS sample. The LFS final weights are converted to basic weights for SLID by taking into account the fact that SLID only uses about one-third of the LFS sample and that of that portion it only attempts to contact those who answered to the LFS.

The “non-response adjustment” then takes into account non-response. In this adjustment, the complete non-respondents are given final weights of zero. Their basic weights are then summed and re-distributed among the persons or households whose answers were found to be complete, or at least partially complete. (Sufficient partial response to the interview is defined according to a minimum set of completed questions.) That is, non-response is compensated for by increasing

the weights of the respondents who participated in the survey. Note that SLID's basic weights already take into account non-response in the LFS.

The non-response adjusted weights in SLID are then further adjusted—this time upwards or downwards and to different degrees—to ensure that estimates on relevant population characteristics respect population totals of another data source (or sources) which is considered very reliable. This is called "calibration.

The source of benchmark demographic data for the household surveys is the Census of Population. Since the Census is only conducted every five years, Statistics Canada "projects" the Census results for later years (up to the present), and then revises those estimates when the next Census becomes available. The projections use a variety of secondary information, including administrative data on births, deaths, and migration.

In the calibration step, weighted totals based on the responses are first calculated according to various cross-tabulations of variables, called "classes". The variables are typically province, age, and sex. The same totals are then calculated using the Census and the related population projections. Then the technique of generalized least squares estimation is used to determine which groups of sample respondents (by province, age, and sex) should get upwards or downwards adjustments in their weights, and by how much.

The idea behind weight calibration is that if the weighted sample can be made more representative according to a few key characteristics of the population, it will also be more representative for a host of other variables which are *related* to those key characteristics.

However, this assumption comes with a lot of caveats or "conditions". It depends on there being a strong association between the controlled and uncontrolled characteristics in the sample and not just in the population it was drawn from. Any so-called "distortions" between characteristics as reflected in the sample will not be corrected when the data are re-weighted after calibration—in fact, they are liable to be worsened.

Even with a minimum of such distortions, there is also the necessity of satisfying the various benchmarks simultaneously. The more benchmarking classes there are, the less likely it is that a solution exists which respects all of them without generating unacceptable weight values. An obvious sign that calibration has been over-applied is that some of the weight values that get generated are negative.

Something which calibration cannot do is to ensure the representation of characteristics which exist very infrequently in the population. Even a well designed sample cannot reliably include respondents that touch on the population's rarest characteristics. No amount of weight adjustments will ensure that such characteristics appear in the weighted data.

1.2 Project to harmonize weight calibration

The improvements to survey weights in SLID and SCF during the 2003 historical revision were part of the Harmonized Calibration of Income Statistics Project. Its objective was to improve the comparability of income estimates and distributions across Statistics Canada surveys and with external sources (Tremblay, 2003). Three other surveys are also involved in the project. They are the Survey of Household Spending (SHS) which is carried out since 1997, the Survey of Family Expenditures (FAMEX) which preceded it, and the Survey of Financial Security (SFS, conducted for 1999 and 2004).

A steering committee consulted with external advisors, namely the Advisory Committee on Labour and Income Statistics, and the Advisory Committee on Statistical Methods. Presentations were made to this latter committee in April 2000 (Webber et al.) and November 2003 (Latouche and LaRoche).

1.3 Early focus on the Survey of Labour and Income Dynamics

During the mid-1990s, as part of regular data quality evaluation, Statistics Canada examined and compared the estimates on income from the household surveys. The Survey of Consumer Finances (SCF) was still the main survey for annual income statistics, while the Survey of Labour and Income Dynamics (SLID) had just been introduced as a major new longitudinal survey designed to produce both longitudinal and cross-sectional statistics. The income component of SLID was closely modelled on the SCF in terms of questionnaire design and income concepts.

After SLID had been successfully implemented for a few years, consideration was given to using it to replace the Survey of Consumer Finances. However, there were in particular two concerns that had to be resolved before this could be done. First, it was imperative to ensure that SLID produced high quality estimates of income not only for its longitudinal panel of respondents, but also for annual cross-sections of the population. (See Box 1.3 below, "Maintaining a cross-sectional sample in a panel survey".) Second, it was necessary to describe clearly the impact from using the T1 file as a partial source of income information on respondents. The T1 file is

produced by Canada Revenue Agency and contains data on the T1 Income Tax and Benefit Return that individuals complete. In SLID, unlike SCF, respondents can give Statistics Canada permission to use their T1 tax information for the purposes of the survey. Over 80% of SLID's respondents give their consent to use their information from the T1 file.

Due to such concerns, both SLID and SCF continued to be conducted for 1996 and 1997, by which time the SLID sample size was equivalent to that of SCF. Their results were extensively compared before Statistics Canada decided to discontinue the SCF.

When the two income surveys were compared,[2] it was found that the estimates of total income (market income, after-tax income) were reasonably similar, as were the estimates of aggregate wages and salaries. In 1996, SLID recorded aggregate wages and salaries that were 1.1% lower than the SCF figure. In 1997, the SLID figure was 1.7% higher. Furthermore, the distribution of income was very similar between the two surveys. Only at the lowest part of the distribution was there a significant difference: SLID captured more people with very small income amounts, including wages and salaries, while SCF had more people with no income. This difference was attributed to the use of T1 data for a large part of the SLID sample.[3]

Box 1.3
Maintaining a cross-sectional sample in a panel survey

In SLID, the cross-sectional sample is derived in large part from the longitudinal sample. Each longitudinal sample, also called a "panel" of respondents, constitutes a cross-section of the population for the year the panel started. As the longitudinal respondents are re-interviewed every year over six years, a current cross-sectional sample of the population is maintained by interviewing all the *new* people found to be living with a longitudinal respondent at that time.

The theoretical basis for using a panel survey to produce cross-sectional estimates the way that SLID does had been established (Lavallée, 1995), but it had not previously been put into practice. One of the concerns with SLID was attrition, i.e., the fact that non-response becomes cumulative in a longitudinal survey. When the same sample of respondents is contacted each year for an interview, there is a growing portion of the sample respondents who become either unreachable (through changes of address and telephone number) or unwilling to continue participating in the survey. Attrition could diminish the quality not only of the longitudinal information, but also the cross-sectional results.

2. Webber et al, April 1999; Cotton et al, December 1999.
3. Ibid.

1.4 Comparisons of survey data with administrative data

One of the tests that had long been used to assess the accuracy of the Survey of Consumer Finances was to compare the estimates of aggregate annual wages and salaries (earnings from employment) and aggregate pre-tax or "total" income with related estimates from the System of National Accounts (SNA), which produces the figures on Gross Domestic Product. For this exercise, the data are adjusted ("reconciled") as much as possible for quantifiable differences in their definitions and coverage of the population. SCF and SLID data were also compared with the estimates from administrative files produced by Canada Revenue Agency (CRA), namely the T1 and T4 files. As mentioned in the previous section, the T1 file contains data from the Income Tax and Benefit Return that individuals complete. The T4 file contains data from the T4 Statement of Remuneration Paid which employers complete. Statistics Canada uses it to produce estimates of total labour income in Canada, for the System of National Accounts.

Table 1.4 and Chart 1.4 illustrate the differences between the SCF-SLID time series on aggregate employment earnings and that of the SNA. Between 1980 and 1991, aggregate wages and salaries in SCF were 1.6% to 3.6% higher than the same SNA estimates, after adjustments. Starting in 1992, the gap widened. It fluctuated between 3.1% and 6.2% over the years 1992 to 2002.

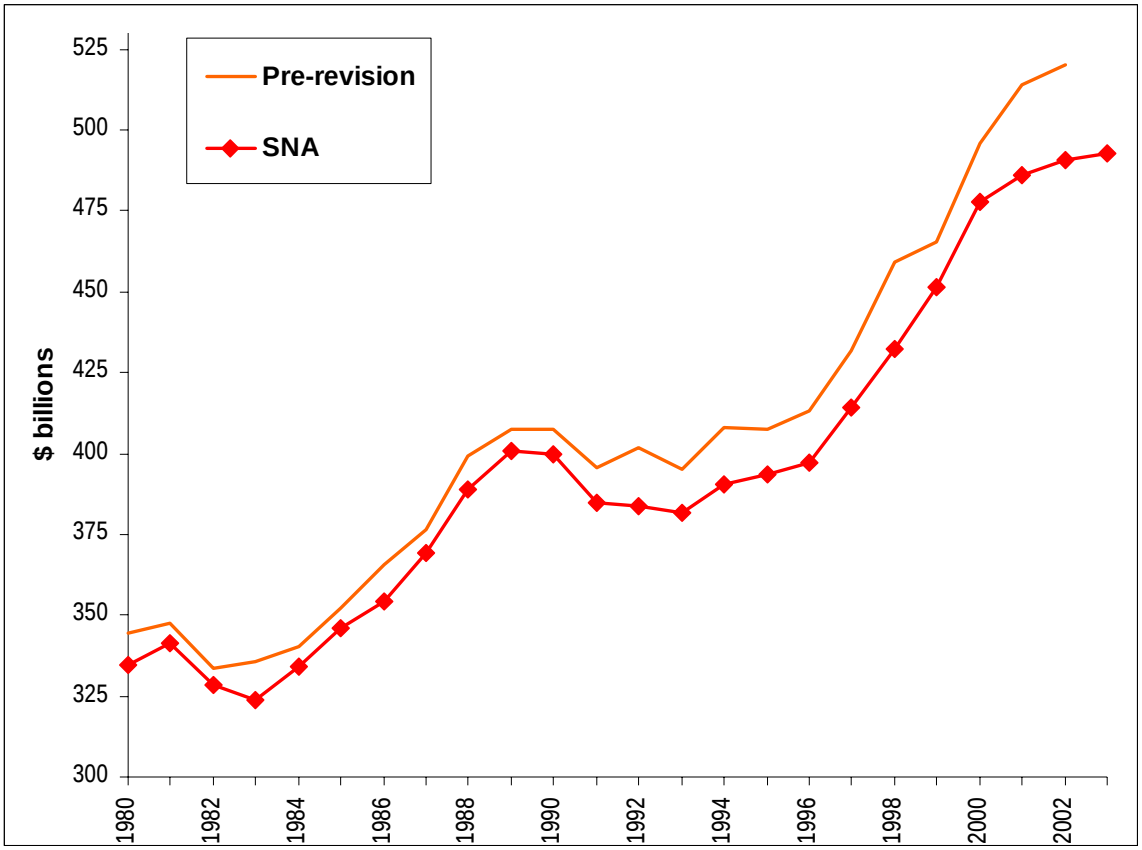
Table 1.4 also shows the percentage difference between the T4 file estimates and the SNA estimates starting in 1990. The two sets are very close because the SNA figures are in large part derived from the T4 file. For the same reason, the summed data from Line 101 (employment income) of the T1 file are very similar to the totals from the T4 file.

Table 1.4: Aggregate wages and salaries, All provinces, 1990 to 2002

	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002
SNA[1] (in 2000 constant dollars)													
($ billions)	400	385	384	382	391	394	397	414	432	452	478	486	491
Difference from SNA (%)													
T4 file	0.3	0.3	0.1	0.0	-0.2	0.5	0.5	0.5	0.6	0.7	1.0	0.6	0.7
SLID-SCF, pre-revision	1.9	2.9	4.7	3.5	4.4	3.5	4.1	4.2	6.2	3.1	3.8	5.7	5.9

1. System of National Accounts (SNA). Source: CANSIM tables 382-0001 and 382-0006, plus a constant adjustment factor to account for quantifiable differences in definitions and population coverage as compared to SCF and SLID.

Chart 1.4: Aggregate wages and salaries, Canada, 1980-2003, in constant dollars of 2000



The distribution of wages and salaries was also examined to identify the differences between the survey and administrative data more precisely. This type of comparison is not possible with the System of National Accounts, which only has estimates of aggregate earnings, but it can be done with the T1 and T4 files, which contain the necessary “microdata” to derive estimates by ranges of earnings levels.[4]

It was observed that, for all the years back to 1980, SCF and SLID overestimated the number of people and families with incomes towards the middle ranges of the income distribution and underestimated the number of people and families with incomes at the very high and very low ends of the income distribution.

For example, in 2000, SLID underestimated the number of people with annual wages and salaries over $90,000 and below $5,000 (approximately), as compared to the T4 file (not shown).

4. Statistics Canada undertook a careful examination of the T4 files to identify and correct obvious errors in individual records on the files. Statistics Canada continues to apply some editing of the files each year as deemed necessary, prior to deriving estimates from them

The number of people in the upper range counts for only a small percentage of all adults. The number of people with low or no wages and salaries is much larger. Overall, the overestimation across the middle of the distribution dominated the other impacts in terms of the aggregate value of wages and salaries. (Evidence of the distributional difference between SCF-SLID and the T4 file is also shown in Section 1.8 below, Chart 1.8.)

The fact that both SLID and SCF revealed the same types of differences with the administrative data showed that the problem could not be attributed to the different collection and sampling methodology in SLID as compared to SCF. It was potentially a more general phenomenon among household surveys. This is the subject of the next section.

1.5 Non-response as a possible source of the discrepancy

A trend of increasing rates of non-response has been observed in several household surveys at Statistics Canada. In longitudinal or "panel" surveys, such as SLID, a cumulative increase in non-response is normal because the same respondents are followed for several years. But the trend of increasing non-response in the cross-sectional surveys was not easy to explain. It has been monitored for some time now, and the reasons seem to include a combination of changes in survey operations and changes in society**.**

Specific efforts are being made to combat non-response in SLID, including some innovative and cost-effective ways to improve tracing of longitudinal respondents as they change address over time. There is evidence that they are being successful, because the decline in the SLID response rate has been stopped. However, although they might reduce the problem, they will probably not be able to bring response rates back to the levels of the early 1990s.

The discrepancy regarding the income distribution between the SCF-SLID data and administrative sources existed at least as far back as 1980. But the widening gap since 1992 might be related to declining response rates. In other words, biases due to non-response may have affected the representation of the income distribution.

An understanding of the typical nature of non-response to income surveys can explain this hypothesis. Non-response can arise either from a failure to contact respondents or refusal to participate among those who are contacted. It has long been known that non-response tends to be concentrated among people who change address and telephone number relatively more often and people who are less likely to be at home and available when an interviewer is likely to call. It is also recognized that surveys with a major focus on personal finance can face challenges due to

the sensitivity of the topic to most people, especially people with either very high or very low incomes. Non-response for either reason could contribute to a concentration of non-response at the high end of the income distribution and at the lower end of the income distribution. We don't always know the current characteristics of the non-respondents, so imputation as a method of correcting the problem is limited.

1.6 Making room for income calibration

The HARC Project included improvements in demographic calibration in the relevant surveys, as well as income calibration. First, in the years leading up to the 2000 historical weight revision, much work was done to see how some shortcomings *not* related to the income estimates could be corrected using better demographic calibration. (As explained in Section 1.1, calibration can sometimes worsen existing distortions in the sample data for uncontrolled characteristics, so problems on the demographic side might otherwise have been worsened by the income calibration.) Then later, in the 2003 historical revision, the number of age-sex categories was reduced, to allow for the addition of several income categories.

It had been observed for some time that significant discrepancies existed between the different household surveys regarding family and household counts. After careful examination, it was concluded that the existing demographic calibration did not fully resolve some biases regarding people's living arrangements. The controls on age groups and sex were effective in terms of the number of individuals, but households or families are also important as units of analysis in SLID and they were not adequately controlled via the person counts, at least not in the relatively recent period.

The 2000 historical weight revision addressed this, by adding two more variables in the calibration, namely family size (i.e., the number of people living together who are related to each other) and household size (i.e., the number of people living together regardless of whether they are related or unrelated). The following categories are controlled via calibration: families of one person; families of two persons; households of one person; household of two persons. Families and households of 3 persons or more are not controlled but are residual categories. Note that data on family size and household size did not become available in Statistics Canada's population projections until after the 1996 Census. But they were also added for inter-censal estimates back to 1976, thus providing the necessary benchmark data to include this new calibration in the SLID and SCF data back to 1980.

Regarding the 2003 revision and the reduction of the number of age-sex categories, there were initially 22 categories for each province (220 classes in total), which were obtained by crossing the following 11 age groups by sex: 0-6, 7-15, 16-17, 18-24, 25-34, 35-44, 45-54, 55-59, 60-64, 65-69 and 70+. They were reduced to 15, as follows. First of all, it was felt that the youngest age group did not need to be calibrated by sex, so it now accounts for just one age-sex category. Secondly, the 16-to-17 age group, being very narrow, was combined with the group of ages 7 to 15. The 55-to-59 and 60-to-64 age groups were combined, as were the 65-to-69 and 70+ age groups. The new categories are as follows: 0-6, 7-17, 18-24, 25-34, 35-44, 45-54, 55-64 and 65+.

1.7 Development of income calibration

Although the research showed that both employment income and total income were over-estimated, the development of the new calibration technique focused on employment income. Wages and salaries of employees represent a high proportion of all personal income received in Canada, so any adjustments at the level of earnings can be expected to have an impact on estimates of total income. Fortunately, the concept of wages and salaries is quite similar among surveys and with tax data.

The annual T4 file was favoured over the T1 file because of better coverage and timeliness. The T4 file includes employees regardless of whether—or when—they file a tax return. (Later versions of the T1 file are produced by CRA, and Statistics Canada took the time to evaluate them as well; the coverage of those files was found to be higher than that of the preliminary T1 file, but still lower than the T4 file.)

Of course, the T4 file is not expected to have complete coverage either. Many jobs exist which are not covered by a formal arrangement with a registered employer, and are therefore not covered by a T4 slip. This is discussed further in the next section. There may also be general administrative errors on the part of employers or CRA.

During development of the 2000 weight revision in SLID, it was concluded that more analysis was necessary to assess the reliability of the control totals based on the T4 file. The assessment of the T4 files was then carried out on files going back to the 1980s. It was concluded that the quality of files was sufficient to allow the historical revision using the T4 files to be made back to 1990.[5]

5. Because the Survey of Household Spending was able to go back only to 1997 with its historical weight revision, the quality of the T4 files back in time was not as great a constraint for that survey as it was for SLID and SCF.

The statistics being used concern the total number of people in each province who had earnings from paid employment during the year and the amounts they made relative to each other, called the earnings distribution. This is described in the next section.

It is important to add that aggregated data are used from the T4 file and no matching of information on individuals between the T4 file and the SLID or SCF data files is done.

1.8 Categories for earnings calibration

The earnings distribution can be defined in a few different ways. For example, after categorizing individuals according to the value of wages and salaries they each received in the year, we can then count the number of individuals in each category, or go a step further and sum up each group's aggregated earnings. Evaluations showed that calibrating on the number of earners in each category would be better than calibrating on aggregated earnings. The number of earners is more stable from year to year than their corresponding aggregated earnings, particularly at the high end of the distribution.

The earnings categories can be defined using either nominal values (in this case, dollars) or percentiles (a "relative" measure). Percentiles have some advantages. Unlike dollar values, they do not need to be adjusted for inflation. The ranges can be defined the same way and yet represent different dollar ranges in each province, simply by recalculating them separately for each province. And in the same way that they adjust automatically for the distribution by year or province, they adjust for such differences between data sources as well.

After evaluating many options, it was determined that seven categories of earners would be used, defined primarily as percentile ranges, as described further below. There is also a residual category, to allow for the fact that the total adult population is already controlled in the demographic calibration. It is described first.

As mentioned in the previous section, some jobs are not covered by a formal arrangement with a registered employer, and are therefore not covered by a T4 slip. It is likely that many of them are casual in terms of schedule and would therefore account for low amounts of annual earnings. It is not certain whether the surveys capture all of these jobs and corresponding earnings, either: most likely there is some under-coverage in either source for very small amounts of annual earnings, but the exact range or extent of under-coverage cannot be assumed to be the same. Therefore, it was decided to combine adults with no earnings from employment and adults with small earnings,

by putting them both in the residual category. For this reason, the first calibration category begins at $1500 (for all provinces), which is adjusted for each year to be $1500 in constant dollars of 1997.

The lower boundaries for the five middle categories are the 10th, 25th, 50th, 65th, and 75th percentiles. For the seventh and last category, the lower boundary is either the 95th, 98th, or 99th percentile. If the sample size in the province is too small to permit calibrating on just the 99th percentile, then it is re-evaluated at the 98th percentile, and if still not large enough the last category starts at the 95th percentile. This helps to maintain the weights of the very highest wage earners in the sample who, because of their large values, need to be treated with greater precision than people whose values have less influence on aggregate and average income.

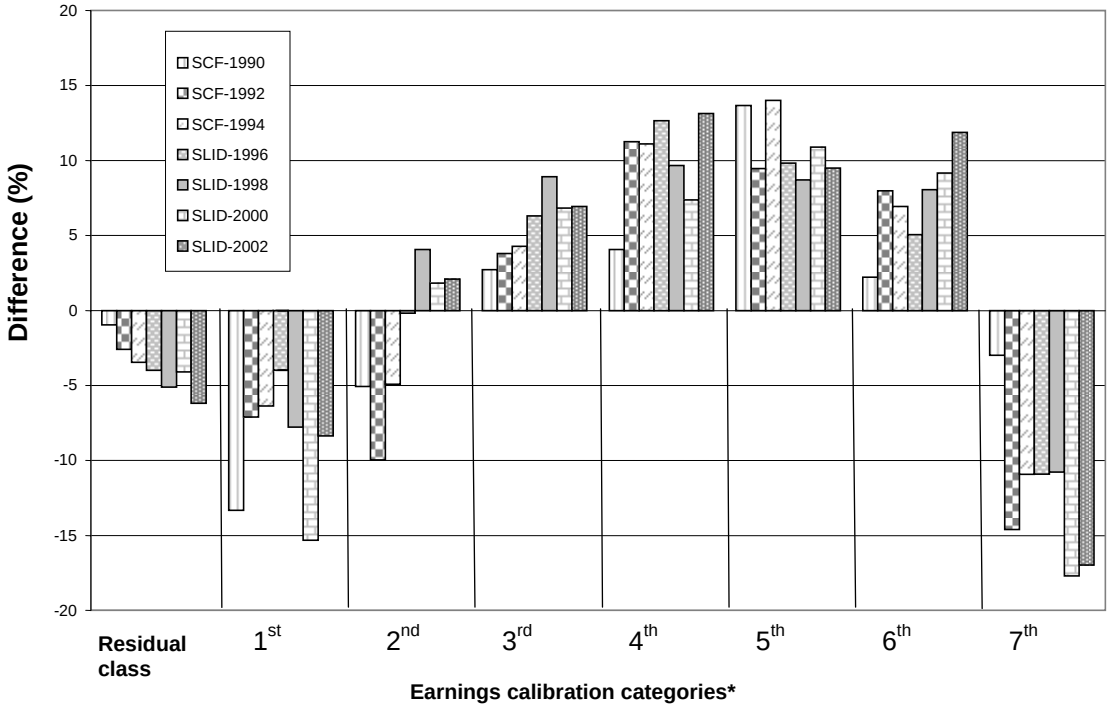
Table 1.8: Categories of annual wages and salaries used in the income calibration in SLID[1]

residual category	Adults with $0 wages and salaries, and up to $1,500
1	Adults with at least $1,500, up to 10th percentile
2	Adults from 10th percentile to 25th percentile
3	Adults from 25th percentile to 50th percentile
4	Adults from 50th percentile to 65th percentile
5	Adults from 65th percentile to 75th percentile
6	Adults from 75th percentile to 95th, 98th or 99th percentile
7	Adults from 95th, 98th or 99th percentile to 100th percentile

1. The Survey of Household Spending uses the same calibration categories except that categories 1 and 2 are combined.

Chart 1.8 presents the seven calibration classes plus the residual class in a bar chart of the percentage differences between SCF or SLID data and the T4 file estimates, for various years. In each category except one, the direction of the difference is the same in all years. In the category labelled $1500-P10, which is the exception, the difference is negative in all years shown for SCF and positive for all years shown for SLID except the year 1996 for which the difference is negligible. This is probably related to the observation stated in Section 1.3 that SLID and SCF differ in one main respect: SLID records a larger number of people with small amounts of earnings, while SCF recorded larger numbers of people with no earnings. All the differences shown in the chart largely disappear after the data have been calibrated based on the T4 file.

Chart 1.8: Aggregate wages and salaries: difference between pre-revision SCF-SLID estimates and T4 file estimates, by earnings calibration categories[1], Canada, selected years



1. See also Table 1.8 for explanations regarding the earnings categories.

Box 1.8
Projection of the T4 file

The T4 file is not available in time to do calibration each year (for the current year), so the previous year's file is used, and the required calibration statistics are projected one year forward. The growth rate used for projecting the number of earners in each calibration category by province is obtained either from the Census-based population estimates (adults only), or from average weekly earnings in the Labour Force Survey. An evaluation showed that the population growth rates produced more accurate projections for the first five earnings categories—those that cover earners up to the 75th percentile, while the weekly earnings growth rates produced better projections for the two highest categories.

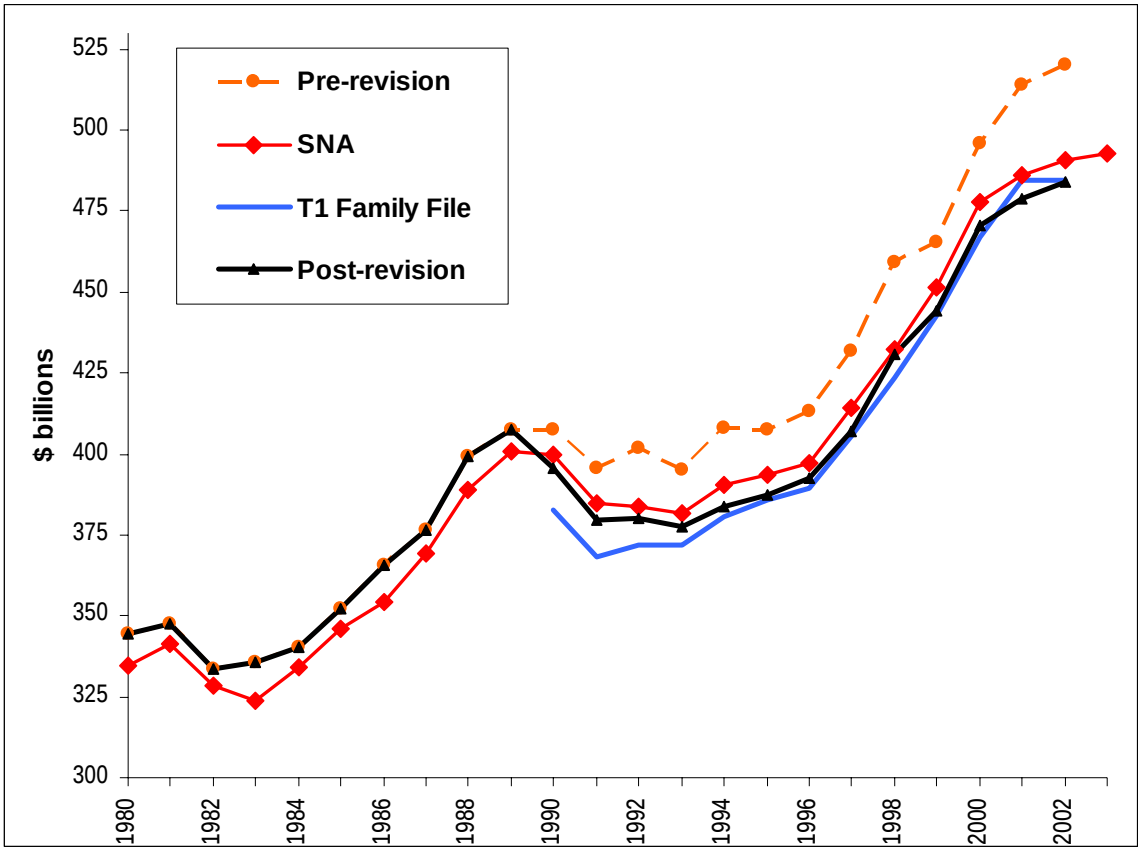
1.9 Evaluation

This section illustrates the effects of including the T4 earnings calibration in SLID and SCF. The effects of reducing the age-sex groupings from 22 to 15, as described in Section 1.6, are included. But since they are so small relative to the income calibration effects, the depictions here are also very representative of the T4 impacts alone. Note that this section focuses specifically on

the controlled characteristics only. Part 2 of this paper extends the evaluation to various standard published tables from SLID.

Chart 1.9 shows that, after the revision, the values of aggregate wages and salaries in SLID and SCF are much closer to related estimates from other sources. It shows both of these series, along with the SNA series and the T1-derived estimates as obtained from the T1 Family File (T1FF) produced by Small Area and Administrative Data Division of Statistics Canada. The T4 file series is not included simply because the SNA and T4 estimates are almost the same (Table 1.4, above).

Chart 1.9: Aggregate value of wages and salaries, Canada, 1980-2002, in constant dollars of 2000



Note that the survey estimate each year is not just closer to the SNA estimate now—it is a little lower. This is normal given that the survey information is based on a sample of the population. In the complete population, the very highest end, or “tail”, of the distribution comprises the few people who have extremely high earnings. As explained in Section 1.1, it is unlikely that the few individuals making up a rare sub-group like this will be represented by even one respondent in the random sample. The huge influence that these individuals’ wages and salaries have on the aggregate value in the population remains totally absent in the survey. (The best we can do is to define the highest earner category in the sample as narrowly as possible, to prevent their weights from being adjusted downwards any more than necessary to match the benchmark number of earners in their category.) Therefore, the administrative data should always be higher than the survey estimate if all other earnings groups are correctly represented.

Part 2
Impacts of the 2003 historical revision

The final weights in the Survey of Consumer Finances (SCF) and the Survey of Labour and Income Dynamics (SLID) were revised from 1990 to 2002 in order to the change from the 1996 Census to the 2001 Census for population benchmarks and to introduce income calibration based on the T4 file. The SCF and SLID results on low income prevalence were also affected by newly derived LICOs from 1980 to the current period. Of the three reasons why the 2003 historical revision was carried out, the new income calibration had by far the largest impact. (The separate impact of each factor for selected estimates is shown in Table A of the Appendix.)

Part 2 of this document gives examples of the impact of the revision on data results. (Section 1.9 of Part 1 also contains some charts illustrating the impacts.)

The historical revision had a relatively small impact on the trends shown by SLID since 1996 and SCF up to 1995. This is because the addition of weight adjustments based on the T4 file, referred to as income calibration, had similar impacts for all years. Over many years, however, such as the entire period from 1990 to 2002, a gradual increase in the size of the revisions may be apparent for certain topics in the SCF-SLID time series. One of these topics is low income prevalence based on low income cut-offs (LICOs), since aggregate income and the distribution of income were both affected by the revision.

The precise impacts of the 2003 historical revision can be seen by comparing the data published before and after. Note however that previously published dollar figures must be compared with their revised values using either current dollars or consistent constant dollars in both cases. Be aware also that rounding of dollar values in published tables may affect the precision of comparisons.

The weight revision was taken back only to 1990, so data users are cautioned about findings that combine data on the years up to 1989 with data on the years after 1989. The data up to 1989 can still be used with confidence, particularly regarding trends, but the potential for a level break in the series between 1989 and 1990 should be evaluated carefully when comparing both periods together.

To facilitate an understanding of the direction and size of the changes since 1990, some key examples are summarized here. In the section on “Low income prevalence”, the period from 1980 to 1989 is also briefly discussed.

These examples do not include 2003, because 2003 was not published on a pre-revision basis, but the dollar values shown and compared in these examples are in constant dollars of 2003. The year 2003 was chosen because the revised figures now available in standard tables are in constant dollars of 2003.

Note: throughout this paper, wages and salaries are before income taxes and other deductions (gross wages and salaries).

Box 2
Mechanism whereby T4 calibration affects potentially all data

Although many people in the population do not have earnings from employment, the calibration based on T4 data corrects biases not only for the wage-earning population but also for the rest of the population. The weights must still adhere to population counts, so an imbalance in one sub-group, such as an over-estimation of the earner population if that is the case, implies an opposite imbalance in the non-earner population. With the enhanced weight adjustment, the offsetting imbalances are both corrected.

The fact that wages and salaries make up a small portion of seniors' incomes means that seniors predominantly fall into a single category, namely the residual category of people with no employment earnings. (The earnings calibration categories are described in Section 1.9 of Part 1.)

The other categories (seniors with earnings) are made up of relatively few persons. The overall impact of T4 calibration on seniors' incomes is therefore small.

Since children live predominantly with non-seniors, the income results concerning children, such as low income status based on family income, are affected in a similar way to non-seniors.

Finally, the new income calibration increases the survey accuracy on characteristics seemingly unrelated to income, via the impacts on earners and non-earners.

2.1 Wages and salaries: distribution

In general, the number of people in the ten provinces with earnings from employment[6] was adjusted downwards. The adjustment increased the number of people in the lowest and highest earnings categories, and reduced the number of people in a very large middle range of the distribution.

Table 2.1 illustrates this point. Prior to the revision, the survey recorded 14.5 million individuals in 2002 as having total wages and salaries in a very large middle range, between $5,000 and $150,000 (in constant dollars of 2003). After the revision, there are 1.1 million fewer people in this

6. The terms earnings from employment, earnings, and wages and salaries are all used interchangeably in this document. Earnings from self-employment are not included.

large range category for 2002, giving a new figure of 13.3 million. In contrast, the number of individuals with wages and salaries outside that range rose by less than one-quarter million. Specifically, there are close to 200,000 more people with wages and salaries below $5,000 (an increase of 8.9%), and about 28,000 more people with wages and salaries of $150,000 or more (an increase of 27%). The latter increase, of 27%, is large but people who received $150,000 or more in wages in salaries still made up only one-half of one percent of all adults, while people who received between $1 and $5000 make up almost one-tenth of all adults.

Table 2.1: Distribution of adults by annual wages and salaries, Canada, 2002, in current dollars of 2003

					Proportion of all adults	
	pre-revision	post-revision	Change		pre-revision	post-revision
$0	7,953,400	8,672,300	718,900	9.0%	32.2%	35.4%
$1 - $4,999	2,199,300	2,396,000	196,700	8.9%	8.9%	9.8%
$5,000 - $149,999	14,453,600	13,316,000	- 1,137,600	-7.9%	58.5%	54.3%
$5,000 - $9,999	1,735,700	1,704,100	- 31,600	-1.8%	7.0%	7.0%
$10,000 - $19,999	2,748,900	2,602,800	- 146,100	-5.3%	11.1%	10.6%
$20,000 - $29,999	2,467,000	2,281,400	- 185,600	-7.5%	10.0%	9.3%
$30,000 - $39,999	2,396,900	2,109,200	- 287,700	-12.0%	9.7%	8.6%
$40,000 - $49,999	1,705,700	1,582,200	- 123,500	-7.2%	6.9%	6.5%
$50,000 - $59,999	1,248,900	1,114,400	- 134,500	-10.8%	5.1%	4.5%
$60,000 - $79,999	1,431,100	1,265,000	- 166,100	-11.6%	5.8%	5.2%
$80,000 - $149,999	719,400	656,900	- 62,500	-8.7%	2.9%	2.7%
$150,000 and over	104,800	133,100	28,300	27.0%	0.4%	0.5%
Total >$0	16,757,700	15,845,100	- 912,600	-5.4%	67.8%	64.6%
TOTAL	24,711,000	24,517,300	- 193,700	-0.8%	100.0%	100.0%

Note that the approximate turning points of $5,000 and $150,000 for the direction of the revisions are indications of the impact for all adults in all provinces combined. Different approximate turning points could be observed by province or by other characteristics. It may not even be appropriate to identify a single turning point at the higher end by province, where the survey sample becomes very small due to the small size of this population in reality. Nonetheless, this overall pattern depicts well the nature of the revision due to the new calibration based on wages and salaries using the T4 file.

In the combined ten provinces, the number of adults including seniors with no wages and salaries in 2002, about 8.0 million before the historical revision, was revised upwards by 9.0% to 8.7 million people. Non-earners represented 32.2% of adults before the revision, and 35.4% after the revision.

2.2 Wages and salaries: aggregate values

Due to the predominant downward revisions in the number of persons by earnings classes and despite the upward revision in the number of high-wage earner adults (which was large in percentage terms but small in terms of the number of people), average and median wages and salaries usually decreased with the revision. Among families of two or more people, aggregate wages and salaries for 2002 were revised downwards by 4.1% to $537.2 billion (in constant dollars of 2003). British Columbia, Prince Edward Island, Saskatchewan, and Newfoundland and Labrador had revisions that were larger than this in percentage terms (see Table 2.2).

Table 2.2: Aggregate annual wages and salaries, Economic families and unattached individuals, Canada and provinces, 2002, in 2003 constant dollars

	pre-revision	**post-revision**	**Difference**	
	$ millions	$ millions	$ millions	%
Canada	**560,317**	**537,190**	**- 23,127**	**- 4.1**
Newfoundland and Labrador	6,607	6,288	- 319	- 4.8
Prince Edward Island	1,747	1,612	- 135	- 7.7
Nova Scotia	13,815	13,537	- 278	- 2.0
New Brunswick	10,687	10,423	- 264	- 2.5
Quebec	120,687	116,670	- 4,016	- 3.3
Ontario	239,281	228,886	- 10,395	- 4.3
Manitoba	17,329	16,882	- 447	- 2.6
Saskatchewan	14,307	13,489	- 818	- 5.7
Alberta	61,953	61,820	- 133	- 0.2
British Columbia	73,904	67,582	-6,322	- 8.6

2.3 Family income, government transfers, and income tax

Since wages and salaries constitute a large part of income, most estimates of market income, total income, and after-tax income were also adjusted downwards. Market income includes wages and salaries, self-employment income, private pensions, investment income, and other, smaller sources of private income.

The average market income in 2002 of families of two or more people was revised downwards by 3.1%, from $67,700 to $65,600. Among unattached individuals, average market income in 2002 was revised downwards by 6.5%, from $26,300 to $24,600. See Table 2.3.1 (Canada) or Table B in the Appendix (Canada and provinces).

After-tax income is composed of market income and government transfers, less income taxes. The level of average after-tax income for families of two or more persons in 2002 changed from $62,200 to $60,400, a revision of -2.9%, and for unattached individuals it was revised by -4.9%.

Table 2.3.1: Average family income (market, total, after-tax income), government transfers, and income tax, Canada, 2002, in 2003 constant dollars

		Pre-revision	Post-revision	difference	% change
Economic families and unattached individuals	**Market income**	**54,400**	**52,000**	**-2,400**	**-4.4**
	Government transfers	6,900	7,000	100	1.4
	Total income	61,300	59,000	-2,300	-3.8
	Income tax	10,600	10,300	-300	-2.8
	After-tax income	50,700	48,800	-1,900	-3.7
Economic families	**Market income**	**67,700**	**65,600**	**-2,100**	**-3.1**
	Government transfers	7,500	7,800	300	4.0
	Total income	75,300	73,400	-1,900	-2.5
	Income tax	13,100	13,000	-100	-0.8
	After-tax income	62,200	60,400	-1,800	-2.9
Unattached individuals	**Market income**	**26,300**	**24,600**	**-1,700**	**-6.5**
	Government transfers	5,400	5,600	200	3.7
	Total income	31,700	30,200	-1,500	-4.7
	Income tax	5,200	4,900	-300	-5.8
	After-tax income	26,600	25,300	-1,300	-4.9

As a consequence of lower annual wages and salaries, government transfers generally increased, particularly those that are most targeted to families and individuals with little private income, such as social assistance benefits, child benefits, and refundable tax credits such as the GST/HST Credit and certain provincial sales tax credits.

In total, the aggregate value of government transfers was revised upwards by 3.5% to $90.5 billion (see Table 2.3.2.). The component of government transfers that underwent the largest correction due to the revision was social assistance benefits, thereby improving greatly the estimation of this income source in both SCF and SLID. The aggregate value of social assistance benefits in 2002 was raised by 20%.

The aggregate amount of child benefits was revised upwards by 7.4% in 2002, and the aggregate amount of federal GST/HST Credit, by 5.7%.

Table 2.3.2: Aggregate government transfers, Canada, 2002, in 2003 constant dollars

	Pre-revision $ million	**Post-revision** $ million	**Difference** $ million	**Difference** %
Government transfers	**87,437**	**90,516**	**3,079**	**3.5**
Old Age Security (pension and supplement)	23,233	23,194	- 39	-0.2
CPP and QPP benefits	26,683	27,443	760	2.8
Child benefits	7,376	7,920	544	7.4
Employment Insurance benefits	13,777	13,629	- 149	-1.1
Workers compensation benefits	4,965	5,231	266	5.3
GST and Harmonized Sales Tax (HST) credits	2,623	2,772	150	5.7
Refundable provincial and territorial tax credits	1,595	1,689	94	5.9
Social assistance	7,186	8,639	1,453	20.2

Average income tax paid by families of two or more people in Canada declined slightly due to the revision, from $13,100 to $13,000. All provinces had downward revisions except Ontario, which had no change. The largest change was -4.5%, for Saskatchewan. Refer to Table 2.3.1 above (Canada) and Table A in the Appendix (Canada and provinces).

2.4 Low income prevalence

In the period 1980 to 1989, the results from SCF and SLID on low income prevalence changed only as a result of revised 1992-based low income cut-offs (LICOs) following a historical re-weighting of the 1992 Family Expenditure Survey. [7] The effect was a small upward revision in low income prevalence: the rate for Canada, all persons, was 1% or 2% higher each year (between 0.1 and 0.2 percentage points). (See Table 2.4.1) By province, there are many small downward revisions as well as upward revisions.

In the period from 1990 to 2002, the low income results changed for three reasons: the revised 1992-based LICOs, the switch from the 1996 Census to the 2001 Census for population benchmarks used in the calculation of survey weights, and the newly introduced income calibration based on the T4 file, also used in the weight calculation starting with the 2003 revision and data release.

7. Information on low income cut-offs can be found in “Survey of Labour and Income Dynamics – A Survey Overview” and in “Low income cut-offs for 2004 and low income measures for 2002”. The new 1992-based low income cut-offs are printed in the second of these two documents.

Consistent with the upward revisions in the number of people with no earnings or very low earnings, the low income rates for persons and families generally underwent upward revisions in all years from 1990 to 2002. The revisions gradually increased in size over that long period, starting with a revision of 8% in 1990 and finishing with a 22% upwards revision in 2002. (See Table 2.4.2)

Nevertheless, many conclusions about trends in low income are still expected to be the same following the historical revision. For example, whether on a pre- or post-revision basis, the prevalence of low income for all persons generally increased between 1990 and 1996, and then declined each year until 2001. (See Chart 2.4.) From 1990 to 1996, the revised numbers show a 33% increase in low income status among persons, while the data published before the revision showed a 28% increase. From 1996 to 2001, the revised numbers show a decrease of 34%, instead of the previous 29%.

For the most recent years of the revision, the low income rate for all persons was revised as follows: in 2000, from 10.9% to 12.5%; in 2001, from 9.3% to 11.2%; and in 2002, from 9.5% to 11.6%. These are changes of 15%, 20%, and 22%, respectively. (See Table 2.4.2)

The low income rates for *families* of two or more people changed as follows: in 2000, from 7.9% to 9.0%; in 2001, from 6.6% to 7.9% in 2001; and in 2002, from 7.0% to 8.6%. These are percentage point increases of 14%, 20%, and 23% respectively. For unattached individuals, the corresponding increases in low income prevalence were 15%, 18%, and 19%.

In 2002, the low income rate for children under age 18 changed from 10.2% to 12.2%, and for adults aged 18 to 64 it rose from 9.7% to 12.1%. The rate for seniors—already smaller than that of the two other large age groups—rose considerably less, from 6.9% to 7.6%. (For an explanation of the smaller impacts regarding seniors, see the Box, “Mechanism whereby T4 calibration affects potentially all data”, at the beginning of Part 2.)

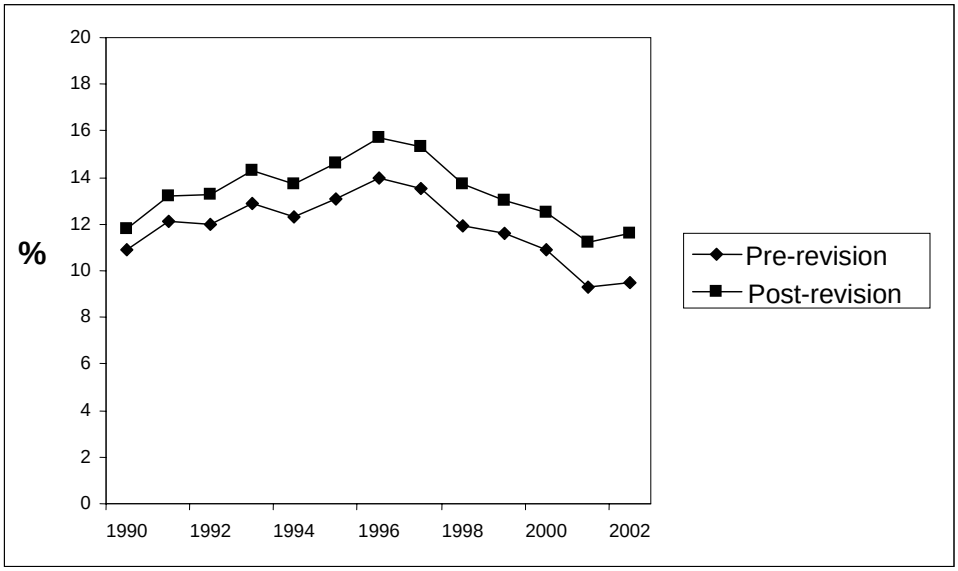
Table 2.4.1: Revisions in low income rates of persons, Canada, 1980 to 1989

	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989
Percentage point change: post-revision minus pre-revision										
All persons	**0.1**	**0.1**	**0.1**	**0.1**	**0.1**	**0.2**	**0.2**	**0.1**	**0.1**	**0.2**
Under 18 years	0.0	0.1	0.1	0.1	0.2	0.1	0.3	0.0	0.1	0.2
18 to 64 years	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
65 years and over	0.5	0.3	0.2	0.4	0.3	0.5	0.6	0.3	0.5	0.4
Percentage change: (post-revision minus pre-revision) / pre-revision x 100										
All persons	**1%**	**1%**	**1%**	**1%**	**1%**	**2%**	**2%**	**1%**	**1%**	**2%**
Under 18 years	0%	1%	1%	1%	1%	1%	2%	0%	1%	2%
18 to 64 years	1%	1%	1%	1%	1%	1%	1%	1%	1%	1%
65 years and over	2%	1%	1%	2%	2%	4%	5%	2%	4%	4%

Table 2.4.2: Revisions in low income rates of persons, Canada, 1990 to 2003

	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002
Percentage point change: post-revision minus pre-revision													
All persons	**0.9**	**1.1**	**1.3**	**1.4**	**1.4**	**1.5**	**1.7**	**1.8**	**1.8**	**1.4**	**1.6**	**1.9**	**2.1**
Under 18 years	0.8	0.8	0.9	1.0	1.1	1.3	1.9	1.8	1.9	1.0	1.4	1.7	2.0
18 to 64 years	1.0	1.3	1.5	1.5	1.7	1.7	1.8	2.0	2.0	1.6	1.9	2.2	2.4
65 years and over	0.6	0.7	0.6	0.7	0.7	0.9	0.7	0.5	0.3	0.3	0.4	0.4	0.7
Percentage change: (post-revision minus pre-revision) / pre-revision x 100													
All persons	**8%**	**9%**	**11%**	**11%**	**11%**	**11%**	**12%**	**13%**	**15%**	**12%**	**15%**	**20%**	**22%**
Under 18 years	6%	6%	6%	6%	7%	8%	11%	11%	14%	7%	11%	16%	20%
18 to 64 years	10%	11%	13%	12%	14%	13%	13%	15%	17%	14%	17%	23%	25%
65 years and over	6%	7%	7%	6%	9%	12%	8%	6%	4%	4%	6%	6%	10%

Chart 2.4: Low income rate, All persons, Canada, 1990 to 2002



2.5 Income inequality

Consistent with the increases in the number of people at the very low and very high ends of the earnings distribution, and the decrease in the number of people in the middle of the distribution, certain measures of income inequality (Gini coefficient, ratio of 5th to 1st quintile) generally increased.

The ratio of average market income of the one-fifth of families with the highest values of market income to the one-fifth of families with the lowest values (5^{th} quintile versus 1^{st} quintile) underwent an upwards revision in each year from 1990 to 2002. The trend, however, is still much the same as before. There is just a slight widening of the difference between the pre- and post-revision ratio over the years, with the largest difference occurring in 2002.

The change in the ratio of the 5^{th} quintile to the 1^{st} quintile is less noticeable when calculated for after-tax income than it is for market income. Nevertheless, the same observations can be made in that the after-tax income ratio is also higher in all years from 1990 to 2002 after the historical revision, and that a widening of the difference is again apparent—and strongest between 2001 and 2002.

The Gini coefficient for economic families of two or more persons is higher in all years from 1990 to 2002 after the historical revision, whether it is calculated on market income or after-tax income. The change in the Gini coefficient is larger for market income, because market income was

affected more by the revision than was after-tax income. Over the long period covered by the revision, the larger impact of the revision on later years than earlier years is apparent.

2.6 Topics other than income

As mentioned earlier, even though the re-weighting aimed at improving the quality of the income estimates, this revision also improved statistics that tend to be somewhat associated with income.

The rate of home ownership shown in the SLID data changed slightly, from 68% to 67% of households in December 2002. (The corresponding percentage of households who rent their dwelling changed from 32% to 33%.) The 2001 Census, conducted in May 2001, recorded an ownership rate of 66%. Data prior to 2002 are available in SLID but are not imputed for non-response, making any comparisons more tenuous for those years.[8]

The distribution of educational attainment in Canada, based on broad categories of "high school or less" and "some post-secondary certificate", was already very close to the Census distribution, but the revision brought it even closer. (See Table 2.6.)

Table 2.6: Educational attainment, Persons aged 25-64, Canada, 2001

	pre-revision %	**post-revision** %	**Census** %
High school or less	46.1	46.7	46.6
Some post-secondary certificate	53.9	53.3	53.4

Conclusion

From the mid-1990s until 2005, Statistics Canada carried out a great deal of research and evaluation on the methods used to produce final weights for the Survey of Labour and Income Dynamics (SLID) and some other household surveys. Called the Harmonization of Income Statistics Project (HARC), this effort led to substantial improvements being made in SLID's demographic calibration using the Census of Population, plus the introduction of a major new type of calibration that uses tax data from CRA regarding people's employment earnings. This paper has covered the various factors that occurred simultaneously to motivate such an in-depth review and make such improvements possible.

8. The Survey of Labour and Income Dynamics receives sponsorship from Canada Mortgage and Housing Corporation for content on housing characteristics, starting with reference year 2002.

The historical weight revision of SCF and SLID data which incorporated income calibration based on the T4 file had a greater impact than typically observed in historical weight revisions. Such a major change demanded that it be undertaken very carefully. To this end, the HARC Project included a thorough assessment of the T4 and T1 files for the intended purpose, testing of several options for income calibration over many years of data, consultation with external experts, and analysis of the impacts on a variety of survey results to verify the integrity of the data following the new procedure.

The one-time historical revision to introduce income calibration in SLID changed the level of estimates from the survey but had relatively little effect on the trends revealed in the data. Although users of the SLID data must deal with the data revision, it was important to carry out the revision back in time to avoid distorting the trends that are reflected in the data. It is also expected that the new estimates are more in line with what users may already know to be true based on their experience with other sources of information.

The HARC project is now virtually complete, but Statistics Canada is continuing to look at ways to improve the Survey of Labour and Income Dynamics in a variety of ways. Weight adjustment methods will continue to be monitored and reviewed. Just as importantly, changes are being made regarding survey collection procedures, and it appears they are making a positive impact. They should at least reduce the size of the weight adjustments necessary to bring the data in line with what we know from the T4 file. Meanwhile, it is expected that this new method of benchmarking will continue to be used for the foreseeable future.

References

Cotton, C., M. Webber, M. Meere, K. Bishop, and P. Hewer, “A Comparison of the Results of the Survey of Labour and Income Dynamics (SLID) and the Survey of Consumer Finances (SCF), 1993-1996.” *Income Research Paper Series*, Statistics Canada, Catalogue number 75F0002MIE-99002. April 1999.

Cathy, C., K. Bishop, P. Giles, P. Hewer, and Y. Saint-Pierre. “A Comparison of the Results of the Survey of Labour and Income Dynamics (SLID) and the Survey of Consumer Finances (SCF), 1993-1997: Update.” *Income research paper series*, Statistics Canada, Catalogue number 75F0002MIE-99007. December 1999.

LaRoche, S. and C. Cauchon. “Longitudinal and Cross-Sectional Weighting of the Survey of Labour and Income Dynamics, 2003 Reference year”. *Income Research Paper Series*, Statistics Canada (forthcoming).

Latouche, M., and S. LaRoche. “The Survey of Labour and Income Dynamics Calibration Strategy.” Technical report presented to the Advisory Committee on Statistical Methods. Statistics Canada, November 2003 (internal document).

Lavallée, Pierre. “Cross-sectional Weighting of the Longitudinal Surveys of Individuals and Households Using the Weight Share Method.” *Survey Methodology*, vol. 21, no 1 (June 1995), 25-32.

Lévesque, I., and S. Franklin. “Longitudinal and Cross-Sectional Weighting of the Survey of Labour and Income Dynamics, 1997 Reference year.” *Income Research Paper Series*, Statistics Canada, Catalogue No. 75F0002MIE-00004. June 2000.

Rea, Willa, and Edith Greenlee. “The Effects of the Revised Estimation Methodology on Estimates from Household Expenditure Surveys. *Income Research Paper Series*, Statistics Canada, Catalogue No. 75F0002MIE, (forthcoming).

Statistics Canada. *Income in Canada*. Catalogue number 75-202-X, May 2005 (free at www.statcan.ca).

Statistics Canada. *Income Trends in Canada*. Catalogue number 13F0022X, May 2005.

Statistics Canada. *Survey of Labour and Income Dynamics – A Survey Overview*. Catalogue number 750011X, May 2005 (free at www.statcan.ca).

Tremblay, J., C. Nadeau, S. Auger, S. LaRoche, and M. Latouche. “Developments on the Harmonised Calibration of Income Statistics Project.” Background document distributed to the Advisory Committee on Statistical Methods. Statistics Canada, November 2003 (internal document).

Webber, M., M. Latouche, and E. Rancourt. “Harmonised Calibration of Income Statistics.” Technical report presented to the Advisory Committee on Statistical Methods. Statistics Canada, April 2000 (internal document).

Appendix

Table A: Published and unpublished data showing separate impacts of the new demographic calibration (new Census) and the introduction of income calibration (T4), Canada, 2002 and 2003

	2002			2003		
	pre-revision Excludes new Census Excludes T4	**unpublished** Includes new Census Excludes T4	**post-revision** Includes new Census Includes T4	**pre-revision** Excludes new Census Excludes T4	**unpublished** Includes new Census Excludes T4	**post-revision** Includes new Census Includes T4
Average after-tax income	(constant dollars of 2003)			(constant dollars of 2003)		
All	50,700	50,700	48,800	na	50,600	48,400
Families (2 persons or more)	62,200	62,000	60,400	na	61,700	59,900
Unattached individuals	26,600	26,600	25,300	na	27,000	25,600
Persons in low income[1]	('000)			('000)		
All persons	2,908	2,954	3,536	na	2,997	3,552
Under 18 years	702	721	839	na	736	843
18 to 64	1,946	1,959	2,413	na	2,013	2,450
65 and over	260	273	284	na	248	258
	(%)			(%)		
All persons	9.5	9.6	11.6	na	9.7	11.5
Under 18 years	10.2	10.5	12.2	na	10.8	12.4
18 to 64	9.7	9.7	12.1	na	9.9	12.1
65 and over	6.9	7.3	7.6	na	6.5	6.8
Families in low income*	('000)			('000)		
All	1,621	1,651	1,999	na	1,661	2,014
Families (2 persons or more)	605	609	736	na	616	726
Unattached individuals	1015	1,042	1,263	na	1,045	1,287
	(%)			(%)		
All	12.7	12.9	15.5	na	12.9	15.4
Families (2 persons or more)	7.0	7.0	8.6	na	7.0	8.4
Unattached individuals	24.8	25.4	29.5	na	25.3	29.4

1. Family income under the low income cut-off (LICO, 1992 base, after tax)

Notes:

"New Census" refers to calibration using the 2001 Census-based population estimates and the new age-sex categories.

"T4" refers to calibration using the T4 file estimates of earners by earnings classes.

Table B: Average family income (market, total, after-tax income), government transfers, and income tax, Canada and provinces, 2002, in 2003 constant dollars

Canada		pre-revision	post revision	difference	% change
Economic families and unattached individuals	Market income	54,400	52,000	-2,400	-4.4
	Government transfers	6,900	7,000	100	1.4
	Total income	61,300	59,000	-2,300	-3.8
	Income tax	10,600	10,300	-300	-2.8
	After-tax income	50,700	48,800	-1,900	-3.7
Economic families	Market income	67,700	65,600	-2,100	-3.1
	Government transfers	7,500	7,800	300	4.0
	Total income	75,300	73,400	-1,900	-2.5
	Income tax	13,100	13,000	-100	-0.8
	After-tax income	62,200	60,400	-1,800	-2.9
Unattached individuals	Market income	26,300	24,600	-1,700	-6.5
	Government transfers	5,400	5,600	200	3.7
	Total income	31,700	30,200	-1,500	-4.7
	Income tax	5,200	4,900	-300	-5.8
	After-tax income	26,600	25,300	-1,300	-4.9

Newfoundland and Labrador		pre-revision	post revision	difference	% change
Economic families and unattached individuals	Market income	39,400	36,600	-2,800	-7.1
	Government transfers	10,700	10,700	0	0.0
	Total income	50,100	47,200	-2,900	-5.8
	Income tax	7,700	7,200	-500	-6.5
	After-tax income	42,400	40,000	-2,400	-5.7
Economic families	Market income	45,300	43,600	-1,700	-3.8
	Government transfers	11,600	11,800	200	1.7
	Total income	57,000	55,400	-1,600	-2.8
	Income tax	8,900	8,600	-300	-3.4
	After-tax income	48,100	46,800	-1,300	-2.7
Unattached individuals	Market income	15,000	13,900	-1,100	-7.3
	Government transfers	7,000	6,800	-200	-2.9
	Total income	21,900	20,700	-1,200	-5.5
	Income tax	3,000	2,700	-300	-10.0
	After-tax income	19,000	18,000	-1,000	-5.3

Prince Edward Island		pre-revision	post revision	difference	% change
Economic families and unattached individuals	Market income	39,700	38,200	-1,500	-3.8
	Government transfers	9,100	9,500	400	4.4
	Total income	48,700	47,700	-1,000	-2.1
	Income tax	6,700	6,500	-200	-3.0
	After-tax income	42,100	41,200	-900	-2.1
Economic families	Market income	49,100	47,700	-1,400	-2.9
	Government transfers	10,100	10,500	400	4.0
	Total income	59,200	58,200	-1,000	-1.7
	Income tax	8,400	8,300	-100	-1.2
	After-tax income	50,800	49,900	-900	-1.8
Unattached individuals	Market income	16,000	14,100	-1,900	-11.9
	Government transfers	6,600	7,000	400	6.1
	Total income	22,600	21,100	-1,500	-6.6
	Income tax	2,400	2,100	-300	-12.5
	After-tax income	20,200	19,000	-1,200	-5.9

Nova Scotia		pre-revision	post revision	difference	% change
Economic families and unattached individuals	Market income	45,100	43,100	-2,000	-4.4
	Government transfers	7,700	7,800	100	1.3
	Total income	52,800	50,900	-1,900	-3.6
	Income tax	9,100	8,800	-300	-3.3
	After-tax income	43,700	42,100	-1,600	-3.7
Economic families	Market income	55,900	54,800	-1,100	-2.0
	Government transfers	8,400	8,500	100	1.2
	Total income	64,300	63,400	-900	-1.4
	Income tax	11,400	11,300	-100	-0.9
	After-tax income	52,900	52,100	-800	-1.5
Unattached individuals	Market income	19,200	18,300	-900	-4.7
	Government transfers	6,100	6,200	100	1.6
	Total income	25,200	24,500	-700	-2.8
	Income tax	3,700	3,600	-100	-2.7
	After-tax income	21,500	21,000	-500	-2.3

New Brunswick		pre-revision	post revision	difference	% change
Economic families and unattached individuals	Market income	41,600	40,000	-1,600	-3.8
	Government transfers	8,400	8,500	100	1.2
	Total income	50,100	48,500	-1,600	-3.2
	Income tax	7,800	7,600	-200	-2.6
	After-tax income	42,200	40,900	-1,300	-3.1
Economic families	Market income	52,000	50,100	-1,900	-3.7
	Government transfers	9,100	9,300	200	2.2
	Total income	61,100	59,300	-1,800	-2.9
	Income tax	9,800	9,600	-200	-2.0
	After-tax income	51,300	49,800	-1,500	-2.9
Unattached individuals	Market income	15,500	15,300	-200	-1.3
	Government transfers	6,700	6,800	100	1.5
	Total income	22,200	22,100	-100	-0.5
	Income tax	2,800	2,800	0	0.0
	After-tax income	19,400	19,300	-100	-0.5

Quebec		pre-revision	post revision	difference	% change
Economic families and unattached individuals	Market income	47,200	45,000	-2,200	-4.7
	Government transfers	7,200	7,300	100	1.4
	Total income	54,400	52,300	-2,100	-3.9
	Income tax	10,200	9,800	-400	-3.9
	After-tax income	44,200	42,500	-1,700	-3.8
Economic families	Market income	59,800	57,900	-1,900	-3.2
	Government transfers	8,000	8,100	100	1.3
	Total income	67,800	66,000	-1,800	-2.7
	Income tax	12,900	12,500	-400	-3.1
	After-tax income	54,900	53,500	-1,400	-2.6
Unattached individuals	Market income	23,600	22,300	-1,300	-5.5
	Government transfers	5,800	5,900	100	1.7
	Total income	29,400	28,200	-1,200	-4.1
	Income tax	5,300	5,000	-300	-5.7
	After-tax income	24,100	23,200	-900	-3.7

Ontario		pre-revision	post revision	difference	% change
Economic families and unattached individuals	Market income	62,800	60,200	-2,600	-4.1
	Government transfers	6,700	6,900	200	3.0
	Total income	69,500	67,200	-2,300	-3.3
	Income tax	12,100	12,000	-100	-0.8
	After-tax income	57,400	55,200	-2,200	-3.8
Economic families	Market income	76,400	74,300	-2,100	-2.7
	Government transfers	7,300	7,500	200	2.7
	Total income	83,700	81,800	-1,900	-2.3
	Income tax	14,700	14,700	0	0.0
	After-tax income	69,000	67,100	-1900	-2.8
Unattached individuals	Market income	30,500	28,300	-2,200	-7.2
	Government transfers	5,300	5,500	200	3.8
	Total income	35,800	33,800	-2,000	-5.6
	Income tax	5,900	5,700	-200	-3.4
	After-tax income	29,800	28,200	-1,600	-5.4

Manitoba		pre-revision	Post revision	difference	% change
Economic families and unattached individuals	Market income	47,500	45,600	-1,900	-4.0
	Government transfers	6,400	6,600	200	3.1
	Total income	54,000	52,200	-1,800	-3.3
	Income tax	8,100	7,800	-300	-3.7
	After-tax income	45,900	44,400	-1,500	-3.3
Economic families	Market income	60,100	58,400	-1,700	-2.8
	Government transfers	7,000	7,200	200	2.9
	Total income	67,100	65,600	-1,500	-2.2
	Income tax	10,200	9,900	-300	-2.9
	After-tax income	56,900	55,600	-1,300	-2.3
Unattached individuals	Market income	22,500	21,000	-1,500	-6.7
	Government transfers	5,300	5,500	200	3.8
	Total income	27,800	26,500	-1,300	-4.7
	Income tax	3,900	3,600	-300	-7.7
	After-tax income	23,900	22,900	-1,000	-4.2

Saskatchewan		pre-revision	post revision	difference	% change
Economic families and unattached individuals	Market income	45,900	43,800	-2,100	-4.6
	Government transfers	7,000	7,200	200	2.9
	Total income	52,900	50,900	-2,000	-3.8
	Income tax	8,800	8,300	-500	-5.7
	After-tax income	44,100	42,600	-1,500	-3.4
Economic families	Market income	58,000	55,900	-2,100	-3.6
	Government transfers	7,600	7,800	200	2.6
	Total income	65,600	63,700	-1,900	-2.9
	Income tax	11,100	10,600	-500	-4.5
	After-tax income	54,500	53,000	-1,500	-2.8
Unattached individuals	Market income	21,700	20,200	-1,500	-6.9
	Government transfers	5,800	5,900	100	1.7
	Total income	27,500	26,100	-1,400	-5.1
	Income tax	4,200	3,900	-300	-7.1
	After-tax income	23,200	22,200	-1,000	-4.3

Alberta		pre-revision	post revision	difference	% change
Economic families and unattached individuals	Market income	58,900	56,600	-2,300	-3.9
	Government transfers	5,700	5,700	0	0.0
	Total income	64,600	62,300	-2,300	-3.6
	Income tax	10,700	10,400	-300	-2.8
	After-tax income	53,900	52,000	-1,900	-3.5
Economic families	Market income	72,900	71,500	-1,400	-1.9
	Government transfers	6,300	6,400	100	1.6
	Total income	79,200	77,900	-1,300	-1.6
	Income tax	13,100	13,000	-100	-0.8
	After-tax income	66,100	64,900	-1,200	-1.8
Unattached individuals	Market income	28,700	27,500	-1,200	-4.2
	Government transfers	4,400	4,400	0	0.0
	Total income	33,100	31,900	-1,200	-3.6
	Income tax	5,300	5,200	-100	-1.9
	After-tax income	27,700	26,700	-1,000	-3.6

British Columbia		pre-revision	post revision	difference	% change
Economic families and unattached individuals	Market income	52,100	49,000	-3,100	-6.0
	Government transfers	6,600	6,900	300	4.5
	Total income	58,700	55,900	-2,800	-4.8
	Income tax	9,200	8,900	-300	-3.3
	After-tax income	49,400	47,000	-2,400	-4.9
Economic families	Market income	67,100	63,500	-3,600	-5.4
	Government transfers	7,400	7,700	300	4.1
	Total income	74,400	71,200	-3,200	-4.3
	Income tax	12,000	11,700	-300	-2.5
	After-tax income	62,500	59,500	-3,000	-4.8
Unattached individuals	Market income	26,000	23,600	-2,400	-9.2
	Government transfers	5,300	5,600	300	5.7
	Total income	31,300	29,300	-2,000	-6.4
	Income tax	4,500	4,100	-400	-8.9
	After-tax income	26,800	25,200	-1,600	-6.0