INTRODUCTION OF A NEW CALIBRATION STRATEGY FOR THE SURVEY OF HOUSEHOLD SPENDING

Sophie Arsenault, José Gaudet, Christian Nadeau and Johanne Tremblay, Statistics Canada
Christian Nadeau, R.H. Coats Building, 16th Floor, Ottawa, Ontario, K1A 0T6, Canada

KEYWORDS: Harmonisation, Calibration, Income distribution, Expenditure survey.

INTRODUCTION

In January 1998, the annual Survey of Household Spending (SHS) was introduced to replace the former occasional Family Expenditure Survey (FAMEX) in order to produce more frequent and more precise provincial and territorial expenditure estimates. The reduction in the level of detail for the data that are collected, the extension of the target population, the increase in sample size and the improvement of the sample allocation methodology are among the main changes that were made to the methodology. For the first two survey years of the SHS, it was decided to use a weighting methodology that is quite similar to the one used for the last survey year of the FAMEX. This former survey, done at the national level every four or five years, did not provide much recent information to identify which modification of the calibration strategy would be suitable. Studies on coverage and representativeness of the sample done on the first survey year of SHS suggested that the use of more detailed demographic controls and the addition of income related controls at the calibration stage would be suitable.

At the same time the SHS was introduced, the Task Force on Income Statistics examined ways of producing income estimates at lower cost, of harmonizing them and of improving their range and quality (see Coombs et al., 1998). The Task Force also developed a conceptual framework for statistic on income distribution, covering also assets, debts and expenditures. One of the Task Force report's recommendations was to harmonize the processing of surveys on personal and household finance in order to reduce the differences in the estimates from different sources. Another recommendation was to increase the use of administrative data, such as tax data from the Canadian Customs and Revenue Agency (CCRA) at the weighting stage. It is thought that such a procedure would improve both the quality of statistics on personal and household finance and their level of harmonisation.

In January 2000, following the recommendations of the Task Force, the project on the Harmonized Calibration was initiated with the objective of identifying and evaluating a calibration strategy that would be similar for the different surveys on personal and household finance (Webber, Latouche and Rancourt, 2000). The efforts that have been made to improve the calibration scheme of the SHS were pursued within this project. In this paper, we put the emphasis on the new calibration strategy that was implemented in the SHS as a result of the work accomplished within the Harmonized Calibration project. A short description of the methodology of the SHS is provided in section 1, before the justification of the need for a new calibration scheme is presented in section 2. Reasons for choosing specific controls for the Harmonized calibration project are presented in section 3. The methodology of the study that led to the adoption of the new calibration scheme is presented in section 4. Evaluation results are provided in sections 5 and 6, and a few implementation details are then provided in section 7. Conclusion on the implemented methodology and potential area for improvement and development are presented in section 8.

1. DESCRIPTION OF THE SURVEY OF HOUSEHOLD SPENDING METHODOLOGY

The SHS is an annual survey that collects detailed information on household expenditure habits, dwelling characteristics and household facilities and equipment. Income is also collected as a variable for analysis and a measure of data quality. The data are collected from a sample of about 20,000 to 24,000 households from the 10 provinces and 3 territories through face-to-face paper interviews. The interviews occur from January to March and cover the expenditures during the previous calendar year of every household member. A stratified multistage design is used to select the sample from the Labour Force Survey area frame. After the questionnaires have been verified, captured, edited and imputed, the weighting is performed.

The weighting includes a number of adjustment steps. To compensate for nonresponse, the respondents and nonrespondents are grouped in a number of nonresponse adjustment groups within each province. The grouping is mainly based on the design and on the level of urbanisation. Within each nonresponse adjustment group, the initial weights of the respondents are adjusted by a factor that is equal to the inverse of the weighted response rate. In the first three years of SHS, the overall response rates at the national level has varied from 73% to 76%.

The following step consists in a weight calibration on a number of controls so that some survey estimates match known population counts. Before the implementation of the new calibration strategy, three age groups and two household sizes (1, 2+) for each province and some major cities were used. Controls of the number of persons by age group were obtained from postcensal population estimates. Controls of the number of households by size were derived by adjusting the 1991 Census figures with a growth factor derived from the monthly LFS. The calibration is done using the integrated weighting approach presented in Lemaître and Dufour (1988), which ensure equal weights for every member of a particular household when controls on person characteristics are used.

Influential observations, which are defined as observations that contribute in a too large proportion to some provincial estimate, are identified and their weights are adjusted in order to reduce the estimates variability. Total income is used at the identification stage. If necessary, their weights are reduced in order to ensure that the estimated number of individuals in some high income classes does not exceed this observed in CCRA tax file. The calibration step is then reapplied to ensure that the control totals are matched. For more details on the methodology of the SHS, see Arsenault and Tremblay (2001).

2. REPRESENTATIVENESS OF THE SHS SAMPLE

A number of considerations could lead to the use of a larger set of controls at the calibration stage: harmonisation of related estimates from different surveys or with external sources, gain in estimate precision (likely to happen when the control are highly correlated with the variable of interest) and improvement of sample representativeness. This last consideration is the one that initiated the work for revising the SHS calibration scheme. It is considered important for the SHS to obtain good sample representativeness in terms of age groups, household composition, income level and geography since these characteristics are important factors in the expenditure level and pattern.

Data from the SHS 1997[1] showed that the representativeness varies substantially from one age group to the other, even after the former calibration with the three age group controls, as shown in Arsenault et al. (1999). The slippage rates, defined as the percentage differences between the survey estimates calculated from non calibrated weights (adjusted to compensate for nonresponse but not adjusted with external data) and postcensal population estimates indicate that the adult population is underestimated by 9.6% in SHS and children are slightly overestimated. The percentage differences between the survey estimates (adjusted with three age groups) and postcensal population counts defined as the residual slippage rates in Table 2.1 shows that the under-representation of persons aged between 25 and 39 remains sizeable, even after adjustment of weights. On the other hand, the over-representation of the persons aged between 15 and 19 increases substantially.

Table 2.1: Slippage Rates (%) and residual Slippage Rates (%) for detailed Age Groups, at the National Level – SHS 1997

Age Groups	Slippage Rates	Residual Slippage Rates
0-14	0.7	0.0
15-64	-9.6	0.0
15-19	0.6	17.3
20-24	-10.4	2.1
25-29	-14.9	-8.1
30-39	-10.7	-6.8
40-54	-9.6	1.3
55-64	-9.8	1.8
65+	-9.6	0.0
All	**-7.5**	**0.0**

Table 2.2: Slippage Rates (%) of the number of household by Household Sizes – SHS 1997

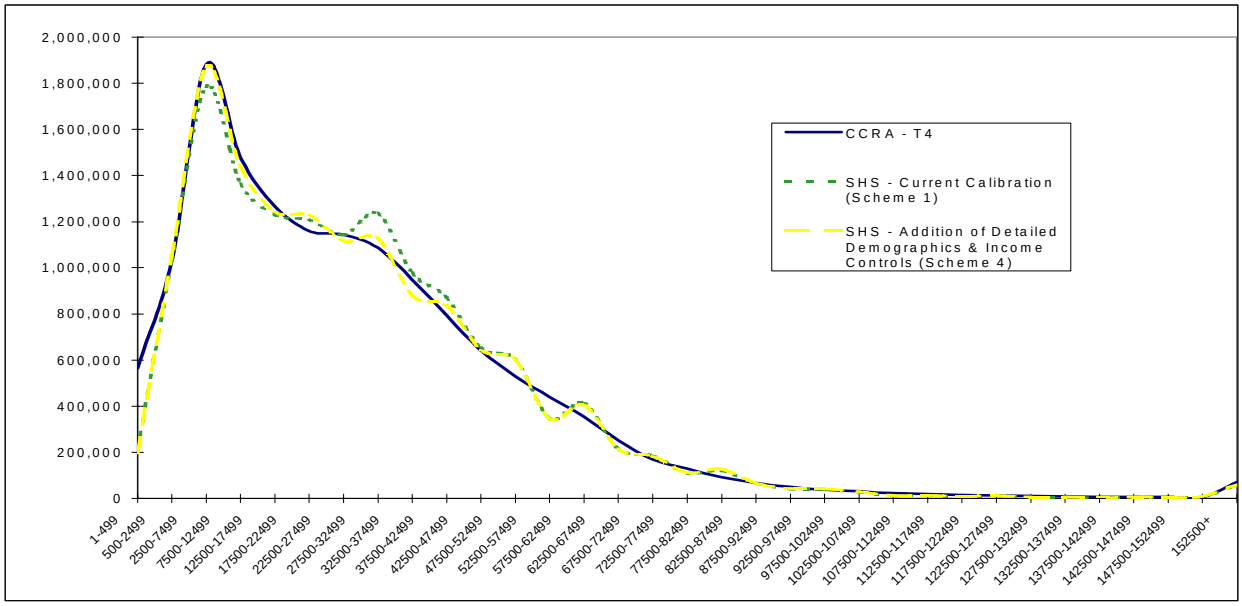
Household Size	**Slippage Rates**
1 person	-8.1
2 persons	-3.2
3 persons and more	-5.6
All	-5.5

Differences are also observed in the representativeness by household size as shown by the slippage rates presented in Table 2.2. Survey estimates of number of households calculated from non calibrated weights are compared with the new postcensal series of estimated number of households and indicate that the number of households is underestimated by 5.5%. The slippage rate differs substantially from one household size to the other, one person households being the most underestimated ones.

When compared to the distribution obtained from the statement of remuneration paid file of the CCRA (T4 file), the income distribution of the 1997 SHS also shows representativeness problems. Low and high wage and salary incomes are under represented while the

[1] The survey year represents the reference period. Data for the SHS 1997 has been collected from January to March 1998.

Figure 2.1: 1997 Population Distribution by Wage and Salary Level



middle wage and salary income classes are over represented, as observed when comparing the SHS (current calibration) and T4 curves in Figure 2.1.

The above results combined with more detailed analysis of the SHS data suggest an increase in the number of age groups and household size groups at the calibration stage, as well as the inclusion of controls on sex groups and income categories.

3. DETERMINATION OF HARMONISED AGE, SEX, HOUSEHOLD AND INCOME GROUPS

A common set of demographic variables and grouping was found to harmonize the calibration strategy of surveys on personal and household finance. Based on previous analysis made on the cross-sectional income survey data and considering the influence age and sex have on both the expenditure level and pattern, 18 age-sex groups were defined.

SHS was the only survey to use household counts (size 1 and 2+) in its calibration. Following the recent development of an official monthly series of postcensal estimated number of households and families by size (1, 2, 3+), it was decided to use these controls in all surveys on personal and household finance in order to improve representativeness and to further harmonize them. For the SHS, the adoption of these new controls was thought to improve data quality and representativeness of the sample.

Three administrative sources of income data have been identified as potential sources for calibration of surveys on personal and household finance. The T4 file, which is made of the statement of remuneration paid by each employer to all employees, has a very good coverage of the salaried work force since by law each employer must deliver this statement. However, wage and salary represent 65% of the total income. The income tax and benefit return file (T1) is made of the annual forms filled by the citizens. This file is edited and it covers most of the various sources of income included in the surveys. However, the coverage of this file is not as good as the T4 since not all individuals have to fill the T1 form. Statistics Canada also produces the T1 Family File where census families are re-constructed using addresses and Social Insurance Number. This file has the advantage of an improved coverage of the adult population since family members with lower income are synthesised from existing family member records. It also has the advantage of providing a family income instead of an individual income.

In a first step of introducing income components in the calibration of these household surveys, Wages and Salaries from the T4 file was retained, mainly for the good coverage and the similarity in concepts from the different surveys. The approach retained was to control on the number of individuals within income intervals. The underestimation or overestimation of the income distribution has been studied for small intervals of income in order to identify income boundaries for calibration. However, the selected approach for the definition of the calibration interval boundaries was to use percentiles. This approach was chosen for its simplicity, its objectivity and its contribution in harmonisation between surveys. Studies have indicated

that the difference between survey specific calibration groups and percentiles was not very large (Webber, Latouche and Rancourt, 2000). The number of groups was restricted to six by province, as defined by percentiles 25, 50, 65, 75 and 95.

4. METHODOLOGY FOR IDENTIFYING A SUITABLE CALIBRATION SCHEME

A study was conducted to decide if it was desirable to include the harmonized common set of demographic characteristics and controls on wage and salary income groups. Another objective of the study was to better understand the effect of adding each set of controls to the calibration.

Table 4.1: Description of the Calibration Schemes

Control Groups	Calibration Schemes			
	1	2	3	4
10 Provinces x 3 Age groups	✓	✓		
10 Provinces x 9 Age groups x 2 Sex			✓	✓
15 large cities x 2 Age groups	✓	✓	✓	✓
10 Provinces x 3 Household sizes	✓	✓	✓	✓
10 Provinces x 6 Wage and Salary		✓		✓

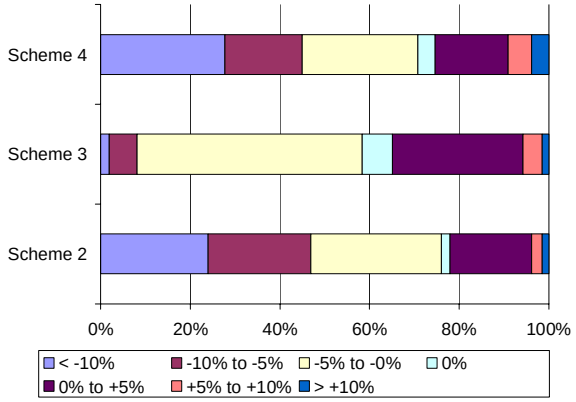
The 1997 SHS data were used for the study. These data were weighted using four different calibration schemes present in Table 4.1. The first calibration scheme is quite similar to the original calibration scheme of the SHS with the exception that it uses 3 household size categories as compared to 2. It was used as a benchmark to measure the impact of adding controls. The second calibration scheme includes controls on wage and salary income groups while the third calibration scheme includes controls on sex and on a larger number of age groups. Finally, the fourth calibration scheme combines the inclusions of schemes 3 and 4 and is the one we wanted to implement.

The impact of using different calibration schemes was evaluated by comparing the expenditure and income estimates along with their CVs, the household composition distribution and the g-factor distributions. The estimates used consisted in level estimates at province and Canada level for 19 major categories of expenditures and 13 income sources. Some results of this evaluation are presented in the next two sections.

5. RESULTS OF THE EVALUATION OF THE DIFFERENT CALIBRATION SCHEME

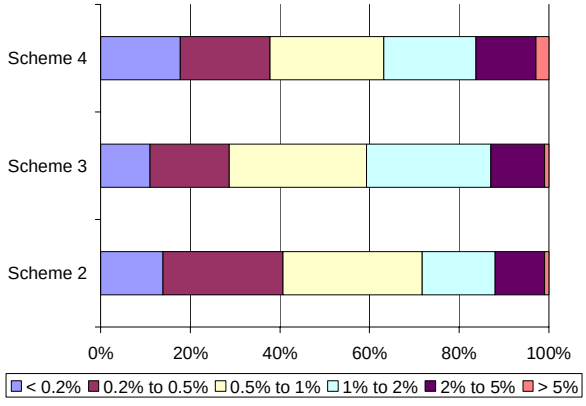
Detailed analysis by components indicated that adding controls to calibration scheme had no major negative impact on expenditure estimates and resulted in a positive effect on CVs. Figure 5.1 and 5.2 illustrate some of the results.

Figure 5.1: Impact of Adding Controls on CVs



In Figure 5.1, we observe that the CVs are reduced by more than 10% for 25% of the expenditure estimates when wage and salary controls are introduced (scheme 2). This proportion is around 5% when only age-sex controls are added (scheme 3) and around 27% when both the age-sex and wage and salary controls are included (scheme 4). It seems that the addition of wage and salary has the most positive effect on CVs. 76% of the estimates studied showed better CVs while 58% of the estimates showed a CV reduction when the more detailed demographic controls are included. When both types of controls are added, 70% of the estimates showed a CV reduction. The graph shows that generally, the addition of both types of controls has a positive impact on CVs since CV reduction is more frequently observed than the CV increase. The impact of adding wage and salary income control is greater than the impact of increasing the number of demographic controls. Moreover, the estimate variability reduction is not very big when going from calibration scheme 2 to 4.

Figure 5.2: Impact of Adding Controls on the Estimates



In Figure 5.2, we observe that, for 14% of the expenditure estimates, the absolute relative difference is under 0.2% when scheme 2 is used instead of scheme 1. It seems that, generally speaking, the inclusion of age-sex groups has more impact on the estimates than the inclusion of wage and salary.

We also note that the distribution of the population according to wage and salary has been improved by using the wage and salary income group controls, as illustrated by Figure 2.1. The middle wage and salary income classes are better represented using calibration scheme 4 than using the current calibration. Low and high wage and salary incomes are also better represented even if the observed problems have not been completely solved.

Adding controls to the calibration procedure results in distributions of g-factor with heavier tails, which was predictable. The increase in the number of demographic controls had more impact on the g-factor distribution than the addition of controls on wage and salary income categories. Adding both types of controls simultaneously also had more impact on the distribution than adding any one of them separately, which again was predictable.

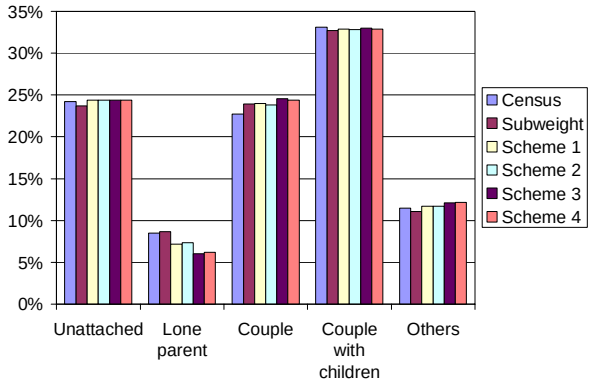
6. IMPACT ON HOUSEHOLD COMPOSITION DISTRIBUTIONS

One major drawback of increasing the demographic controls was the impact on household composition. Prior to the harmonized calibration project, a study was conducted to compare calibration schemes involving no control on household size and different number of age groups. Overestimation of the number of households of size 1 was observed when increasing the number of age groups. It was thought that the problem would be resolved when controlling for household size at the same time. The problem now seems to show up in the household composition distribution.

In Figure 6.1, we observe that the proportion of lone parent family household estimated using the SHS nonresponse adjusted and not calibrated weight (subweight) was quite similar to the one observed in the Census 1996. However, after calibration, this proportion goes down. Moreover, it is worst in the case of calibration schemes 3 and 4 where the number of demographic controls is increased.

The assumption is that, given that all the members of a household get the same final weight and given that globally the weights of adults have to be inflated more than the weights of children (as shown by the slippage rate in Table 2.1), households with a smaller ratio of children over adults have their weight inflated more than the weight of households with a larger ratio of children over adults. These ones may even decrease. Moreover, when controlling for household size, household consisting of couple without children will have their weight inflated to simultaneously compensate for under representing size 2 households and adults. This probably results in size 2 lone parent family households having their weights deflated or not inflated sufficiently. A similar phenomenon should be observed in size 3+ households.

Figure 6.1: Household Composition Distributions



7. IMPLEMENTATION OF THE NEW CALIBRATION SCHEME

Following the study, it was decided to implement a modified version of calibration scheme 4. As explained at the end of the previous section, the integrated weight methodology along with the use of a large number of demographic controls are key factors in deteriorating the household composition distribution. Further analysis is needed for a better understanding of the causes. To overcome the negative impact on the household composition distribution, two other sets of controls were included: provincial counts of lone parent family households and provincial counts of couple with children. These were obtained using projections of Census counts.

A few problems were experienced with the T4 files, which are used to produce the wage and salary income group controls. The province of residence variable on the T4 file was missing for a high proportion of records. The province of employment variable was used to impute the province of residence when missing. These two variables, however, do not necessarily correspond. During the production of the controls from the T4 files, it was discovered that the counts included a number of records with invalid identifiers and duplicates. Because of the tight deadlines, the T4 file did not go through a

major cleaning procedure before being used as controls. However the impact of the unclean data is considered small compared to the improvement of the income distribution provided by the new approach.

Another major issue concerning the T4 file is that the file corresponding to the survey year under processing is not available on time to be used. The T4 file corresponding to the year prior to the survey is used to produce controls for wage and salary income group. Based on LFS wage and salary data, trends are calculated between the two years and they are applied to the categories derived from the available T4 file in order to produce the controls.

The calibration scheme described above was implemented for each province but not for the Territories for several reasons. The sample size in the three territories was not sufficient to support as many controls. Another reason is that there was some evidence of poor quality for some of the controls for these regions.

In addition to its implementation for the SHS 1999, the new calibration scheme has also been applied to the 1997 and 1998 SHS data and to the 1996 FAMEX data in order to maintain historical comparability.

8. CONCLUSION AND FUTURE WORK

For SHS, the Harmonized Calibration project helped in keeping the momentum required to evaluate and implement a new calibration scheme, which improves the representativeness of the sample.

Between the weighting of the SHS 1999 data and the weighting of the SHS 2000 data, better editing procedures of the T4 files have been implemented and the postal code has been used as the basis for assigning a province of residence to each record.

The model that is used to derive wage and salary controls from the previous year T4 files is under evaluation. The results might lead to improvement of this model.

A better understanding of the factors that led to the deterioration of the household composition distribution, should be targeted. In particular we want to better understand the role played by the integrated weight methodology. Other weighting methodologies and solution should be investigated in the future. For example, a better understanding of the nonresponse mechanism and better methods to adjust for nonresponse would help in reducing the magnitude of the problem of sample representativeness.

ACKNOWLEDGEMENTS:

The authors would like to thank Michel Latouche and Ioana Schiopu-Kratina for their time and many helpful comments.

REFERENCES:

Arsenault, S., Beaulieu, S., Bernier, J.-L., Fillion, J.-M., Tremblay, J. and Vandermeer, B. (1999). 1997 Survey of Household Spending Quality Indicators. Household Survey Methods Division, Statistics Canada.

Arsenault, A. and Tremblay, J. (2001). Methodology of the Survey of Household Spending. Catalogue no. 62F0026MIE-01003, Statistics Canada.

Coombs, J., Champion H., Hole, G., Love, R., Macredie, I., Meere, M., Sheridan, M., Singh, M.P., Standish, L. and Webber, M. (1998). Report of the Task Force on Income Statistics. Statistics Canada.

Lemaître, G. and Dufour, J. (1987). An Integrated Method for Weighting Persons and Families. *Survey Methodology*, Vol.13, no. 2, pp. 211-220, Statistics Canada.

Webber, M., Latouche, M., Rancourt, É. (2000). Harmonized Calibration of Income Statistics. Statistics Canada.