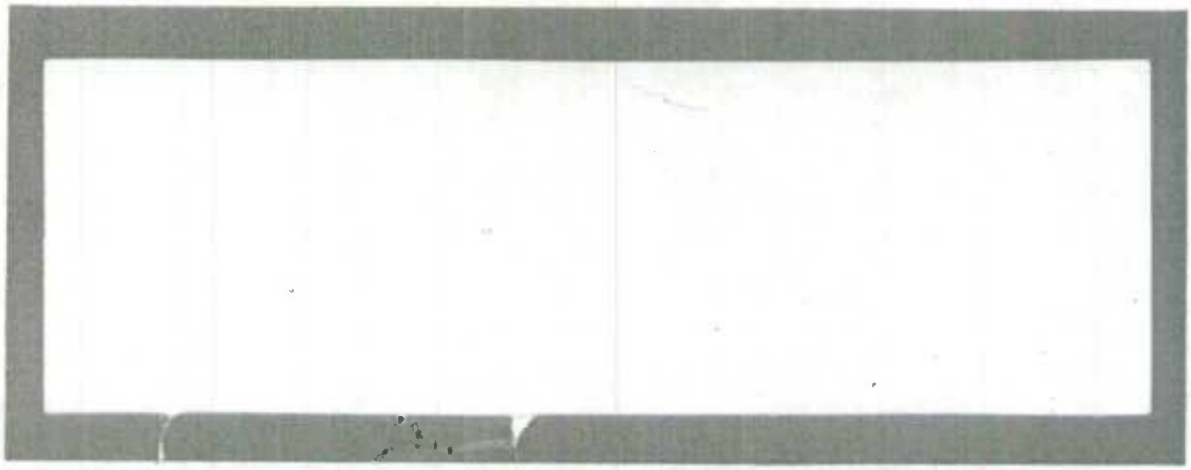

Methodology Branch

Household Survey
Methods Division

Direction de la méthodologie

Division des méthodes
d'enquêtes des ménages

Canada

WORKING PAPER
METHODOLOGY BRANCH

CALIBRATING THE SURVEY COMPONENT OF THE SOCIAL POLICY SIMULATION DATA BASE: A CASE STUDY IN VARIABLE SELECTION FOR CALIBRATION IN COMPLEX SURVEYS

HSMD - 2001 - 005E

Takis Merkouris and Khushnood Khan

Household Survey Methods Division
Statistics Canada

May 2001

The work presented in this paper is the responsibility of the author and does not necessarily represent the views or policies of Statistics Canada

Calibrating the Survey Component of the Social Policy Simulation Data Base: A Case Study in Variable Selection for Calibration in Complex Surveys.

Takis Merkouris and Khushnood Khan[1]

ABSTRACT

This paper describes the revised calibration scheme for the Survey of Consumer Finances component of the Social Policy Simulation Data Base (SPSD), involving survey data for the reference year of 1997. For later reference years this calibration will be applied to data from the Survey of Labour and Income Dynamics, the new survey component of SPSD. The revision included in the calibration income variables for the first time, and was carried out in harmonization with the new calibration schemes of the current income and expenditure surveys of Statistics Canada. The extensive empirical study conducted for the revised calibration included subject-matter and statistical considerations related to the usefulness of the various auxiliary calibration variables, analysis of slippage rates for these variables, as well as various diagnostics related to the effect of the revised calibration on the survey weights and on various estimates and their variances. In the literature on calibration in surveys very little formal attention is devoted to the selection of the set of auxiliary variables to be used in the calibration procedure. A formal approach in the context of regression estimation is only tractable in simple random samples. In this paper issues concerning multi-way calibration of survey weights are discussed in detail, and a practical guide for the selection of an effective set of calibration variables according to suitable criteria is put forward. Thus, this fairly comprehensive empirical study can serve as a case study in variable selection for calibration in complex surveys.

Key words: Calibration, residual slippage rate, weight adjustments, control totals.

---

[1]Takis Merkouris and Khushnood Khan, Household Surveys Methods Division, Statistics Canada, Ottawa, Ontario, K1A 0T6

Calage de la composante d'enquête de la Base de données de simulation de politique sociale : Une étude de cas de la sélection de variables pour le calage dans des enquêtes complexes.

Takis Merkouris et Khushnood Khan[2]

RÉSUMÉ

Le présent exposé décrit le schéma de calage révisé pour la composante Enquête sur les finances des consommateurs de la Base de données de simulation de politique sociale (BDSPS), comprenant des données d'enquête pour l'année de référence 1997. Pour les années de référence subséquentes, ce calage s'appliquera à des données de l'Enquête sur la dynamique du travail et du revenu, la nouvelle composante d'enquête de la BDSPS. La révision a intégré au calage des variables du revenu pour la première fois, et elle a été effectuée en harmonie avec les nouveaux schémas de calage des enquêtes courantes sur les revenus et les dépenses de Statistique Canada. L'étude empirique poussée menée pour le calage révisé a comporté des considérations spécialisées et statistiques liées à l'utilité des diverses variables de calage auxiliaires, une analyse des taux de glissement pour ces variables, de même que divers diagnostics liés à l'effet du calage révisé sur les poids d'enquête et sur diverses estimations et leur variance. La documentation sur le calage dans les enquêtes accorde très peu d'attention à la sélection de la série de variables auxiliaires à utiliser pour le calage. Une stratégie formelle dans le cadre de l'estimation par régression n'est possible que moyennant des échantillons aléatoires simples (Silva et Skinner, 1997). Dans le présent exposé, les questions liées au calage à plusieurs voies des poids d'enquête sont abordées en détail, et un guide pratique de sélection d'une série efficace de variables de calage est présenté en fonction de critères appropriés. Ainsi, cette étude empirique relativement détaillée peut servir d'étude de cas pour la sélection de variables en vue du calage dans les enquêtes complexes.

Mots clés : Calage, taux de glissement résiduel, rajustements de poids, totaux de contrôle.

---

[2]Takis Merkouris et Khushnood Khan, Division des méthodes des enquêtes auprès des ménages, Statistique Canada, Ottawa (Ontario) K1A 0T6

1. Introduction

This document describes the revised calibration scheme for the Survey of Consumer Finances (SCF) component of the Social Policy Simulation Data Base and Model (SPSD/M). Calibration is a weight adjustment procedure whereby the survey weights are adjusted so as to make the estimated totals of certain auxiliary variables equal to known population totals (control totals) for those auxiliary variables[3]. The new calibration scheme incorporates as controls the counts in certain groups of wage earners, as well as the count of earners of self-employment income (nonfarm). Changes in the set of demography controls for SCF have also been made, most notably the elimination of the economic region controls and the addition of economic-family size controls. The new calibration scheme for the SCF has been revised in harmonization with the new calibration schemes of the other income and expenditure surveys of Statistics Canada. The studies referred to below in relation to the calibration of the SCF weights were based on data from the SCF reference year of 1997, the last year of this survey. The revised calibration scheme is to be used in the Survey of Labour and Income Dynamics (SLID), which has since replaced the SCF. As usual, the revised calibration was carried out separately for each province.

The various demography and income controls are described in Section 2, along with subject-matter and statistical considerations related to the usefulness of the various controls and to the choice of the particular calibration groups. The results of a study on the size and statistical significance of the "slippage" for the new calibration variables are presented in Section 3. Diagnostics related to the impact of the calibration on the weight adjustment factors are described in Section 4. The effect of the revised calibration on various estimates and on the variance of these estimates is discussed in Section 5. Concluding remarks are made in Section 6.

2. Calibration Controls for SPSD/M

2.1 *Demography controls*

The demography variables used in the calibration for the SPSD/M were **age/sex, Census Metropolitan Areas/Urban Centers, sample rotation, and economic-family size**. All demography control totals used in the final calibration of the SCF data are based on the 1996 Census, and correspond to the provincial population totals for January 1998 — instead of the totals for the traditional survey month of April — in harmonization with the other income and expenditure surveys. A few minor differences between the calibration scheme for the SPSD/M and the harmonized calibration scheme are due to the ready availability of the controls at the time of the SPSD/M production, and to the special requirements of the SPSD/M. All demography controls exclude the military personnel living in households (and in barracks). Since this segment of the population can be identified in the SCF sample, it is also excluded from the calibration. In contrast, the income controls include the military of all types.

---

[3] The calibration procedure referred to here is based on the method of generalized regression.

Age/Sex

The number of age/sex groups has been reduced to 24, from the 30 groups traditionally used by the SCF and the Labour Force Survey (LFS); see Table 1. These 24 age/sex groups are currently used by the LFS (historical revision in January 2000), from which they were readily obtained in time for the SPSD/M production.

The harmonized calibration scheme to be used by the other income and expenditure surveys involves at least 18 age/sex groups; for the SLID there will be 22 groups. The difference in the age grouping between the SCF and the SLID is mainly in the lower tail of the age distribution; see Table 1. The specification of the young-age groups for the SLID was based on subject matter requirements, while the reason for dropping certain groups from the old SCF scheme was the concern for an excessive number of age/sex controls. For future production, using SLID data, it may be advisable to monitor the survey slippage[4] for age, and modify the number of groups if necessary. Slippage will always be a more serious concern for the SLID than was for the SCF. Periodically one SLID panel will be coming from an old LFS design. Also, unlike the LFS, the SLID does not incorporate cluster-growth and stabilization weight adjustments at each wave. For surveys of the size of the SLID a uniform age/sex grouping of 22 or 24 groups for all provinces is all right, although the sample count in one or two old-age groups is barely sufficient in Newfoundland and PEI.

Table 1. Age grouping, by sex

SCF	0-14	15-16	17-19	20-24	25-29	30-34	35-44	45-54	55-59	60-64	65-69	70+
SLID	0-6	7-15	16-17	18-24	25-34		35-44	45-54	55-59	60-64	65-69	70+

Sub-provincial geography

All 26 Census Metropolitan Areas (CMA's) and 38 Urban Centers (for people of age greater than fifteen) previously used in the SCF calibration were used again in the calibration of the SCF 1997 data. For the purpose of the harmonization studies all income and expenditure surveys were (as of May 15, 2000) to use only the 16 CMA's included in the calibration for the Survey of Household Spending (SHS). Subject matter specialists were to update the list later. Differences between surveys could be possible. Since then, the SHS has decided to use 16 CMA controls for the populations of ages below and above eighteen; this will double the number of the CMA controls. The SLID will use the same number of CMA controls, but not necessarily for the same CMA's.

Reasons for including CMA's in the calibration include subject matter considerations (i.e., level of publication of income statistics, strong correlation with income variables — especially the Low Income Cut off (LICO)), and correction of the sample coverage of the CMA population. Note that the sampling design variable is geographical, so that any sample deficiency or deterioration

[4] Slippage here denotes the discrepancy between the actual and the estimated population total for a particular characteristic.

over time is reflected directly on the sample geography. This must be especially true in the SLID because of undercoverage, attrition and moves within its panels. Inclusion of CMA in the calibration is expected to improve as well the precision of income estimates at that level. Note also that the assumption for the CMA as calibration variable in the SLID is that a valid code is available for each individual in either panel (including movers and cohabitants), whether it comes from an old or the current LFS design. This CMA code should match the definition of the corresponding "current" CMA control. Using controls for urban centers in the SLID may improve the representativity of its panels.

The controls for Economic Regions (ER) have been dropped from the calibration. There are no compelling subject matter or methodological reasons to include them in the income and expenditure surveys. Note that the non-CMA aggregate geography will automatically be controlled. With little additional auxiliary information the ER's (72 in total) might do more harm than good, in terms of weight distribution and variability of estimates.

Rotation groups

The control totals for the four rotation groups (based on the LFS design) comprising the SCF sample were retained in the revised calibration scheme. For balancing the sample, the total weight (for all ages) of each of the four rotation groups was controlled to one fourth of the total (provincial) population.

This type of weight control is applicable only to the SCF (LFS) data. Although the panels of the SLID are based on LFS rotation groups, for a number of technical reasons related to the dynamic nature of the SLID it is not sensible to use rotation groups in the calibration of this survey.

Economic-family size

On subject matter recommendation, counts for three economic-family size groups were included in the calibration, namely, counts of families of size one, two and more than two.

Reasons for incorporating family size in the calibration include large slippage rates for family size counts, consistency of estimates at person and family levels, and improved precision due to high correlation of family size with income characteristics. For consistency with the corresponding controls the SCF family size variable was modified to exclude the military.

2.2 *Income controls*

Wage /Salary Income

Large differences between the estimated distribution of wage/salary earners and the corresponding distribution derived from the revenue T4 file have been observed over the last few years. Analogous differences have also been observed for the distribution of the wages. To improve the estimation of the distribution of both the count of wage earners and the wage earnings, six calibration groups of wage/salary earners for each province were used, with group intervals defined by the 25^{th}, 50^{th}, 65^{th}, 75^{th}, and 95^{th} /98^{th} /99^{th} percentiles. The upper percentile was the 95^{th} for Newfoundland and PEI, the 99^{th} for Quebec and Ontario, and the 98^{th} for the rest of the provinces. The upper percentile for

each province is the highest possible for controlling the underestimation of the count of high-wage earners, given the available sample size of the corresponding calibration group. The group intervals corresponding to the percentiles vary by province.

Although the wage/salary control totals include individuals of age less than fifteen, in the SCF data only those of age greater than fourteen are wage/salary earners.

An extensive study (based on SCF data from the years 1996, 1997) of various options regarding the extent of calibration has lead to the choice of a small number of calibration groups, namely six, in order to avoid potential negative impact on statistical properties of derived estimators. On a related point, note that a distinguishing feature of this calibration variable is the non negligible response error in reporting the amount of wage/salary. Broad wage groups reduce the risk of a misclassification error in the calibration set-up.

Control totals based on percentiles were used in order to achieve harmonization of all income and expenditure surveys with respect to standard distributional features of the controlled variable across provinces and across time. "Optimal" control totals for each individual survey were only slightly better in terms of matching the distribution of wage earners. Evidently, practicality is a major advantage of the percentile-based controls. Moreover, accurate net annual differences for the percentiles, more importantly the median, are thus obtained.

The first group, defined by the 25th percentile, excludes earners of a wage/salary less than \$500. These earners were excluded in order to avoid outlying calibration weight adjustments, and also because the reported amounts in that range are not reliable. The percentiles had to be adjusted to reflect this exclusion. A publication footnote giving the operational definition of the percentiles would thus be in order.

Calibration groups for the wages were not used for the following reasons. To begin with, the calibration to counts of wage earners results in a substantial improvement of the estimation of the distribution of the wages as well; see discussion below. Furthermore, the practice in household surveys is to use for calibration only auxiliary count variables, thereby correcting the representation of people in the corresponding controlled groups. Controlling the count (in certain groups) of wage earners can be justified as a further balancing of the sample with respect to this survey characteristic. However, including also the continuous variable wage stretches the calibration to control a main survey characteristic, thus blurring the distinction between survey and administrative data. This case of over-calibration, with respect to the type of controlled variable, is the main qualitative reason for excluding such a variable from the calibration. On more technical grounds, note that given the very wide range of wage values at the micro level there is a risk of outlying calibrated weights, and hence bias of derived estimators. For the same reason, the resulting (regression) estimators are unstable (having a large sampling variance) for characteristics other than the controlled ones. Finally, serious numerical difficulties are encountered in calibrating values of a continuous variable with as wide a range of values as that of the wage variable.

Self-employment/nonfarm income

One group of earners of self-employment (nonfarm) income greater than \$500, for each province, was added to the calibration of the SCF data to improve the estimation of this income characteristic. The group with this type of income between -\$500 and \$500 was excluded for the same reasons as

noted above for the wage/salary controls. Controlling self-employment/nonfarm is a special SPSD/M requirement — it is not part of the harmonized calibration scheme.

The option of adding the group with income below -$500 was eliminated because of inadequate sample size, and because of the unacceptably large calibration weight adjustments that would result from correcting the extreme underestimation of this domain. The use of one calibration group for all self-employment (nonfarm) incomes, excluding only those in the range -$500 to $500, was also eliminated because it increased the variability of the calibrated weights and because of its undesirable effect on slippage. Table 2 below presents slippage rates at Canada level for three groups of earners of self-employment/nonfarm income for four different calibration schemes, starting with the demographic controls, and progressively adding the wage controls, the control for self-employment/nonfarm income greater than $500, and, alternatively, the control for self-employment/nonfarm income less than -$500 and greater than $500. The second column shows the rate of underestimation of the three groups by the SCF data even after controlling the demography, while the second column shows reduction of the slippage resulting from adding the wage control. As is evident from the last two columns, the slippage is eliminated for the controlled self-employment group, but while controlling the group with income in the range [500,) reduces the slippage of the group with income in (, -500]∪[500,) from 31.64% to -22.68%, controlling instead the later group results in overestimation of the [500,) group by 27.62%. Note that the control for the [500,) group comprises 72.50% of the control for the (, -500]∪[500,) group.

Table 2. Slippage rates for earners of self-employment (nonfarm) income. Canada level

SELF_EMP	DEM	DEM +WAGE	DEM +WAGE +SELF_EMP>500	DEM +WAGE +Abs(SELF_EMP)>500
>500	-18.73	-12.42	0	27.62
<-500	-82.6	-82.31	-82.45	-72.81
<-500 and >500	-36.3	-31.64	-22.68	0

The problems associated with the inclusion of earners with self-employment/nonfarm income less than -$500 in the calibration of the SCF data will be encountered also in all income and expenditure surveys if counts of earners of total income are used in the calibration. In addition to these problems, the exclusion of any problematic range, say (, -500]∪[500,), for total income will necessitate an adjustment of the percentiles for the distribution of earners of total income, which will effectively treat incomes in that range as zeros. This will have implications in the calculation of such income statistics as means, medians and LICO's. Note that in the SCF the negative incomes were treated as zeros in the calculation of LICO's at the family level, while zero-income units were excluded from the calculation of means and medians.

3. The slippage of SCF estimates for the new calibration variables

The magnitude of the discrepancy between control totals and their estimates for the new calibration variables (i.e., family size, wages and self-employment/nonfarm) has been assessed by computing the corresponding residual slippage rates. These slippage rates are defined as the residual relative differences between the control totals for the new calibration groups and the corresponding estimated counts that already incorporate the calibration to the rest (old set) of the controls. The assessment of the statistical significance of the discrepancies between controls and their estimates is based on the standardized (by the standard error of the estimates) observed differences. A simple working rule (based on Normal distribution theory) is to consider an observed difference as significant if its magnitude is more than twice its standard error. The residual slippage rates and the standardized differences for the new calibration variables, denoted by %(est-cnt)/cnt and (est-cnt)/se, respectively, are shown in Table 3 for Canada, and Table 4 for Ontario.

Table 3. Residual slippage rates (Canada)

Count Variable	%(est-cnt)/cnt	(est-cnt)/se
fmsz		
1	4.88	3.27
2	6.40	6.09
3+	-0.50	-0.95
total	3.18	8.57
wage		
total (>=500)	3.75	9.81
self-empl		
[500,)	-18.57	-11.11
(, -500]	-82.55	-64.46
total	-36.16	-29.2

At Canada level, the SCF using the "basic" calibration scheme (including the demographic controls — except family size) overestimates the number of families of size one and two by 4.88% and 6.40%, respectively. The differences are statistically significant, as they correspond to 3.27 and 6.09 standard errors. The total number of families is overestimated by 3.18%, the difference being 8.57 times its standard error. The count of earners of a wage/salary of more than $500 is overestimated by 3.75%, which corresponds to 9.81 standard errors. The counts for the two self-employment (nonfarm) income groups, namely more than $500 and less than -$500, are grossly underestimated, with the differences being statistically very significant, as shown in Table 3. The count for the combined group of self-employment (nonfarm) is underestimated by 36.16%.

The residual slippages for the various provinces are in the same direction, and more or less of sizes comparable to those at the Canada level. Table 4 shows the slippage rates for Ontario. In particular,

it reveals the typical pattern of slippages for the six wage groups that correspond to the chosen percentiles. The lower tail of the distribution of wage earners is underestimated, the middle range is overestimated, and the upper tail is grossly underestimated. The observed differences for all six groups are statistically quite significant. A graphical comparison, at Canada level, of the estimated distribution of wage earners with the corresponding distribution based on the T4 data is provided by Figure 2.

Table 4. Residual slippage rates (Ontario)

Count Variable	%(est-cnt)/cnt	(est-cnt)/se
fmsz		
1	6.51	2.22
2	4.06	2.17
3+	0.18	0.20
total	3.18	4.59
wage		
[500, 10500)	-8.27	-4.47
[10500, 26500)	6.86	3.11
[26500, 35500)	15.49	4.98
[35500, 43500)	15.44	4.00
[43500, 129500)	5.69	2.32
[129500,)	-34.89	-3.94
total	4.36	6.68
self-empl		
[500,)	-16.45	-5.44
(, -500]	-81.54	-35.26
total	-36.41	-17.1

It has been observed by analysts of the Social and Economic Studies Division that when the design (uncalibrated) weights are used, SCF underestimates the low-earning population, but estimates reasonably well the middle-and higher-earning population. It appears then that the calibration of the design weights to the demography controls (excluding the family size controls) increases substantially the size of the middle-earning population (see Figure 2), thereby resulting in the overestimation of the total count of wage earners, as shown in Table 3.

In a systematic study of the proposed calibration scheme, rates of residual slippages resulting from adding in a step-wise manner a new calibration variable at a time were computed. The effect of controlling the family size on the wage/salary slippage was negligible, while the effect of adding the

wage/salary controls on the self-employment/nonfarm slippage, as well as the effect of adding the single self-employment/nonfarm control on the combined self-employment group (, -500]∪[500,) have already been discussed, and shown in Table 2 above (columns three and four).

4. The impact of the calibration on the weight adjustment factors

The new calibration scheme has resulted in an appreciable increase in the variability of the weight adjustment factors (known as "g-weights"), as expected given the increase in the number of the auxiliary variables used in the calibration and the magnitude of the slippage associated with the additional calibration variables. It is to be noted that the g-weights should be tightly distributed around one for the resulting estimators to have good statistical properties. The impact of each of the new calibration variables on the distribution of the g-weights at Canada level is depicted in Figure 1.

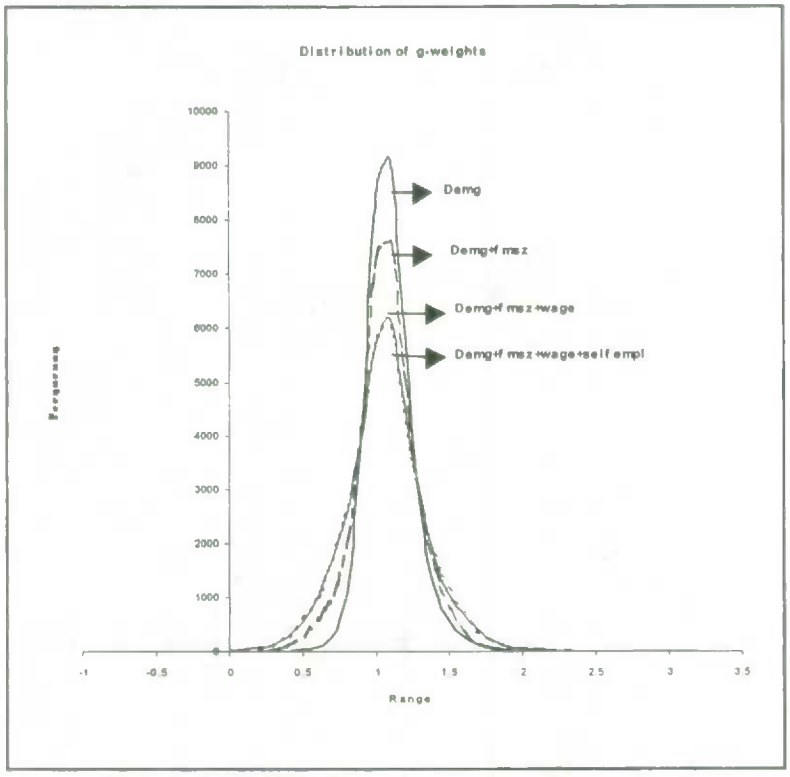

Figure 1. Distribution of g-weights, Canada level

It is observed that the addition of the family size controls (to the basic demographic ones) flattens the distribution of the g-weights, and the addition of the wage/salary controls flattens it even more. On the other hand, the addition of the self-employment/nonfarm control has very little effect.

Some descriptive statistics of the g-weights, at Canada level, are shown in Table 5. Notably, most of the increase in the variability (as measured by the CV) of the g-weights is due to the addition of

the family size controls, while the single self-employment/nonfarm control adds very little variability to that already added by the wage/salary controls. The family size controls also increase the range of the g-weights, because they produce some negative weights. The range is further increased by the wage/salary controls because of some large positive g-weights. The family size controls result into a total of seven negative g-weights, while adding the wage/salary controls results into a total of twenty negative g-weights. No additional negative weights result from the addition of the self-employment/nonfarm control. Most of the negative weights are in Alberta. The large positive g-weights (most of them in Alberta) produced by the wage controls are due to the correction of the gross underestimation of the upper tail of the wage distribution. With all new controls included, at Canada level 99% of the g-weights are below 1.83.

Table 5. Descriptive statistics of g-weights, Canada level.

	MEAN	CV	MIN	MAX
DEM	1.09	13.94	0.34	2.45
DEM+FMSZ	1.09	20.11	-0.61	2.33
DEM+FMSZ+WAGE	1.09	24.42	-0.55	3.16
DEM+FMSZ+WAGE+SELF_EMP>500	1.09	25.31	-0.53	3.12

The effect of the various controls on the variability of the final calibrated weights is shown in Table 6. Interestingly, the addition of the new controls increases the variability of the final weights at a rate that is lower than the rate corresponding to the variability of the g-weights. Table 6 also demonstrates the unpredictable effect of outlying g-weights, that is, the potential coupling of an outlying g-weight with a relatively large design weight. For instance, although the largest g-weights for the last two calibration schemes are practically the same (3.16 and 3.12), the largest final weights differ substantially (2561.96 and 3298.91). With all new controls included, at Canada level 99% of the final weights are below 1376.

Table 6. Descriptive statistics of final weights, Canada level

	MEAN	CV	MIN	MAX
DEM	339.3	79.94	19.75	2339.8
DEM+FMSZ	339.3	82.07	-219.8	2087.4
DEM+FMSZ+WAGE	339.3	84.22	-165.6	2562
DEM+FMSZ+WAGE+SELF_EMP>500	339.3	84.9	-161.7	3298.9

Large variation in g-weights may cause an increase in the variance of derived estimators of characteristics that are not strongly correlated with the controlled variables, possibly offsetting the decrease in variance effected by the incorporation of the auxiliary information through calibration. Since the variation of g-weights generally increases with the number of control totals, caution was exercised in the selection of the set of controls for the final calibration scheme. In this connection note that although new controls have been added, other controls have been eliminated, namely, some age group controls and all economic region controls. The number of controls in the new calibration scheme is smaller than in the old one in most provinces.

Large g-weights may also result in biased estimators of characteristics of interest associated with subpopulations that contain units with such large weight adjustments.

Whereas the impact of g-weights on variances of estimators of characteristics of interest can be assessed directly by computing those variances (see Section 5), the potential bias effect is neither analytically tractable nor in general empirically detectable. Judicious selection of the calibration groups provides protection against potential biases due to outlying g-weights or due to calibration groups of very small sample size. An example of such judicious choice is the exclusion of the group of earners of self-employment/nonfarm income of less than -$500 from the calibration; see discussion in Section 2.

Finally, the following (standard in the LFS) procedure was used for handling the negative weights. The households with negative calibrated weights were identified. Those negative weights were substituted by the original design weights for all household members, and the complete new set of weights (calibrated and substituted) was then recalibrated. Except restoring the original positive weights, the effect of the recalibration was otherwise negligible.

5. The impact of the new calibration on estimates and their variances

The effect of the new calibration scheme, at Canada level, on the distribution of wage/salary earners and on the distribution of wages is shown in Figures 2 and 3, respectively. The new calibration scheme improves considerably the estimation of the frequency distribution, by wage classes, for both

wage/salary earners and wages. This is more so for the middle-income range that suffers most from overestimation, while estimation of the upper tail of the frequency distribution is only marginally improved. Note that the distributions corresponding to the old calibration scheme (SCF:weights-published) shown in Figures 2 and 3 are based on the Census 1991 demography controls for the survey month of April.

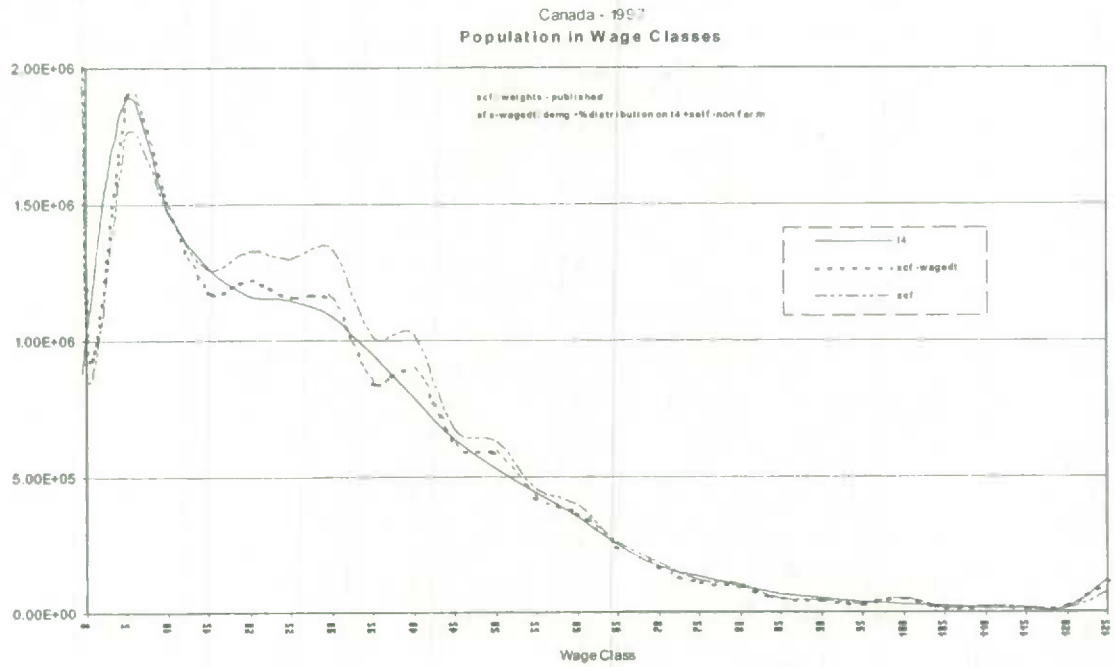

Figure 2. Distribution of wage/salary earners, Canada level

Table 7 summarizes the effect of the new calibration scheme on the estimates and their variances for a few selected income variables, including the calibration income variables (i.e., wage and self-employment/nonfarm income), as well as the total investment income, the pension income and the total income. The table shows relative differences between estimates, of counts and total earnings, before and after the addition of all new calibration variables. Relative differences for sampling variances and differences of CV's are also shown.

For wages/salaries larger than \$500 the estimated count is reduced by 3.6%, and the estimated total earning by 5.7%. On the other hand, the estimated counts for self-employment/nonfarm income larger than \$500 and total (larger than \$500 and smaller than -\$500) are increased by 22.8% and 21.1%, respectively; comparable are the increases for the corresponding estimated total earnings. The reduction of sampling variance is 100% for both controlled counts, and 91.5% for the total count of self-employment/nonfarm, while the reduction is substantial, 87.6%, only for the total of wage earnings. Substantial is the reduction in CV only for the counts of the self-employment (nonfarm) income larger than \$500 and total, by 2.05% and 1.47%, respectively. For the rest of the studied income variables, the only appreciable effect is the increase of the estimated total investment income

by 1.3%, and the decrease of the estimated total income by 2.3%. The change in the CV's of these estimates is insignificant.

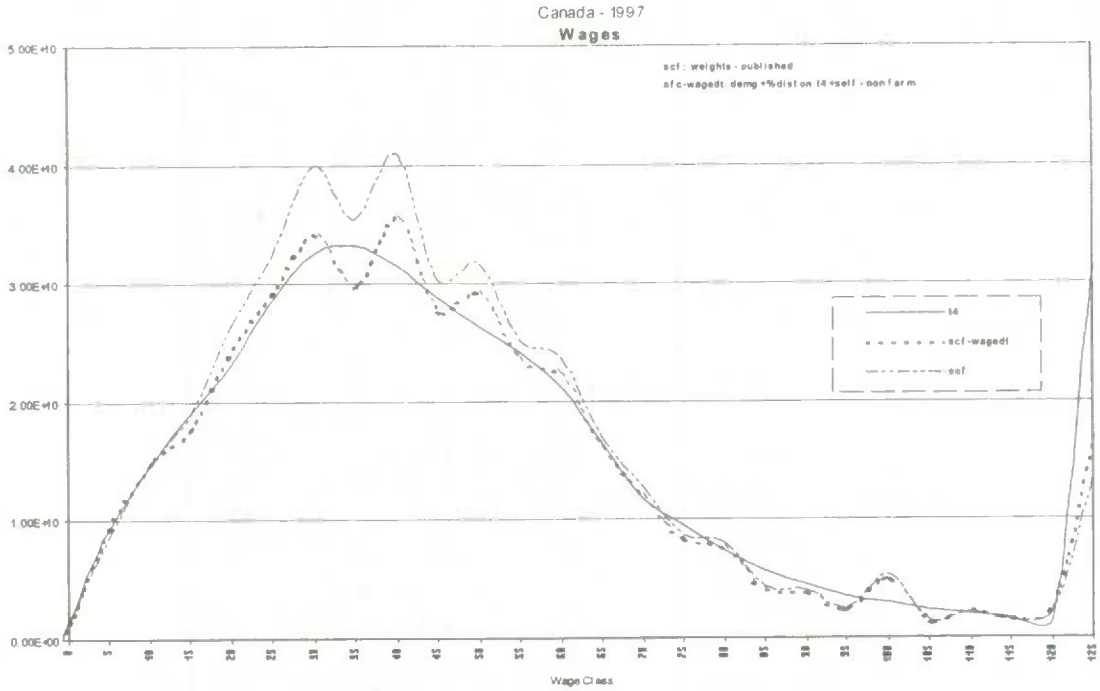

Figure 3. Distribution of wages, Canada level

Table 7 shows the compound effect of all additional calibration variables on income estimates. A more detailed analysis (based on similar tabulations not shown here) reveals that the family size controls have no appreciable effect on the studied income estimates. On the other hand, the wage controls increase the estimated counts and total earnings for the self-employment/nonfarm income by 7% and 9%, respectively, for both the larger than $500 and total (larger than $500 and smaller than -$500) incomes. The rest of the increase in those estimates, as well as all the reduction of the corresponding sampling variance is due to the self-employment control. All other substantial effects on income estimates and their variances noted above are primarily due to the wage controls.

Table 7. Effect of the new calibration on income estimates, Canada level

B=dem; A=dem+fmsz+wage+self-empl						
	Count			**Earning**		
Variable	% B-A / B	% V(B)-V(A) / V(B)	CV(B)-CV(A)	% B-A / B	% V(B)-V(A) / V(B)	CV(B)-CV(A)
wages>=500	3.6	100.00	0.37	5.7	87.6	0.42
sempl_nf>=500	-22.8	100	2.05	-23.2	2.4	0.68
sempl_nf<=-500	-0.6	-5.1	-0.14	-1.2	-7.7	0.32
total sempl_nf	-21.1	91.5	1.47	-23.7	2.7	0.73
tinvinc	0.2	-3.5	-0.02	-1.3	6.2	0.17
peninc	-0.6	-12.9	-0.05	-0.8	-13.1	-0.05
totalinc	0.2	10.3	0.01	2.3	52.7	0.15

Table 8. Effect of the new calibration on wage counts and total earnings, Ontario

B=dem;A=dem+fmsz+wage+self-empl								
	Count				**Earning**			
wage	% B-A / B	% V(B)-V(A) / V(B)	CV(B)	CV(A)	% B-A / B	% V(B)-V(A) / V(B)	CV(B)	CV(A)
[500, 10500)	-9.01	100	2.01	0	-9.22	47.59	2.47	1.64
[10500, 26500)	6.42	100	2.07	0	6.39	93.84	2.12	0.56
[26500, 35500)	13.42	100	2.69	0	13.46	99.41	2.7	0.24
[35500, 43500)	13.38	100.00	3.35	0	13.37	99.77	3.35	0.19
[43500, 129500)	5.38	100.00	2.32	0	5.37	94.06	2.39	0.62
[129500,)	-53.58	100.00	13.59	0	-54	70.01	14.64	5.21
total	4.17	100	0.62	0	5.28	86.39	1.13	0.44

The effect of the calibration on estimates of counts and total earnings for the six wage calibration groups is illustrated for Ontario in Table 8. It is observed that the relative differences for the estimated earnings are very similar to those for the estimated counts. Furthermore, besides eliminating the differences between estimated counts and corresponding controls for the six wage groups, with the implied reduction of sampling variability by 100%, the calibration of the wage counts results as well in a dramatic reduction of the sampling variability of the estimated total earnings for the six groups, as indicated by the relative differences of the variances and the decrease in the CV's.

An interesting complement to Table 8 involves differences between estimated wage totals and the corresponding totals in the T4 file, before and after the introduction of the six wage count controls. Table 9 shows the relative differences (with respect to the estimated totals) and their statistical-significance. The relative differences for the wage totals are very similar to those for the wage counts, shown in table 8, for all but the last open ended group. The last group represents 1% of the wage earners but 9.32% of the total wage in Ontario. The calibration to the wage counts reduces the relative differences for the wage totals substantially for all six groups. The difference for the first and third groups is no longer statistically significant. On the other hand the calibration shifts the sampling weight from the overestimated middle range to the upper tail of the distribution of the wage earners, which changes the relative difference for the total wage from a statistically non-significant 1.58% to a statistically significant -3.90%.

Table 9. Relative differences between estimated and true wage totals, Ontario

	Before		**After**	
	%		%	
wage	E-T / E	E-T / SE(E)	E-T / E	E-T / SE(E)
[500, 10500)	-10.54	-4.27	-1.21	-0.74
[10500, 26500)	8.37	3.95	2.11	3.77
[26500, 35500)	13.31	4.93	-0.17	-0.72
[35500, 43500)	13.87	4.14	0.57	3.02
[43500, 129500)	4.42	1.85	-1.01	-1.63
[129500,)	-151.94	-10.38	-63.60	-12.21
total	1.58	1.4	-3.9	-8.86

The effect of the new calibration, at Canada level, on estimates of a few selected count variables is shown on Table 10, which has the same format as Table 7. The estimated count of workers in the classes of "paid/large business" and "paid/government business" is reduced by 3.2% and 6.1%, respectively. This may well be explained by the relationship of these count variables with the calibrated count of wage earners. Indeed, the reduction in the count of workers in these two classes

is entirely due to the wage controls. Note that about 60% of all workers belong to one of these two classes, while about 30% of workers belong to the last class. The last class includes people who have never worked and people who worked for the last time more than one year before the reference year.

Table 10. Effect of the new calibration on count estimates, Canada level

B=dem; A=dem+fmsz+wage+self-empl			
	%	%	
Count Variable	B-A / B	V(B)-V(A) / V(B)	CV(B)-CV(A)
class of worker			
paid/large business	3.2	48.7	0.12
paid/government business	6.1	14.1	0.02
self-empl/nincorp/paid help	-23.4	-41.9	0.16
self-empl/nincorp/no paid help	-18.4	25.7	0.57
unpaid family worker	-19.9	-49.2	-0.17
Other	-2.7	58.6	0.24
marital status			
married	-0.7	38.5	0.1
single	-1	25.6	0.11
widow	1.4	-1.5	-0.04
separated/divorced	8.8	15	-0.03
household size			
1	3.6	55.5	0.46
2+	2.5	70.1	0.13
total	2.8	71.6	0.14
family type			
2	3.3	66.3	0.43
3+	1.8	56.8	0.17
total	3.1	100.00	0.36

The total count of workers is equal to the population total of individuals of age greater than fourteen. This population total is an aggregate age control total (see Table 1), and remains constant when more

control totals are added to the calibration. The effect of this is that the reduction of the counts in the first two classes of workers necessarily implies a shift of the aggregate weight to the other classes, which as shown in Table 10 results in increased counts in all these other classes.

The large increases in the estimated counts of the two types of self-employed workers, by 23.4% and 18.4%, is mostly due to the self-employment/nonfarm control, with the rest of the decrease effected by the wage controls. These findings are similar to those presented above in connection with Table 7.

The large increase, by 19.9%, in the estimated count of unpaid family workers is mostly due to the family size and wage controls, and to a lesser degree due to the self-employment/nonfarm control.

The large decrease, by 8.8%, of the estimated count of separated/divorced individuals is almost entirely due to the family size controls.

The substantial decreases in the estimated counts for the household sizes and family types shown in Table 10 are due to the family size controls.

Although large reduction of the sampling variance for most of the estimated counts is observed in Table 10, the reduction of the CV's is very small.

6. Concluding Remarks

The empirical study for the revised calibration scheme for the SCF 1997 data has confirmed the need for the incorporation of the family size controls and the controls for the two income variables. A judicious choice of income calibration groups has been made, which improved the estimation of the distribution of the controlled income variables while minimizing negative effects of the calibration. Various diagnostics for the effect of the new calibration scheme on the weights, on estimates of a few selected variables and on the variances of those estimates have shown no appreciable negative impact of the calibration. In particular, although the addition of the family size controls and the two income controls increase the variability of the weight adjustment factors, the change in the sampling variability (in terms of CV) of various estimates is negligible, with the exception of estimates that are directly related to the controlled variables — for which the gains in efficiency are substantial. Furthermore, changes in the estimates for certain characteristics are explainable, given the strong correlation of these characteristics with the controlled variables. Concerning potential bias effect of large weight adjustments, based on the present study there is no evidence of such effect on the limited number of study variables. More conclusive results would involve a more detailed study of all the main estimates of subject-matter interest, including estimates at subprovincial levels — at which any adverse calibration effects are more likely to manifest themselves. Of course, definite answers to such concerns are possible only to the extent of the availability of relevant benchmarks. As for the benchmarks used in the calibration of the SCF 1997 data, it has been assumed that the various controlled totals were highly accurate.

Although the revision of the calibration scheme for SPSD started with the SCF 1996 data, the empirical study of the new harmonized scheme has been based on the SCF 1997 data. The particular grouping for each of the new calibration variables is more substantively than empirically motivated; see discussion in Section 2. On the other hand, the specific grouping for the basic demographic

variables could be more adaptive to changes in sample coverage or sample size over time. It is advisable to monitor the slippage (calculated using design weights) separately for each calibration variable for changing patterns over time, especially when the sampling design changes or when revised demography controls are used. The quality of any new controls should be confirmed. It should be emphasized that elimination of gross coverage problems for some variables through calibration carries the risk of distorting the distribution of other variables; an instance of this involving the count of wage earners was noted at the end of Section 3. This may necessitate a preliminary ranking of the various potential calibration variables in order of importance, on the basis of expected calibration benefits, followed by a stepwise-forward calibration procedure in which a new calibration variable is added at a time until adverse calibration effects are detected. Such a procedure, which involves calculation of residual slippage rates at each step, will provide an insight into possible causes of the distortion of the distribution of variables of interest, whether they are survey design deficiencies or adverse effects of calibration to the usual demography controls. Clearly, addressing coverage issues at the design stage is the ideal course of action. Even then, though, calibration will still be important for ensuring consistency between estimated counts and known population totals (a weaker requirement than correction of coverage bias), for harmonization of comparable estimates between various surveys, and last but not least for the reduction of sampling variability.

Acknowledgment

The authors wish to thank Emile Allie for his valuable assistance on matters related to the study of the distribution of the income calibration variables and the choice of the calibration groups, as well as for producing the income controls and many useful graphics — including Figures 2 and 3.