SPSD/M

Growth and Validation Guide

This guide is designed to provide SPSD/M users with information relating to the growth of the SPSD/M to represent different years and the validation of SPSD/M results.

Statistics Canada Statistique Canada

Canada

Table of Contents

Introduction

SPSD/M's base year (2007) demographic data comprises a probability sample from provincial populations representing individuals and their living arrangements (i.e., individuals grouped in families and households). The demographic structure of the database is subsequently "grown":

1. By adjusting household weights according to the change in Statistics Canada's provincial age-sex population estimates from the base year to each succeeding year. Provincial age-sex counts used to generate household weights for off-base years come from Statistics Canada's Annual Population Estimates at July 1st, and from Population Projections for Canada, Provinces and Territories.   These estimates are re-based to December 31st in order to match the SPSD reference period prior to the weighting process. To derive the provincial population by age and sex for 2005, for example, the SPSD population is adjusted by the change in the Annual Population Estimates' population from the base year to 2005.

2. By simultaneous adjustment of household weights to agree with Labour Force Survey data on annual average unemployment rates and employment to population ratios by age and sex and by province and sex for the years in which survey data are available. For future years, the participation rate is kept constant while the unemployment rate by age, sex, and province are modified by the growth rate of the projected national unemployment rate which come from published estimates.

Statistics Canada's population estimates are revised on an on-going basis between Censuses,

consequently the population of off-base years may differ between versions of the SPSD/M.

In addition, database "growth" involves adjustment of dollar denominated components of income and tax to account for inflation, real growth and/or other trends affecting aggregate amounts. This is done by:

1. Adjusting individual dollar denominated items in order to bring the growth rates into line with growth rates from T1 tax data, National Accounts data, or published projections.

Static demographic and economic "growth" of this form assumes that the population remains unchanged in relative terms. For example, provincial population growth due to migration is accounted for by adjusting the base year household weights. That is, the characteristics (e.g., family size, income, etc.) of persons already resident in a province are used to represent the characteristics of migrants to that province. Thus, some aspects of the relative distributional impacts of variant tax/transfer models will represent the base year regardless of the nominal year in which they are implemented or the growth parameters used.

Static database adjustments force agreement with aggregate control totals by minimal change (if any) to the underlying micro data. More ambitious attempts to project micro data would have to explicitly assign new characteristics to individuals, families and households. These characteristics would have to represent labour market activity, personal finances and demographic events that are likely in light of 2007 micro data. Moreover, the assigned characteristics have to make sense in combination: migrants often change jobs and income levels and are more likely to be young and unmarried than the general population. Each modification made to micro data will either weaken or distort correlations that should be present or introduce spurious correlations. Keeping micro data changes to a minimum guarantees that the correlations present in the data have an empirical basis.

Projection Parameters

GROWTH FACTOR PARAMETER CREATION AND PROJECTION

For historical years, the year to year change in each growth factor parameter is forced to reproduce the growth of a stable T1 tax series - all on a provincial basis. This technique assures that each database variable never falls out of alignment with the growth of its proxy T1 series. For example, if the T1 self-employment income series for Ontario grows by 6% from 1997 to 1998, then the corresponding database variable, *idise*, grows at exactly the same rate for all Ontario residents.

There are two exceptions to this method. CPP/QPP income is adjusted using average benefit by year of age. This data comes from Human Resources and Social Development Canada. All SHS variables are grown using personal income which results in no distortion of the income/expenditure equality.

As T1 data is unavailable for subsequent and future years, we base the growth factors on projections done by the private sector provided to us by the Office of the Parliamentary Budget Officer. The income and expenditure parameters are grown by projected personal income.

MODEL PARAMETER PROJECTION

Projection of a model parameter only occurs when no official estimate is available. For example, the year 2016 model parameters governing old age security have not yet been released by the federal

government. Therefore, we must project these parameters for 2016 from the last year in which we have an official estimate.

The Office of the Parliamentary Budget Officer provided us with projected Average Weekly Industrial Wage, Personal Income, and CPI.

Users can determine how each model parameter is projected by examining the parameter's accompanying UPDATE statement in the .mpr files. For example, a FACTOR=CPIM3 statement means that the parameter will be projected at the CPI value minus three. Users can examine the *factors.txt* file in the *SPSD* directory for more information on the implicit growth underlying these three series.

SPSD/M Validation

Model validation has always been a critical component in the construction of the SPSD/M. The validation of results occurs for both the base year and for subsequent years in which the base year data is grown. Base year validation involves a reconciliation with many different sources of detailed data which are readily available. The underlying population is not being subjected to growth methodologies and many conceptual differences have been accounted for.

Validation becomes more difficult in subsequent projection years due to the increasing lack of availability of reference sources as well as the shortcomings of the static aging methodology to account for more variables and interactions. When dealing with simulation in future years, projections can only be validated with other projections. The further out in time one goes, the less a handful of aggregate growth factors applied can accurately account for complex shifting multivariate distributions in a population.

Nevertheless, there is a need for these projections and a need for some agreement on the "accuracy" of resulting estimates. There are three classes of reference standards which can be used for validation in either the database year or subsequent years; Micro-Data, Aggregate Time-series data, and other simulation model results. The first includes other survey and administrative micro-data sources (eg. Census, T4 administrative file). The System of National Accounts is an example of time-series data while the Department of Finances Tax/Transfer model is another microsimulation model.

In any validation exercise using any of the reference classes, one must be sure that the conceptual differences are accounted for (e.g. the employment income of the comparison data is conceptually equivalent to the employment income concept on the SPSD/M). In addition, any ways in which the population underlying the comparison data source differs from the population underlying the SPSD/M must be accounted for. The next section shows some validation tables for the SPSD/M 19.0 while the final sections give examples of conceptual and coverage problems by discussing the three reference classes.

VERSION 19.0 BASE YEAR (2007) VALIDATION TABLE

Overview

The base year (2007) of Version 19.0 has been validated against a variety of other sources. The results are presented here.

The first table shows how both the Survey of Labour and Income Dynamics and the SPSD/M compare to National Accounts Data.   In general, the SPSM estimates are closer to the National Account estimates due to adjustments made in the database creation process.

The following tables show how the SPSD/M compares to other published aggregate sources. Estimates come from the following sources: the Survey of Labour and Income Dynamics, the System of National Accounts, Human Resources and Social Development Canada, and Taxation Statistics. When more than one estimate was available, all estimates are shown. The first two columns of these tables show the percentage and dollar difference between the SPSD/M and the most comparable data available (these are underlined in the table). There are conceptual differences between the sources.

Market Income

This table shows the different estimates for market income. The difference between the SPSD/M and the estimate that most closely matches the SPSD/M concept is shown in the first two columns. The estimates that are compared are underlined in the table.

Comparison of Market Income, 2007 (millions of dollars)

				A	B	E	F	G
		Difference			**Source of Estimate**			
		%	*$*	SLID	SPSM	T4	Greenbook	T1
1	Wages and Salaries	100.2	1,357	660,776	684,184	682,827	654,235	
2	Self-employment Income	104.0	2,194	55,967	57,681		50,141	55,487
4	Total Earnings	105.3	37,489	716,743	741,865		704,376	
5	Investment Income	97.0	-3,010	69,226	96,736			99,746

Notes:

Columns

A	The SLID figures are from the 2007 SLID.
B	The SPSM figures are for the SPSD/M version 19.0 Release.
E	Special tabulation of T4 file.
F	Income Statistics publication, Canada Revenue Agency.
G	Special tabulation of the T1 file.

Rows
5. SPSD/M investment income is higher than the SLID due to the imputations of interest income.

Transfer Income

This table shows the different estimates for transfer income. The difference between the SPSD/M and the estimate which most closely matches the SPSD/M concept is shown in the first two columns. The estimates that are compared are underlined in the table.

Comparison of Transfer Income, 2007 (millions of dollars)

				A	B	C	D	E	F	G

		Difference		Source of Estimate						
		%	*$*	SLID	SPSM	SNA	HRSDC	Admin	Greenbook	T5007 or T4 or T1
	Transfer Income									
6	Government Transfers to Persons			105,738	116,119					
7	Worker's Compensation	101.1	59	5,382	5,531	<u>5,472</u>			4,680	
8	OAS, GIS and SPA	101.3	420	29,259	32,305	<u>31,885</u>	32,570			
9b	Social Assistance	104.2	398	8,871	9,950	11,383			8,275	<u>9,552</u>
	Federal Transfers									
11a	Child Tax Benefit	85.4	-1,363	7,848	7,989	9,449	<u>9,352</u>			
11b	UCCB	106.7	153	2,424	2,425				<u>2,273</u>	
12	OAS	99.6	-111		24,601		<u>24,711</u>		23,574	
13	GIS	97.9	-155		7,191		<u>7,346</u>			
14	SPA	100.2	1		514		<u>513</u>			
15	EI Benefits+QPIP	98.8	-171	14,523	13,897	12,491				<u>14,067</u>
15b	QPIP	91,9	-126		1,293					<u>1,419</u>
16	CPP/QPP Income	98.3	-585	32,484	34,647	<u>35,232</u>	35,617		34,118	
17	Sales Tax Credit	110.8	335	3,233	3,425					<u>3,090</u>
	Provincial Transfers									
18	Family Programs			423	393					
19	Elderly Programs				433					

Notes:
Columns:

A   The SLID figures are from the 2007 SLID.

B   The SPSM figures are for the SPSD/M version 19.0 Release.

D   Social Assistance, Child Tax Benefit and Provincial Transfers come from the Strategic Policy and Research Directorate.   The remaining figures are from Statistics Related to Income Security Programs in Canada.

E   Official Quebec Statistics, government of Quebec

F   Income Statistics publication, Canada Revenue Agency.

G   Special tabulation of T5007, T4 files and T1 files.

Rows:

9b   SNA column is direct relief at the provincial and local level.

Income Taxes

This table shows the different estimates for income taxes. The difference between the SPSD/M and the estimate which most closely matches the SPSD/M concept is shown in the first two columns. The estimates which are compared are underlined in the table.

Comparison of Income taxes, 2007 (millions of dollars)

				B	C	F	G

		Difference		Source of Estimate			
		%	*$*	SPSM	SNA	Greenbook	T1
	Federal						
21	Total Income assessed	102.4	22,626	956,774			934,148
22	Total Deductions	106.0	5,198	91,933			86,735
23	Taxable Income	103.1	25,945	864,914			838,969
24	Basic Federal Tax	103.9	4,282	112,987		108,705	
26	Federal Income Tax	102.5	2,705	112,078	116,430	107,697	109,373
27	Number of Filers (000)	97.4	-626	23,811		24,437	
28	Number of Taxable Filers (000)	106.8	1,132	17,677		16,545	
	Provincial						
29	Québec provincial income taxes	102.7	560	21,424	22,550	20,864	
30	Provincial income tax, excl. Québec	99.1	-408	46,906	51,095	47,313	

Notes:
Columns:

B   The SPSM figures are for the SPSD/M version 19.0 Release.
C   Provincial Economic Accounts.
F   Income Statistics publication, Canada Revenue Agency.
G   Special tabulation of the T1 file.

Rows:

21   Both SPSM and Greenbook are restricted to Taxable Filers Only. The Greenbook figure is net of Tax Exempt Income (Social Assistance, Workers Compensation and GIS/SPA.)
22   Both SPSM and Greenbook are restricted to Taxable Filers Only. The Greenbook figure is net of Tax Exempt Income (Social Assistance, Workers Compensation and GIS/SPA.)
23   Taxable Filers Only.
24   GB has non-provincial netted out (all returns).
26   GB does not include Québec non-refundable tax abatement, matches SPSM (imtxf net of imqtaa).
29   Québec figures from the Portrait de la fiscalité des particuliers au Québec
30   GB does not include Québec taxes.

Payroll Taxes

This table shows the different estimates for payroll taxes. The difference between the SPSD/M and the estimate which most closely matches the SPSD/M concept is shown in the first two columns. The estimates which are compared are underlined in the table.

Comparison of Payroll Taxes, 2007 (millions of dollars)

				B	C	F	J
		Difference		Source of Estimate			
		%	*$*	SPSM	SNA	Greenbook	T1
31	EI Contributions	109.9	691	7,654	7,049	6,623	6,963
32	CPP/QPP Contributions	106.4	1,381	22,812	21,431	19,873	

Notes:

Columns:

B   The SPSM figures are for the SPSD/M version 19.0 Release.

F   Income Statistics publication, Canada Revenue Agency.

Commodity Taxes

This table shows the different estimates for commodity taxes. The difference between the SPSD/M and the estimate which most closely matches the SPSD/M concept is shown in the first two columns. The estimates which are compared are underlined in the table.

Comparison of Commodity Taxes, 2007 (millions of dollars)

				B	C	E
		Difference		**Source of Estimate**		
		%	*$*	SPSM	SNA	COMTAX
	Federal					
33	Goods and Services Tax	100.0	-2		34,449	34,451
34	Households Only	100.0	-1	26,442		26,443
35	Other Federal Commodity Taxes	99.7	-37		13,762	13,799
36	Households Only	99.9	6	10,000		10,005
	Provincial					
37	Sales Taxes	100.2	95		38,015	37,920
38	Households Only	100.0	-8	25,702		25,709
39	Other Provincial Commodity Taxes	100.1	19		17,770	17,751
40	Households Only	100.0	-7	14,808		14,815

Notes:

Columns:

B   The SPSM figures are for the SPSD/M version 19.0

C   Provincial Economic Accounts.

E   Commodity Tax Model

COMPARISONS WITH OTHER MICRO-DATA SOURCES

The main advantage of comparing the SPSD/M against other micro-data sources lies in the ability to examine various demographic and economic distributions in great detail. For example, the Canada Revenue Agency's "Greenbook" file, which is a stratified sample of T1 returns, could be used to compare the number of individuals receiving the working income tax benefit by income class.

 However, the Greenbook, as well as most other micro-data sources, suffer from both a lack of timeliness and conceptual differences in the population on which it provides estimates. The lack of timeliness means that validation exercises can only be performed for historical years while a great deal of policy analysis focuses on the future. If, for instance, one is constructing a new benefit system for the year 2012, most people would prefer to simulate in 2012 (complete with heroic growth assumptions if necessary) and not use, say, 2009 as a proxy. A further problem for non-database years is that the Greenbook provides distributional information for a particular population which is conceptually different from either the SPSM base year or projected year.

The following example is provided to demonstrate how the conceptual differences in coverage can affect the estimates. By comparing the Wages & Salaries estimates, by province, in 2007 in the

Greenbook with T4 estimates some substantial differences arise. These are outlined in the table below.

Table 1 - Comparison of Wages and Salaries, by Province, 2007 (millions of dollars)

Province	T4	Greenbook	Difference	Percentage Difference
Nfld.	7,410	7,834	425	6%
P.E.I	2,032	2,013	-19	-1%
N.S.	15,223	14,818	-405	-3%
N.B.	12,008	,888	-120	-1%
Que.	135,316	134,415	-901	-1%
Ont.	279,201	262,526	-16,675	-6%
Man.	20,822	20,138	-683	-3%
Sask.	17,710	17,760	50	0%
Alta.	106,528	99,478	-7,050	-7%
B.C.	86,578	83,364	-3,213	-4%
Canada	682,827	654,235	-28,592	-4%

It is clear from this table that the Greenbook does not provide good wage and salary estimates for most provinces. For Canada as a whole, the Greenbook underestimates the T4 figures by almost $29 billion or 4%. On a provincial basis, the Greenbook estimates for Ontario, Alberta, and Newfoundland are off by more than 5% while the estimates for Saskatchewan are quite accurate. Since there exists no conceptual differences in the definition of wages and salaries in each data source one must conclude that the differences in the estimates stem from the coverage of the Greenbook sample. Therefore, to compare the wages and salaries figures from the Greenbook file to those in the SPSD/M would not be appropriate.

There exist other features of the Greenbook coverage which make comparisons with the SPSD/M difficult. For example, the Greenbook file includes the tax returns of the deceased provided that the estate filed a return. As the SPSD/M contains no deceased it becomes difficult to compare certain estate-sensitive income components such as the income from capital gains.

In conclusion, all micro-data sources suffer from the same kind of limitations that have been outlined above for the Greenbook. These limitations show the need for caution in interpreting results from any validation exercise.

COMPARISONS WITH AGGREGATE TIME SERIES STATISTICS

An alternative approach to validation is to compare the SPSD/M results with aggregate time series statistics such as those published by Statistics Canada in the System of National Accounts. Although these estimates are much more timely than micro-data sources, there are often serious conceptual and coverage differences in the series. Consider the example described in Table 2 that compares provincial income tax revenues from the System of National Accounts with the Greenbook file.

Table 2 - Comparison of Provincial Income Tax, by Province, 2007

Province	SNA ($ millions)	Greenbook ($ millions)	Difference ($ millions)	Percentage Difference
Nfld.	896	882	-14	-2%
P.E.I	218	215	-3	-1%
N.S.	1,736	1,684	-52	-3%
N.B.	1,258	1,205	-53	-4%
Que.	22,550	20,864	-1,686	-7%
Ont.	28,444	25,735	-2,709	-10%
Man.	2,334	2,171	-163	-7%
Sask.	1,812	1,786	-26	-1%
Alta.	7,667	7,336	-331	-4%
B.C.	6,730	6,299	-431	-6%

The table above shows major differences in provincial income tax estimated from these two data sources. For Ontario alone, the Greenbook estimate is almost $2.7 billion, or 10% lower than the SNA estimate. In addition, the estimates for most provinces would appear to be poor.

There are, however, conceptual differences in the treatment of refundable tax credits between the two data sources. The SNA income taxes figures come from the government revenue table and are gross of most refundable tax credits. These credits are netted out in the calculation of the provincial balance through a compensating entry on the expenditure side through transfer payments to persons. The Greenbook has already eliminated the provincial refundable tax credits from its estimate.

There exist two profound conceptual differences between the SNA estimates and the conceptual basis of the SPSD/M. First, the SPSD/M models household cash incomes and cash transactions between households and governments while the SNA estimates only forms of income accrued from production or transfers payments from businesses and governments. Therefore, total personal sector income includes many imputed (non-cash) items such as the employer contributions to pension plans, rental income accruing from owner-occupied dwellings, and imputed interest stemming from the System of National Accounts' treatment of the financial industry. It would not include income from capital gains, income from trusteed pension plans, gross income from insurance claims, or alimony income.

Secondly, the personal sector, as defined by the SNA, does not correspond with the household universe underlying the SPSD/M. Charitable, churches and other religious organizations, universities, and trusteed pension funds are all components of the SNA personal sector which are not found in the SPSD/M population.

In conclusion, the above discussion makes clear many of the pitfalls of verifying the SPSD/M results with aggregate time series data.

COMPARISONS WITH OTHER MICRO-SIMULATION MODELS

Unfortunately, comparisons with other micro-simulation models of the same nature of the SPSD/M are time-consuming tasks and rarely performed. The high number of assumptions to be reconciled and their interrelated impacts have prevented an exhaustive comparison of two models. The models are under constant development and would change substantially over the time frame of a validation exercise. Moreover, there are only a handful micro-simulation models of this kind with which comparisons could be performed. Ultimately a reconciliation of two models would be very useful

especially in future years.

Adjusting SPSD/M Results to Control Totals

There are a number of ways in which users can adjust SPSD/M results to control totals coming from other sources.

POST HOC ADJUSTMENT

In this method, SPSD/M is run using off-the-shelf parameter files, either using files pertaining to the base year or to a projected year. Tables are produced (using UVAR and XTSPEC) and the results turned into proportions. These proportions are then applied to corresponding control totals to give distributional results that are consistent with the imposed control totals.

Consider a scenario that gives the number of gainers and losers by income group. If the total number of families in each income group differs between the SPSM run and a desired benchmark source, then the gainers and losers can be scaled by the proportion of total SPSM counts to total benchmark counts. This is an operation that can be performed most simply using a spreadsheet. Please see the section on the import facility in the *Tools User's Guide* for more information on facilitating the import of data from SPSM output tables into a spreadsheet.

This method can be adapted even if exactly corresponding totals between SPSM and the benchmark data source are not available, by using corresponding indicator variables. For example, if for a particular scenario SPSM gives a certain result for provincial tax, but the provincial employment income is 10% lower compared to a benchmark source, the SPSM provincial tax result can be scaled up by 10%, since *ceteris paribus* taxes follow employment income fairly closely.

CREATING A NEW WEIGHT FILE

A new weight file can be created if the user wishes to adjust the distribution of the population in SPSD to reflect a structural change. For example, if the Social Assistance population has increased by 20% compared to SPSD/M default results, the user can create a new weight file where the desired population has been increased, and the remainder of the population decreased to maintain a constant total population. This method involves three steps:

1. Perform an SPSD/M run that gives the number of households classified by the variable of interest. In our simple example, we can create a classificatory variable that partitions the population into households that receive Social Assistance and those that do not. The following control parameter settings will produce a table similar to the one illustrated below.

```
UVAR
      label(sastat) = "Household SA status";
      levels(sastat) = "No SA", "Received SA";
      if (HH:imisa > 0) sastat = 1;
      else sastat = 0;

XTSPEC
HH: sastat+ * {units};

Table 1U: Unit Count (000) by Household SA status (2007)
+-----------+----------+
```

```
|Household  |Unit Count|
|SA status  |  (000)   |
+-----------+----------+
|No SA      |   12235.8|
|Received SA|    1253.8|
+-----------+----------+
|Both       |   13489.6|
+-----------+----------+
```

In our example, we wish to scale up the Social Assistance population by 20%. If we wish to maintain the total number of households constant, we must scale down the non-SA population by a factor of 0.888. This factor is derived in the formula below using numbers from the above table.

$1 - ((1253.8 * 1.20) / 13489.6) = 0.888$

2. Next we must produce an output file that scales the weight of each household by the desired factor and outputs the results to an ASCII output file. The following control parameter settings will accomplish this.

```
UVAR
        label(sastat) = "Household SA status";
        levels(sastat) = "No SA", "Received SA";
        label(newwgt) = "Adjusted household weight";
        if (HH:imisa > 0) sastat = 1;
        else sastat = 0;
        if (idhhrh == 0) {
        if (sastat == 1) newwgt = 1.200 * hdwgthh;
        else             newwgt = 0.888 * hdwgthh;
        }

ASCFLAG     1
OUTASC      saadj.prn
ASCUNIT     4
ASCSTYLE    4
ASCDELIM
ASCEXTPRC   3
ASCVARS     newwgt

XTFLAG      1
XTSPEC
HH: sastat+ * {units};
```

It is necessary to ensure that newwgt is set for only one member of the household, since the ASCII output facility automatically aggregates analysis variables to the level (in this case household) given by ASCUNIT. The setting of ASCEXTPRC ensures that the newwgt will retain its fractional information rather than being rounded to an integer.

3. The next step involves converting the ASCII file into an SPSM weight file. The utility bldwgt.exe will perform this operation for us if we issue the following DOS command:

```
c:\> bldwgt 0 saadj.prn saadj.wgt
```

The bldwgt.exe utility is described in more detail in the *Tools User's Guide*.

The weight file saadj.wgt can now be used in SPSD/M runs using the INPWGT control parameter.

ADJUSTING GROWTH FACTORS

The SPSD is constructed by combining data from a number of sources for a base year. In order to

apply this information to subsequent years, dollar-denominated quantities (such as employment income) generally have to be increased to reflect per capita increases due to inflation and productivity changes. Through the database adjustment parameter files (.apr files) users can change these growth factors to force quantities to hit desired benchmark totals. Every dollar-denominated database variable in SPSD/M can be manipulated in this way, and the growth factors can be varied by province.

Note that this technique increases mean amounts per individual or per family. If the changes responsible for the discrepancy between SPSD/M results and the target source result from distributional differences, this approach may lead to false inferences. For example, if total Social Assistance payments need to be adjusted by 20% because the SA population has increased by 20%, this technique would instead increase mean SA benefits by 20% and keep the number of recipients unchanged.

RECORD CONVERSION

It is possible to change SPSD data as it is being read in. This technique can be used to selectively modify records that meet certain criteria. For example, a desired proportion of records of low-income families not receiving Social Assistance could be 'converted' into SA recipients by setting a non-zero value to the idisa variable. There are two very different ways of doing this. One way is to use the SPSM 'glass box' mode, which allows the user to add or modify the C language routines that read in data. A special routine, named adju.c, has been expressly provided for this purpose. Please see the *Programmer's Guide* for more information on this approach.

Another approach is to use SPSM reference variable facility. This facility lets the user replace database variables by variables contained in an SPSM results (.mrs). This is a two step procedure. First, SPSM is run to produce a results file containing a modified value of a variable for each person on the database. The following example set of control parameters produce a new variable named newsa that 'converts' 20% of non-SA families with family disposable income less than $15,000 by giving them $5,000 in SA income. For all other families, newsa retains its previous value. Note that the $5,000 benefit is given to the head of the Census Family. This ensures that the benefit is given only once to each qualifying Census Family. The newsa variable is written to a conventionally created SPSM results file named newsa.mrs. Please note that this examples assumes that the parameter SEED in the base year parameter file (e.g. c:/spsd/ba00.cpr) had three rows and than a fourth random number was to be added. The SEED parameter should usually be incremented since the SPSM model will use the other random number streams when modeling the tax and transfer system.

```
SEED         4
      0
      1
      2
      3
UVARFLAG     1
UVAR
      if (CF:idisa==0 && idrand3<0.20 && idcfrh==0 &&            CF:immdisp<15000)
      newsa = 5000;
      else newsa = idisa;
OUTMRSFLAG   1
OUTVARMRS    newsa.mrs
OUTMRSVARS   newsa
```

The next step involves using the SPSM reference variable facility to replace the database variable idisa with the newsa variable from the previously produce results file newsa.mrs. The following control

parameters accomplish this:

```
REFFLAG    1
INPREF     newsa.mrs
REFVARS    idisa=newsa
```

Any SPSM run that includes these control parameters will use the 'converted' social assistance that was written to the newsa.mrs results file. The REFVARS control parameter tells SPSM to obtain the newsa variable from the file named by the INPREF parameter and rename it to idisa before performing any tax/transfer calculations. Please see the *User's Guide* for more information on the SPSM reference value facility.

RAKING WEIGHTS

All of the weight files included as part of SPSD/M (with the exception of the base year) have been produced using a generalized version of a technique known variously as 'raking' or 'iterative proportional adjustment'. This technique modifies the weight of each household in the database so as to match desired arbitrary control totals. It is a computationally-intensive technique, and uses additional software programs constructed by the SPSD/M development team that require some time to master. If you wish to explore this option further, please contact an SPSD/M team member.