Catalogue No. 94-13

CALCULATION OF FAMILY INCOME FOR SLID

Product Registration Number 75F0002M

May 1994

Philip Giles, Household Surveys Division

Chantal Grondin, Social Survey Methods Division

Sylvie Michaud, Social Survey Methods Division

The SLID Research Paper Series is intended to document detailed studies and important decisions for the Survey of Labour and Income Dynamics. These research papers are available in English and French. To obtain a summary description of available documents or to obtain a copy of any, please contact Philip Giles, Manager, SLID Research Paper Series, by mail at 11-D8 Jean Talon Building, Statistics Canada, Ottawa, Ontario, CANADA K1A 0T6, by INTERNET (GILES@STATCAN.CA), by telephone (613) 951-2891, or by fax (613) 951-3253.

EXECUTIVE SUMMARY

As an individual's economic well-being is closely linked with his/her family circumstances, many researchers are interested in "family" data from SLID. However, in general, the family is not a stable unit through time, and thus raises new issues for a longitudinal survey. As it is arguably the most important family variable in SLID, family income is used as an example of the approach to be adopted for other family variables as well.

Three options for the calculation of family income are presented:

- The Snapshot Approach is the "traditional" method of calculating family income for a particular year based on the family composition on January 1 of the following year;
- The Prorated Approach considers the time during the year in which various people lived together;
- The Subannual Approach extends the Prorated Approach by also considering when the various income sources for each person were earned.

It is essential to note that the last two approaches lead to a calculation of family income which is different for every individual. This is the consequence of introducing time into the definition as families can change over time.

Introduction

The collection of longitudinal survey data raises several new issues for survey staff to address. One of these, the calculation of measures at the family or household level, is discussed in this document. The purpose of this document is to motivate discussion on the approach which should be taken by the Survey of Labour and Income Dynamics (SLID). Although "family income" is used throughout the document, most of the principles are readily extended to other measures at the family and household level. In this sense, "family income" is an example of SLID's approach to family and household level data.

The New Issue

Traditional cross-sectional surveys have derived various measures such as family income, as these are useful analytical variables and, in many cases, interesting data items in their own right. The family income, usually an annual measure, is calculated as the sum of the annual income of each individual in the family. The family is defined by the persons identified as related family members when the survey data are collected. In most (if not all) surveys, no questions are asked to determine whether the current family composition was, in fact, the same composition throughout the entire reference year. Implicitly, this was assumed by both those responsible for producing survey output products and data users. Arguments supporting this approach are:

- to add questions to allow adjustment of family income for family composition changes would reduce survey content elsewhere;
- at an aggregate level, the differences would be negligible due to "netting out", even if a large proportion of families experienced a compositional change during the year.

One of the objectives of SLID is to provide information on changes in family composition. Therefore, some information on changes is available as part of basic survey content. Also, for analysis of longitudinal survey data, microdata take on increased importance in comparison with traditional cross-sectional surveys. Therefore, the above arguments no longer hold. SLID must decide how to calculate family income (and other measures at the family and household level) for inclusion in its output data files. The remainder of the document examines various possibilities, and discusses advantages and disadvantages.

Snapshot Approach

The first approach is essentially the traditional approach. For each person, a value for family income is derived for every year in the survey.

> Calculate family income for a particular year based on the family composition on January 1 of the following year. All members of the family at that time will have the same value for family income -- the sum of every individual's income during the reference year.

This approach does not consider any changes in family composition during the reference year. The same problems would exist -- the family income of certain individuals will be "misleading". This is illustrated in the example given below. However, it is straight-forward to calculate, and provides data users with a measure they are used to. The calculation of family income equivalents (division of family income by family size, as used in poverty analysis) is simple. The same argument applies to the derivation of income-to-needs ratios.

Because researchers are accustomed to this approach and various income studies assume it, SLID will provide family income calculated according to this method. The issue of interest here is whether other measures should also be provided.

Prorated Approach

As compared with the Snapshot Approach, this approach takes into account one additional piece of information -- dates of changes in family composition.

> The family income is calculated separately for each individual, considering the time spent living with other individuals. In situations where there is no change in family composition, the result of this calculation will be the same as in the Snapshot Approach.

It is important to realize that family income is now calculated at the individual level. A person living four months in one family and the remaining eight months in another family will have one-third of his annual family income derived from the first family and two-thirds from the second family. The advantage to this approach is that it will provide a "more accurate" measure of family income for individuals whose economic situation changes greatly from one family to another. If the person's family income is viewed over the full six years of the panel, an income change associated with a family transition will appear more gradually; the Snapshot Approach would, in contrast, show a more dramatic change between two years. It does make a basic assumption -- that all income is earned uniformly throughout the year -- which will not be true for some individuals. It is a different concept for data users to understand, which may or may not pose problems. (It is a fact that longitudinal data are inherently different from cross-sectional data, and different analytical tools are needed. Dealing with different concepts is just part of the shift in thinking.)

An Example

The following example illustrates the differences between the Snapshot and Prorated approaches.

On January 1 of the reference year, there is a three person family: Dad, Mom, and Junior. On October 1, Junior moves out to live alone, resulting in two families. There are no other changes to either family. During the reference year, the total income of each of them is 50K, 22K and 12K, respectively.

Under the Snapshot Approach, the family income of each of the three individuals is:

Dad:	72K [50 + 22]
Mom:	72K [50 + 22]
Junior: 12K	[since he was living alone on January 1 following the reference year]

Under the Prorated Approach, the family income of each of the three individuals is:

Dad:	81K [50 + 22 + (9/12) * 12]
Mom:	81K [50 + 22 + (9/12) * 12]
Junior: 66K	[12 + (9/12) * (50 + 22)]

It could be argued that Junior's family income for the year is better expressed by the Prorated Approach than by the Snapshot Approach. Over a three-year period, covering the year of the transition and the years on either side, the Prorated Approach will show a more gradual change in Junior's family income than that with the Snapshot Approach (i.e., the large drop will be noted in the year after he

moves out).[1] The latter would identify a large change in family income in the same year as Junior moves out.

Example 2: Complicating the situation

Now assume that, instead of moving out to live alone, Junior moved in with his girlfriend Juniper, who had lived alone between January and September. Juniper's total income for the reference year was 36K.

Under the Snapshot Approach, the family income of each of the four individuals is:

Dad:	72K	[50 + 22]
Mom:	72K	[50 + 22]
Junior: 48K		[12 + 36]
Juniper:	48K	[12 + 36]

Under the Prorated Approach, the family income of each of the four individuals is:

Dad:	81K	[50 + 22 + (9/12) * 12]
Mom:	81K	[50 + 22 + (9/12) * 12]
Junior: 75K		[12 + (9/12) * (50 + 22) + (3/12) * 36]
Juniper:	39K	[36 + (3/12) * 12]

This shows that the Prorated Approach can be used for all types of changes in family composition. To determine the value of the Prorated Approach, a decision must be made whether a better measure of Junior's family income is 48K or 75K.

Subannual Approach

---

[1] Assuming that Junior lived with his parents the entire year prior to the reference year and alone the entire year following the reference year, and also that each individual's personal income does not vary greatly over this three-year period.

Assuming one believes that the Prorated Approach is intuitively better than the Snapshot Approach, one possible cause for errors in the Prorated Approach is income sources which are not received uniformly throughout the year. A large percentage of the population earn a high percentage of their annual income from employment and government transfers. Among those in this group, a high percentage have one job (at any time) throughout the year and are full-time full-year workers. Government transfers are mostly paid monthly or quarterly. For this group, the assumption of uniform income receipt is likely a reasonable one. However, there are also many individuals where the assumption will not be reasonable, and it is likely that many of these will be the focus of analysis. Therefore, the Subannual Approach is proposed to deal with this drawback.

> For every individual, a monthly income will be derived (some details given below). The annual income for an individual will be calculated as in the Prorated Approach with one exception. The monthly incomes of all other family members are summed for those months when they are in the same family.

The subannual approach does complicate the calculation of family income equivalents. User consultation will be required to determine the acceptability of deriving this based on the number of months of income contributed by each person to the calculation of family income. (The example below shows this.)

Example of the Subannual Approach

More information will be provided for example 2 to illustrate the difference. Assume that Junior earned all his income between September and December (equal amounts each month). Dad, Mom and Juniper earned their income uniformly throughout the year. The family incomes calculated using the Snapshot and Prorated Approaches will not change.

Under the Subannual Approach, the family income of each of the four individuals is:

Dad:	75K [50 + 22 + (1/4) * 12]
Mom:	75K [50 + 22 + (1/4) * 12]
Junior: 75K	[12 + (9/12) * (50 + 22) + (3/12) * 36]
Juniper:	45K [36 + (3/4) * 12]

To calculate family income equivalents for the Snapshot Approach, divide the estimate of family income by the family size: for Dad and Mom this would yield a value of 36K, and Junior and Juniper would each be assigned a value of 24K. With the Prorated Approach, one must calculate monthly family income, divide each of them by the number of family members that month, and sum over the twelve months. Under Example 2 with the Prorated Approach, the following values are derived: Dad and Mom (30K) , Junior (27K), Juniper (33K). Under the subannual approach where Junior's income was earned uniformly from September to December, the values change to: Dad and Mom (28K), Junior (28K), Juniper (36K).[2]

---

[2] To explain the calculation of the family income equivalents, consider Junior using the subannual approach. Junior lived with his parents from January through August where the family income was \$6K per month (4.17 from Dad, 1.83 from Mom, 0 from Junior). In September, Junior started earning income (\$3K per month). In September, his family income was \$9K (4.17 + 1.83 + 3). From October through December, his monthly family income was \$6K (3 from

Example 3: What happens to the calculation of family income when a marriage ends?

This example features a husband-wife family with two young children (no income). The family lived together throughout Year 1. Dad had an income of 60K (5K per month), all from his full-time full-year job. Mom had no income. On May 1 in Year 2, Dad and Mom separate with the children moving with Mom. Dad agrees to pay $1250 per month in child support ($15K annually). Mom obtains a job paying $30K annually ($2500 per month) starting on May 1. There are no further changes in household composition or in income through to the end of Year 3. The following tables show the calculations of family income for Dad and Mom, plus the family income equivalents for each of the three proposed approaches.

SNAPSHOT APPROACH						
	Family Income ('000)			Income Equivalents ('000)		
	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3
Dad	60	60	60	15	60	60
Mom	60	30	45	15	10	15

himself and 3 from Juniper). Therefore, the family income equivalent for Junior is calculated as: $6K / 3 persons * 8 months + $9K / 3 persons * 1 month + $6K / 2 persons * 3 months = 16 + 3 + 9 = $28K.

PRORATED APPROACH						
	Family Income ('000)			Income Equivalents ('000)		
	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3
Dad	60	70	60	15	47.5	60
Mom	60	50	45	15	14.2	15

SUBANNUAL APPROACH						
	Family Income ('000)			Income Equivalents ('000)		
	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3
Dad	60	60	60	15	45	60
Mom	60	50	45	15	15	15

One interesting observation from these tables is that in Year 2, the Prorated Approach gives Dad a portion of Mom's income since they were in the same family for part of the year, even though all her income was earned while they were in different families. Part of this income is child support payments received from Dad, so is counted twice in Dad's family income for Year 2. His income looks to be higher in the year of separation compared to the year before and the year after.

Calculation of Subannual Income

SLID has never expressed the calculation of subannual income to be one of its objectives. However, using some assumptions, the current survey content will allow the calculation of subannual income. The quality of some of these estimates will be better than others. If the Subannual Approach is used for determining a variable on the microdata file, the calculation of monthly income estimates is necessary for the purpose of calculating annual family income. However, monthly income might not be available on any data file unless the data quality is sufficiently high.

SLID collects information on a series of income sources and derives total income as the sum of the income received from all sources. One possible categorization (although there are several definitional issues which arise) is: employment income, government income, pension income, investment income and other income. More work will take place to examine the various issues and establish clear definitions. However, a preliminary look indicates that the problems may not be insurmountable:

- Employment income: SLID will be collecting detailed information on dates of employment, wages and salaries earned from each employer, dates of change in wages and salaries. Using these dates, SLID can accurately calculate monthly employment income. The one problematic area is self-employment income for owners of unincorporated businesses.

- Government income: Some income sources from government (such as tax credits) can be considered to be earned uniformly throughout the year, even when recipients do not receive monthly cheques. They are clearly calculated on an annual basis and divided according to the characteristics of

the particular programs. Recipients of some government transfers such as CPP and OAS receive monthly cheques. Others are clearly subannual, pertaining to a particular period of time. Examples include UI benefits, Social Assistance, Workmens' Compensation. For some of these programs, SLID collects information on the months in which each person received benefits. Therefore, it would be relatively straightforward to calculate monthly government income.

- Pension income: Most (perhaps all) pensions are paid monthly. It appears that the derivation of monthly pension income should not be problematic.

- Investment income: The calculation of monthly investment income is the most difficult, both from a conceptual as well as an implementation point of view. It is likely that the assumption of uniform receipt will be reasonable (more or less) for investment income such as bank accounts and savings certificates such as Canada Savings Bonds and GICs. When one gets into such items as dividends and capital gains and losses, the reasonableness of this assumption may be more questionable. However, if SLID calculates monthly investment income, there does not appear to be any alternative at this time to using the assumption of uniform income receipt.

- Other income: These income sources are also problematic since this is a very heterogeneous group. A high proportion of people will not report income sources in this group. However, for those who do, the amounts may be quite significant. As with investment income, the calculation of monthly income would require the use of the assumption of uniform income receipt.

Cohabitants

SLID "following rules" – who is traced and for whom data are collected – are an important consideration for the calculation of family income. Details of these following rules are provided in SLID Research Paper 94-01 *SLID Following Rules: Who to Trace and Who to Interview.* Basically, these rules state that all persons in the sample when a panel is introduced are followed for the life of the panel (six years). These persons are called longitudinal respondents. In addition, all persons living with a longitudinal respondent (called cohabitants) are interviewed for a given reference year if they lived with a longitudinal respondent for all or part of that reference year. Thus, a cohabitant is not included in the calculation of family income using the Snapshot Approach for the year in which he/she moves out, but is included in the calculation for the Prorated and Subannual approaches. (This observation holds for all persons according to the definitions. Cohabitants are specifically mentioned here since they will generally only be useful to analysts for their contribution to the family and household characteristics of the longitudinal respondents.)

Effect of Non-response

The discussion in this document so far has focused on the survey data content, examining the possibilities for the calculation of family income based on the various items of information requested of every respondent. When calculating derived variables from survey data, one must always consider what to do with respondents for whom incomplete information was collected, due to non-response, interviewer error, processing errors, etc. Particular problems may also arise when attempting to trace cohabitants who move away from a longitudinal respondent.

Annual income will be calculated for all persons. In some cases the value will be imputed, but every person will have a valid value recorded on the output file. Therefore, it will always be possible to calculate family income using the Snapshot

Approach. Calculation using the Prorated Approach will not pose many problems due to non-response. Dates of changes of family composition will likely be fairly accurate and fairly complete. The Subannual Approach requires many more input data items, and will pose many problems for the calculation of family income.

The following steps will be taken for dealing with non-response:

- Dates for changes in household composition (both for additions and reductions in household members) must be present and consistent. This means that missing dates must be imputed and that inconsistencies must be resolved. Inconsistencies may exist since the same question is asked of more than one person. Referring to example 2, when contacting the household with Dad and Mom, the interviewer will ask "When did Junior leave?". When contacting the household containing Junior and Juniper, the interviewer will ask (about Junior) "When did you move to this address?". The answers should generally be the same, but may not be since different people are responding. Differences may be due to recall errors or to different perceptions as to when the move actually took place (moves can be complex, involving overlapping ownership/tenancy or short-term interim arrangements).

- As noted above, total income must be present for individuals. Imputation will be required if income information is not collected. For these persons, monthly income will be calculated as if it were earned uniformly throughout the year.

Using Data on Dates

Theoretically, SLID will be using the first day of each month for fixing family composition. In particular, January 1 will be the date used for the calculation of family income using the Snapshot Approach for the previous calendar year. Although SLID will be collecting date information on moves to the exact day, the day will be used only as a guide. Thus, anyone moving in January between the 1st and 10th will be deemed to have moved on January 1. Anyone moving between the 11th and 31st will be deemed to have moved on February 1. (This will also be the case with other months.)

Several reasons led to this approach. In a majority of cases, it is expected that the day will not be known or will be guessed. Even those knowing the exact date may round to the 1st, 15th, or 31st of the month. It is likely that most moves take place during the first and last weeks of the month; by not fixing an exact day allows for such factors as the shifting dates of the first and last weekend of each month and dates of holidays near the beginning and end of months (January 1 being a notable example).

Some initial results

Although the examples are useful as illustrations of the various situations and options, it is important to examine the effects on survey results. Some preliminary studies were done using income data collected in the SLID May 1993 field test. For more information on this field test, see SLID Research Paper 93-04 *SLID Income Interview - May 1993: Questionnaire and Data Collection Procedures*.

Total incomes were calculated for all persons responding to the income test. Family incomes, using both the Snapshot Approach and the Prorated Approach, were calculated for all persons in families for whom all persons responded to the test (i.e., for which every person's total income was available). As more work is

needed to define the calculation of family income using the Subannual Approach, it was not possible to include it in this study. (This work will be completed prior to future evaluation studies.)

Over 90% of the economic families in the test had no change in composition during 1992. The following table shows the number of economic families in January 1993 by the types of changes during 1992. The movers are restricted only to those persons 15 years of age and older. Therefore, families in which a baby was born are recorded as "No change". This table also includes families for which family composition was reported, but for which, at least one person's total income was not reported.

Economic Family Changes during 1992	Families	Total Persons	Longitudinal Persons	Cohabitants
No change	1113 (91.2)	2382 (92.6)	2382 (95.5)	0 (0.0)
Movers in only	54 (4.4)	114 (4.4)	53 (2.1)	61 (77.2)
Movers out only	40 (3.3)	42 (1.6)	42 (1.7)	0 (0.0)
Both movers in and movers out	13 (1.1)	34 (1.3)	16 (0.6)	18 (22.8)
TOTAL	1220	2572	2493	79

Some univariate results are given in the following table:

While the numbers are similar, they are slightly higher for family incomes calculated using the Prorated Approach.

Comparison of Family Incomes (1993 SLID field test)		
	Snapshot Approach	Prorated Approach
Number of individuals	2 313	2 313
Mean	48 976	49 553
Standard deviation	33 337	33 747
Minimum	-12 162	-12 162
First quartile	24 839	25 355
Median	42 500	43 302
Third quartile	66 337	66 354
Maximum	333 045	337 613
Interquartile difference	41 498	40 999

Of the 2313 individuals, 233 were in households with a change in composition during the reference year. The following table shows the distribution of the differences between the two approaches for those persons with a change in family composition during the year. For cohabitants, it is necessary to assume that those joining the same household had previously lived together, since no information on their prior family situation was collected in the test.

The generally-higher family income under the prorated approach is again evident; it is higher for 55% of the individuals in families with composition changes, and lower for 23%.

Difference between Prorated Family Income and Snapshot Family Income
(Individuals in families with changes in composition during reference year)
(1993 SLID field test)

Difference (thousands)	Longitudinal persons plus cohabitants	Longitudinal persons only
< -25	6 (2.6)	1 (0.5)
-25 to < -20	1 (0.4)	1 (0.5)
-20 to < -15	5 (2.1)	0 (0.0)
-15 to < -10	13 (5.6)	8 (4.1)
-10 to < -5	9 (3.9)	7 (3.6)
-5 to < 0	19 (8.2)	12 (6.2)
0	53 (22.7)	42 (21.5)
> 0 to 5	39 (16.7)	36 (18.5)
> 5 to 10	32 (13.7)	32 (16.4)
> 10 to 15	17 (7.3)	17 (8.7)
> 15 to 20	13 (5.6)	13 (6.7)
> 20 to 25	11 (4.7)	11 (5.6)
> 25	15 (6.4)	15 (7.7)
TOTAL	233 (100.0)	195 (100.0)

Only incomes for 1992 were collected in the test. However, by assuming that the 1991 income was the same as the 1992 income and that there were no family composition changes during 1991, one can see the year over year effect of each of the Snapshot and Prorated approaches. Almost half the cohabitants were assigned to a separate economic family in the household they joined, due to a lack of information on them. It is not known to what extent this assumption is valid. For these persons, the incomes would be the same in each of the two years, since there was no change in the family composition. Also, the other economic family in the household would have the same family composition for the two years, so the

difference again would be zero. Therefore, the number of zero differences found in this study is much higher than what one would expect in practice.

The following table gives the difference of the 1991 and 1992 incomes.

Difference between 1991 Incomes and 1992 Incomes
(Individuals in families with changes in composition during 1992 and assuming no family composition changes in 1991)
(1993 SLID field test for 1992 income, assuming 1991 income was the same as 1992)

Difference (thousands)	Snapshot (frequency)		Prorated (frequency)	
	Longitudinal + Cohabitants	Longitudinal only	Longitudinal + Cohabitants	Longitudinal only
< -25	51	51	28	28
-25 to < -20	14	14	7	7
-20 to < -15	9	9	15	15
-15 to < -10	9	9	13	10
-10 to < -5	28	28	18	17
-5 to < 0	4	4	47	45
0	72	54	45	35
> 0 to 5	6	4	28	20
> 5 to 10	5	4	10	9
> 10 to 15	3	3	3	1
> 15 to 20	4	0	6	4
> 20 to 25	13	10	3	0
> 25	15	5	10	4
TOTAL	233	195	233	195

This table illustrates a point made earlier in the document. When a change in family composition occurs, the Snapshot Approach provides an immediate change

(in the year in which the change takes place). The Prorated Approach shows a change over two years, the year in which the change takes place and the following year. One can see that there are greater numbers of respondents for which the change was high (both negative and positive) using the Snapshot Approach.

Future Actions

1. Since it corresponds to traditional measures and it is the least problematic to derive, SLID output files will contain family income calculated using the Snapshot Approach.

2. At some point in the future (but not for the release of the first wave of data), SLID output files will contain a measure of family income calculated using the Subannual Approach (or a variant of it). Evaluation studies will be conducted to aid the detailed specifications of this variable, with results disseminated through the SLID Research Paper series.

3. User consultation will be undertaken to:

   a) help in deriving the detailed specification of the Subannual Approach;

   b) determine whether the family income variable using the Subannual Approach will replace the one using the Snapshot Approach or will be an alternate measure added to it.