

WORKINIG PAPER NO. SSMD-89-022 E
METHODOLOGY BRANCH

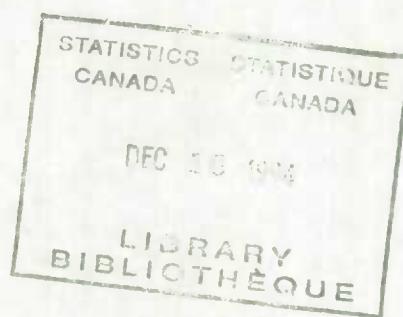
WP 358

**VARIANCE ESTIMATION METHODOLOGY AND
GENERAL PURPOSE VARIANCE ESTIMATION SYSTEM
FOR THE LABOUR FORCE SURVEY**

SSMD-89-022 E

Hyunshik Lee *
Social Survey Methods Division

November, 1989



* The author is now in the Business Survey Methods Division.

TABLE OF CONTENTS

1. INTRODUCTION	1
2. VARIANCE FORMULA FOR RAKING RATIO ESTIMATES OF THE LFS	2
2.1 One-Iteration Variance Formula for LF Monthly Estimates...	3
2.2 Variance Formula for a Linear Combination of LF Monthly Estimates	6
3. VARIANCE ESTIMATION COMPUTER SYSTEM FOR THE LFS ..	7
References	12

APPENDIX

A.1. Definition of the Recorded PSU Number	13
A.2. Record Layout of the VARCOVAR File	15
A.3. Definitions of Variables Used in the GP Computer System	16
A.4. Listing of the GP Computer System	19

Abstract

In the redesigned Labour Force Survey, the raking ratio estimation had been used until 1989. To provide the variance estimates of the raking ratio estimates, a new methodology and computer system had been developed for the published data. However, it is often necessary to calculate variances of user specified characteristics for a special purpose on an ad hoc basis. The existing system does not provide the flexibility to meet such needs. A general purpose variance estimation system was developed to meet these needs. This paper describes briefly the new variance estimation methodology and gives a detailed explanation of the system. It also includes the whole listing of the computer system.

Résumé

Dans le cadre du remaniement de l'Enquête sur la Population Active, un estimateur itératif par le quotient a été utilisé jusqu'en 1989. Dans le but de calculer la variance des estimations présentés dans la publication, il a été nécessaire de développer une nouvelle méthodologie et un système informatique appliquant cette méthodologie. Malheureusement, ce système n'est pas assez flexible pour répondre aux demandes d'estimations spéciales faites de la part des utilisateurs. Un second système a donc dû être élaboré afin de répondre à ces besoins particuliers. Dans cet article, les auteurs présentent la méthodologie de ce second système. Une description détaillée du système est également fournie. Finalement, une copie du programme informatique est jointe en appendice.

1. INTRODUCTION

The variance estimation methodology for the old Labour Force Survey (LFS) was the generalized Keyfitz method (Platek and Singh 1975). Keyfitz (1957) described a simple but useful method of estimating sample variance for a survey design where two units are selected from each stratum independently. His main contribution in the paper was the derivation of a very simple variance formula for post-stratified ratio estimates. In the derivation, he used the first order Taylor series approximation to linearize the ratio estimates and a simple trick of reversing the order of summation to simplify computation. Even though the method is simple and quite efficient, it has a serious limitation because it relies on the assumption that two units are selected independently from each stratum. Violations of the independence assumption can be circumvented by allowing some positive bias in the variance estimate or by using a finite population correction (fpc). But a complication arises in applying the fpc in the case of multi-stage sampling because the fpc should be applied only to the first stage variance component, while the Keyfitz formula gives the overall variance including lower stage components. However, the effect on the bias due to the violation of independence is not a serious problem as long as the fpc is small and the size of the first stage component of variance relative to the total variance is not large (Rust 1985). In the case of more than two units per stratum, Keyfitz suggested collapsing. But collapsing will bring some positive bias and cases with odd number of units will still be problems.

Woodruff (1971) generalized the Keyfitz method to the case of more than 2 units per stratum and also to more complicated estimates such as regression coefficients and correlation coefficients. He also used Taylor linearization. However, he used the replication formula to tackle the case of more than two units per stratum. This method was used in the old LFS to estimate variances of post-stratified estimates. The method, often called the Taylor method or the delta method, is given a theoretical treatment by Wolter (1985).

The raking ratio estimation procedure had been used in the redesigned LFS until 1989. For the variance estimation of the raking ratio estimates, a

method with successive application of Taylor linearization was adopted. In the next section, a variance formula obtained by this method will be given.

In Section 3, the development of a computer system for general purpose variance estimation using the formula is described.

2. VARIANCE FORMULA FOR RAKING RATIO ESTIMATES OF THE LFS

The successive application of Taylor linearization to obtain a variance formula for raking ratio estimates was used by Arora and Brackstone (1977a,b) and Brackstone and Rao (1979) for the case of simple random sampling of units or clusters. The application of this method to the LFS was first considered by G. Gray and an actual formula for one-iteration raking ratio estimate was derived by Choudhry and Lee (1987). The iteration here means a complete cycle of row and column ratio adjustments in this order. The row dimension is subprovincial area (CMA and non-CMA part of ER) and the column dimension is provincial age/sex groups (see Table 1 for age group classification). The raking procedure adjusts the survey weight iteratively in such a way that the marginal totals of the adjusted weight are equal to the projected population totals. The first adjustment is done on the row dimension using the subweight (denoted W_0) and this produces the adjusted weight, say W_1 . Then the procedure is continued for the second adjustment on the column dimension using W_1 and this produces another weight, say W_2 . This completes the first cycle (iteration). The same procedure is repeated once more using W_2 to obtain another pair of weights, say (W_3, W_4) . LFS estimates are based on the final weight W_4 .

In the following, we sketch the derivation of the variance formula for the one-iteration raking ratio estimate (based on W_2) which is used as an approximation for the W_4 -weighted estimate (for details, see Choudhry and Lee 1987). This formula will be referred to as the one-iteration variance formula.

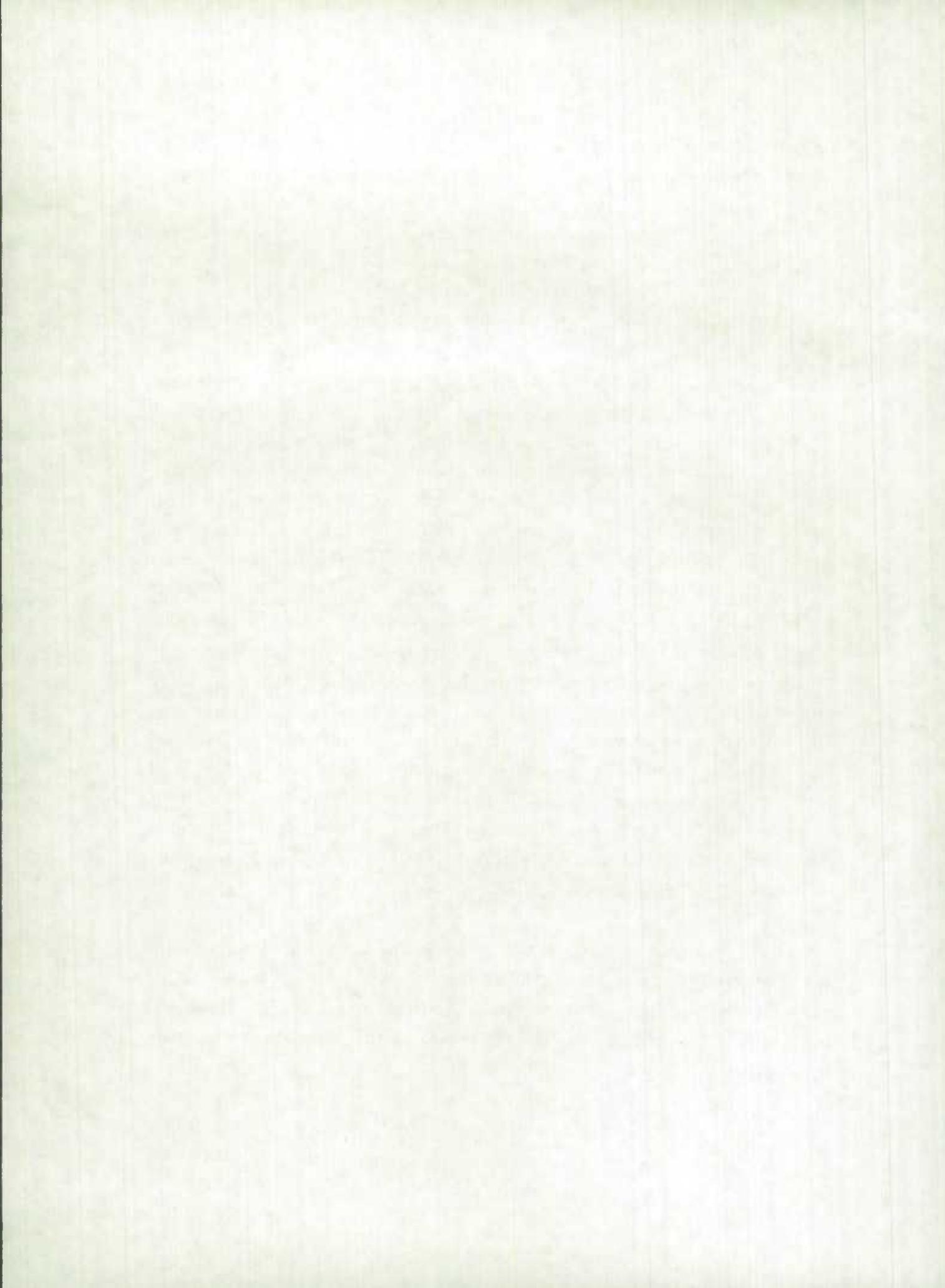


Table 1
Thirteen Age Group Definitions Used in the LFS

Group No.	Age Range
1	<15
2	15
3	16
4	17-19
5	20-24
6	25-29
7	30-34
8	35-44
9	45-54
10	55-59
11	60-64
12	65-69
13	70+

2.1 One-Iteration Variance Formula for LFS Monthly Estimates

Let $x^{(0)}$, $x^{(1)}$, $x^{(2)}$ be monthly estimates of a LF characteristic X in a certain province based on w_0 , w_1 , w_2 , respectively. (Note that obtaining a variance formula for a provincial estimate will suffice because the variance of a national estimate for X is simply the sum of the variances of the provincial estimates.)

Then $x^{(2)}$ can be expressed as:

$$x^{(2)} = \sum_a \frac{x_a^{(1)}}{p_a^{(1)}} p_a \quad (1)$$

where $x_a^{(1)} = w_1$ -weighted estimate of X for an age/sex group a in the province,

$p_a^{(1)} = w_1$ -weighted estimate of population 15+ (LF target population consisting of the population of 15 years of age and over excluding people in the armed forces, institutions and Indian reserves) for the age/sex group a ,

P_a = population 15+ for the a/s group a (from an external source such as a census projection).

Using the first order Taylor approximation to $\frac{x_a^{(1)}}{P_a^{(1)}}$ at $(E(x_a^{(1)}), E(P_a^{(1)}))$, we obtain

$$V(x^{(2)}) \doteq V\left\{\sum_a \frac{P_a}{E(P_a^{(1)})} \left(x_a^{(1)} - \frac{E(x_a^{(1)})}{E(P_a^{(1)})} P_a^{(1)} \right)\right\}. \quad (2)$$

The W_1 -weighted estimates $x_a^{(1)}$ and $P_a^{(1)}$ in (2) can be written as

$$x_a^{(1)} = \sum_s \frac{x_{sa}^{(0)}}{P_s^{(0)}} P_s \quad \text{and} \quad P_a^{(1)} = \sum_s \frac{P_{sa}^{(0)}}{P_s^{(0)}} P_s$$

where s denotes a sub-provincial area (CMA or non-CMA part of ER) in the province. Applying Taylor approximation to these ratios again, the following expression is obtained:

$$\begin{aligned} V(x^{(2)}) &\doteq V\left[\sum_a \frac{P_a}{E(P_a^{(1)})} \sum_s \frac{P_s}{E(P_s^{(0)})} \left\{ \left(x_{sa}^{(0)} - \frac{E(x_{sa}^{(0)})}{E(P_s^{(0)})} P_s^{(0)} \right) \right. \right. \\ &\quad \left. \left. - \frac{E(x_a^{(1)})}{E(P_a^{(1)})} \left(P_{sa}^{(0)} - \frac{E(P_{sa}^{(0)})}{E(P_s^{(0)})} P_s^{(0)} \right) \right\} \right]. \end{aligned} \quad (3)$$

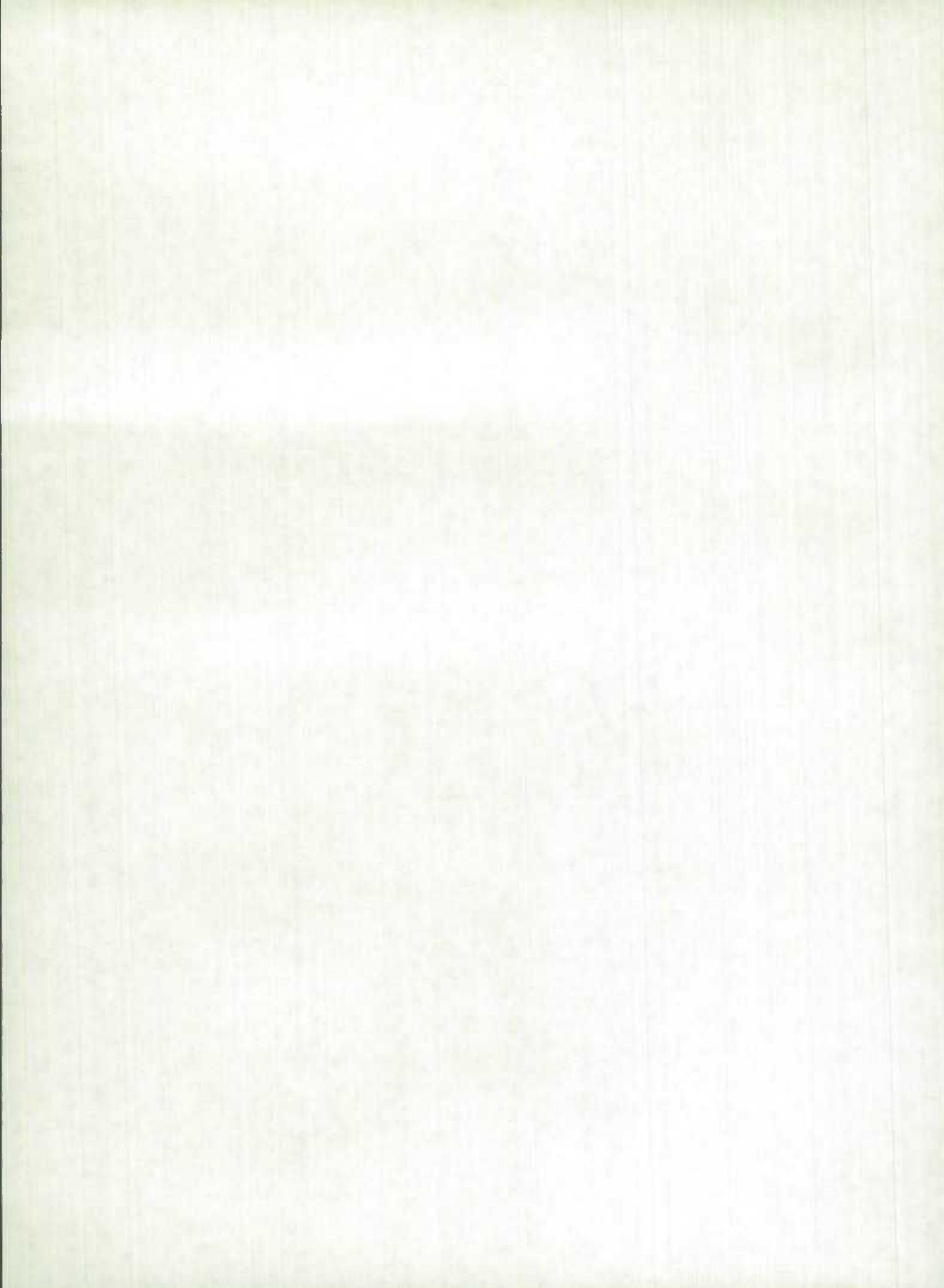
Writing (3) in terms of replicate level estimates and reordering the summations, we have

$$V(x^{(2)}) \doteq V\left(\sum_s \sum_{h \in s} \sum_{i=1}^{n_h} D_{shi}^{(0)}\right) \quad (4)$$

where

$$\begin{aligned} D_{shi}^{(0)} &= \sum_a \frac{P_a}{E(P_a^{(1)})} \frac{P_s}{E(P_s^{(0)})} \left\{ \left(x_{shia}^{(0)} - \frac{E(x_{shia}^{(0)})}{E(P_s^{(0)})} P_{shi}^{(0)} \right) \right. \\ &\quad \left. - \frac{E(x_a^{(1)})}{E(P_a^{(1)})} \left(P_{shia}^{(0)} - \frac{E(P_{shia}^{(0)})}{E(P_s^{(0)})} P_{shi}^{(0)} \right) \right\}, \end{aligned}$$

and h represents a stratum in the s-th subprovincial area.



Apart from special area strata, $(\sum_{i=1}^{n_h} D_{shi}^{(0)})$'s are independent because they are based on the subweight. But, they are highly correlated for the special area strata because of the way survey data from the special areas are treated, as explained in the following. There are 3 types of special areas, military establishments, institutions, and remote area. Each type of special area forms a stratum at the provincial level except in Quebec and Alberta where remote areas are further divided to form two strata. Thus there are at most 4 different special area strata in a province. The survey records corresponding to the selected units from a special stratum are attributed to every subprovincial area which contributes some part in forming the special stratum, with survey weight deflated proportionally to the population size of the part contributed by the subprovincial area. Those records are treated as if they were from a stratum in the subprovincial area. The same records with different weights can be found from other subprovincial areas which contribute some part to the special area stratum. Therefore, the sums of $D_{shi}^{(0)}$ -values for the special area strata are highly correlated.

Now let's write (4) as sum of variances of non-special area strata and special area strata as follows:

$$V(x^{(2)}) = V_1(\text{non-special area strata}) + V_2(\text{special area strata}).$$

Then,

$$V_1 = \sum_s \sum_{h \in s} \frac{n_h}{n_h - 1} \sum_{i=1}^{n_h} (D_{shi}^{(0)} - \bar{D}_{sh}^{(0)})^2 \quad (5)$$

where $\bar{D}_{sh}^{(0)} = \frac{1}{n_h} \sum_{i=1}^{n_h} D_{shi}^{(0)},$

and

$$V_2 = \sum_k^4 \frac{n_k}{n_k - 1} \sum_{i=1}^{n_k} (D_{ki}^{(0)} - \bar{D}_k^{(0)})^2 \quad (6)$$

where $D_{ki}^{(0)} = \sum_s \sum_{k \in s} D_{ski}^{(0)}$, and $\bar{D}_k^{(0)} = \frac{1}{n_k} \sum_{i=1}^{n_k} D_{ki}^{(0)}$. In (5) and (6),

h stands for a non-special stratum and k for a special stratum. Note that $D_{ki}^{(0)}$ is the sum of $D_{shi}^{(0)}$ -values for a special area stratum k scattered within the province. (From now on, the over bar will indicate the usual mean.)

An estimate of $V(x^{(2)})$, then, can be obtained by substituting expected values in (5) and (6) with corresponding estimates as follows:

$$\begin{aligned}\hat{V}(x^{(2)}) &= \sum_s \sum_{h \in s} \frac{n_h}{n_h - 1} \sum_{i=1}^{n_h} (D_{shi}^{(2)} - \bar{D}_{sh}^{(2)})^2 \\ &\quad + \sum_{k=1}^4 \frac{n_k}{n_k - 1} \sum_{i=1}^{n_k} (D_{ki}^{(2)} - \bar{D}_k^{(2)})^2\end{aligned}\quad (7)$$

where $D_{shi}^{(2)} = \sum_s \sum_{h \in s} \sum_a \left\{ \left(x_{shia}^{(2)} - \frac{p_{shi}^{(0)}}{p_s^{(0)}} x_{sa}^{(2)} \right) \right.$

$$\left. - \frac{x_a^{(2)}}{p_a} \left(p_{shia}^{(2)} - \frac{p_{shi}^{(0)}}{p_s^{(0)}} p_{sa}^{(2)} \right) \right\}$$

and $D_{ki}^{(2)}$ is the sum of $D^{(2)}$ -values for the special stratum k from the contributing subprovincial areas in the province. H. Choudhry pointed out that the expression in (7) can be simplified (see Choudhry and Lee (1987)). However, we adopted (7) for it is easier to implement. The expression in (7) is the variance formula for w_2 -weighted estimates and requires two weights w_0 and w_2 . However, the formula with (w_0, w_4) replacing (w_0, w_2) will be used to estimate the variances for w_4 -weighted estimates approximately. The approximate variance estimator was shown to perform reasonably well in Choudhry and Lee (1987).

2.2 Variance Formula for a Linear Combination of LF Monthly Estimates

Let $\sum_{m=1}^M c_m x_m^{(2)}$ be a linear combination of M monthly LF estimates $x_1^{(2)}, \dots, x_M^{(2)}$. A variance estimate of the linear combination can be obtained from the following formula:

$$\begin{aligned}\hat{V}\left(\sum_{m=1}^M c_m x_m^{(2)}\right) &= \sum_s \sum_{h \in s} \frac{n_h}{n_h - 1} \sum_{i=1}^{n_h} (L_{shi}^{(2)} - \bar{L}_{sh}^{(2)})^2 \\ &\quad + \sum_{k=1}^4 \frac{n_k}{n_k - 1} \sum_{i=1}^{n_k} (L_{ki}^{(2)} - \bar{L}_k^{(2)})^2\end{aligned}\quad (8)$$

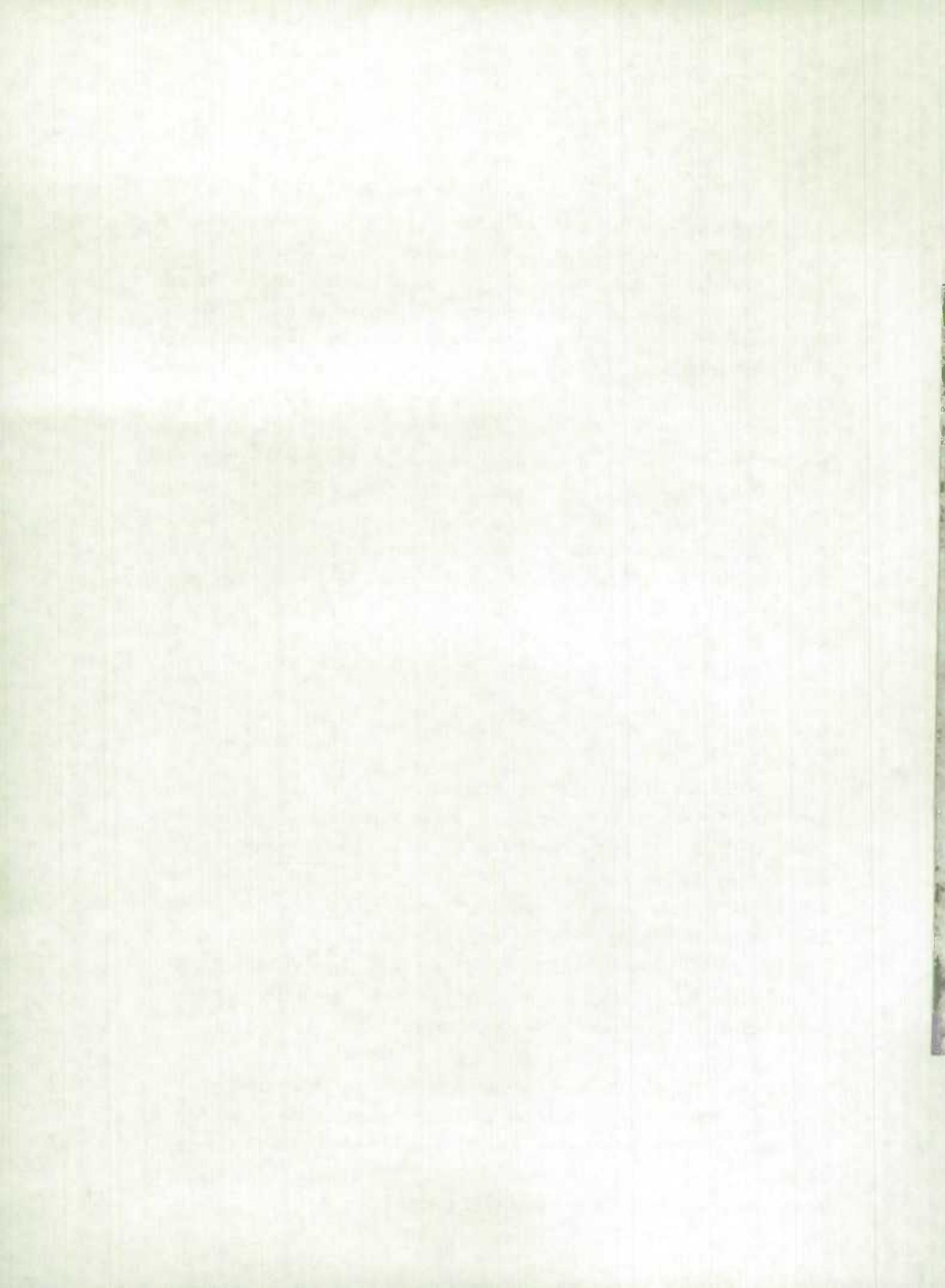
where $L_{shi}^{(2)} = \sum_{m=1}^M c_m D_{mshi}^{(2)}$ and $L_{ki}^{(2)} = \sum_{m=1}^M c_m D_{mki}^{(2)}$, and $D^{(2)}$ -

values are obtained as in Section 3.2 for each month. The formula (8) can be obtained easily by applying the same method used in Section 3.2 to each of $x_m^{(2)}$, $m=1, \dots, M$, assuming that the same replicates stay over the M -month period. (It is trivial to see that the formula (7) is a special case of (8)). The formula is strikingly simple because there is no covariance term. The simple algebraic manipulation of reordering the summation signs makes this simple form possible. The formula can be used even when there is rotation of Primary Sampling Units (PSU) (generally, PSU's are taken as replicates) as long as the number of replicates (PSU's) is the same and the incoming PSU's retain the identification numbers of the outgoing PSU's. In practice, however, there are a small number of cases where the number of replicates changes over time. Even for this case, the formula can be used with some bias implications assuming imaginary replicates are substituted for missing ones whose $D^{(2)}$ -values are taken as zero. Again this formula with (W_0, W_4) will be used to obtain approximate variance estimates of linear combinations of W_4 -weighted estimates.

3. VARIANCE ESTIMATION COMPUTER SYSTEM FOR THE LFS

The formula (7) was used to develop the variance estimation system "VARCOVAR" which computes variance estimates for prespecified LF characteristics. The prime purpose of the system is to provide reliability indicators in the monthly publication "The labour force" (Cat. 71-001) based on six month averages. For this reason, the system is not flexible enough to be easily modified to meet customized needs. Moreover, the VARCOVAR system cannot compute variance estimates for linear combinations of monthly estimates. That is why a general purpose (GP) variance estimation computer system was developed using the formula (8).

In the following the basic structure of this GP computer system is described. The two systems differ in the treatment of apartment frames. The VARCOVAR system employs the Kish adjustment (1965) to reduce the impact of uneven replicate sizes while the GP system does not. The improvement due to this adjustment seems to be minimal.



The GP system consists of five steps, each of which executes a program, as follows:

1. Step 0 (FSPLIT)

This step separates data from the master file using the utility program FSPLIT. This step is optional.

2. Step 1 (FORTGCLG)

This step executes a FORTRAN program which defines necessary parameters and selected LF characteristics and accumulates various weights at replicate, subprovincial, and provincial levels. Detailed description of the step will be given later.

3. Step 2 (SORT)

This step executes the utility program SORT to sort the replicate level file created in Step 1 by the recoded PSU number (the definition is given in the Appendix) and survey month.

4. Step 3 (SORT)

Another SORT program is executed to sort the subprovincial level file created in Step 1 by the first 5 digits of the recoded PSU number and survey month.

5. Step 4 (FORTGCLG)

This is another FORTRAN program which computes $D^{(2)}$ and $L^{(2)}$ values and variance estimates for selected characteristics and then prints the results. Detailed description will be given later.

The major steps are Steps 1 and 4 whose basic structure is described in Figures 1 and 2. The GP system can compute variances of up to 200 characteristics and handle up to 12 months data. The input file should be sorted by survey month and the recoded PSU number and contain some essential variables such as the final weight (W_4), subweight, the recoded PSU number, age.sex code, survey year and month. All these variables can be obtained from the VARCOVAR file which is used as an input file in the VARCOVAR system. The record layout of the VARCOVAR file is given in the Appendix. The complete listing of the system is given in the Appendix with definitions of variables used in the system. The listing also provides an example which will produce variance estimates of the employed and the unemployed for ten provinces. Basic explanations are imbedded in the program itself as comment statements which the user should read carefully, so that no major difficulty is faced in using the system.

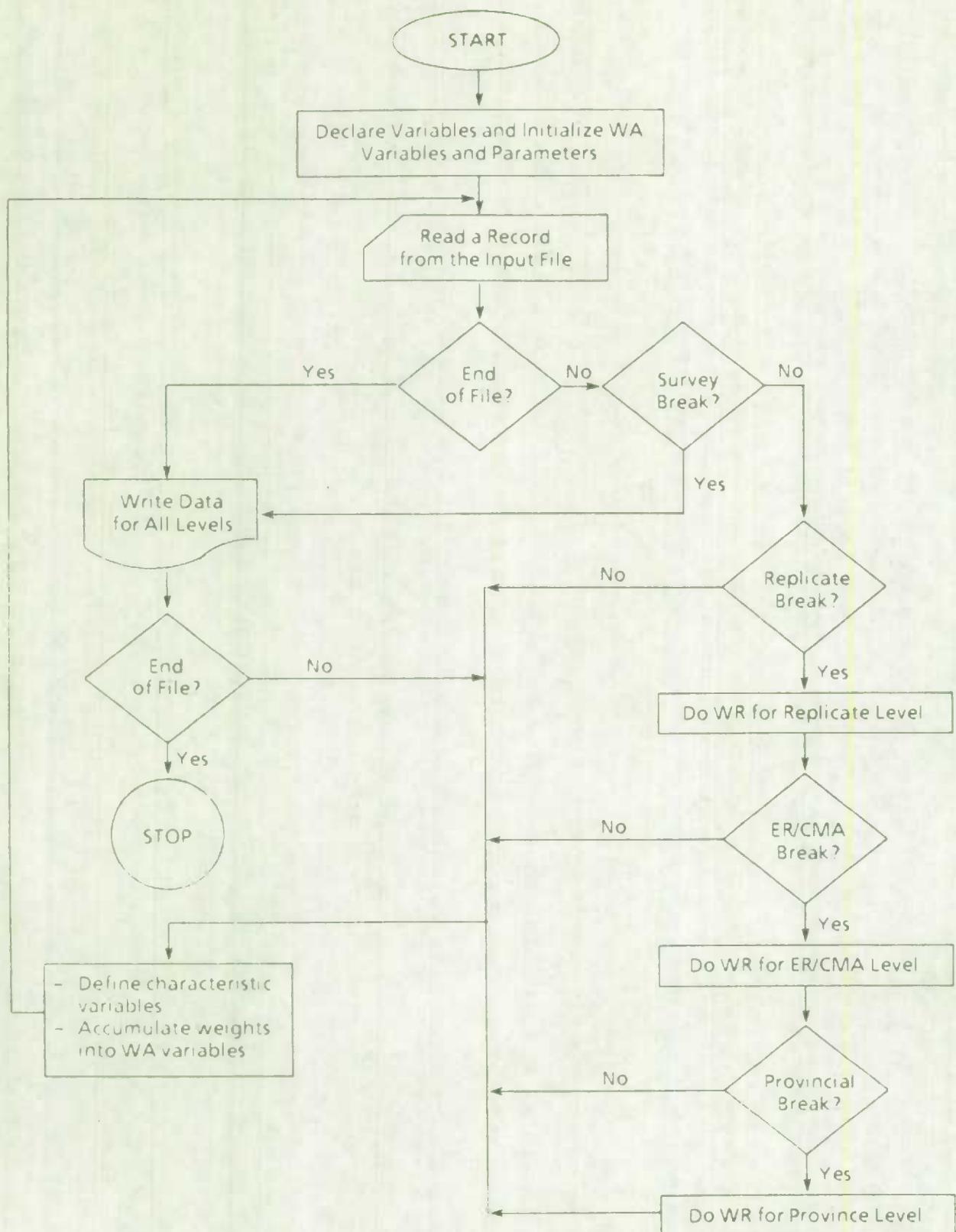


Figure 1 Flow Chart of the FORTRAN Program in Step 2 WA stands for "Weight Accumulation" and WR for "Write data and Reinitialize the WA variables"

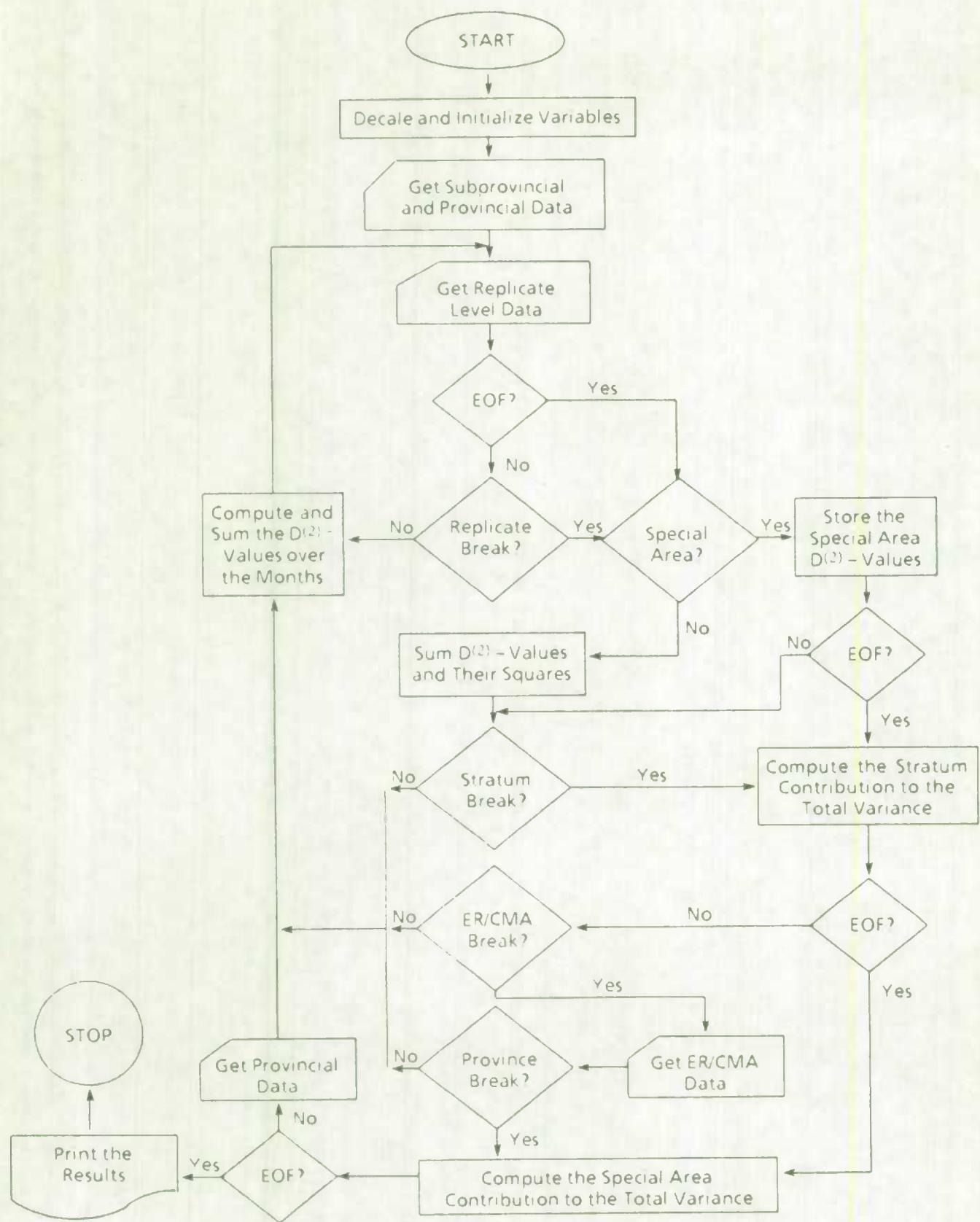
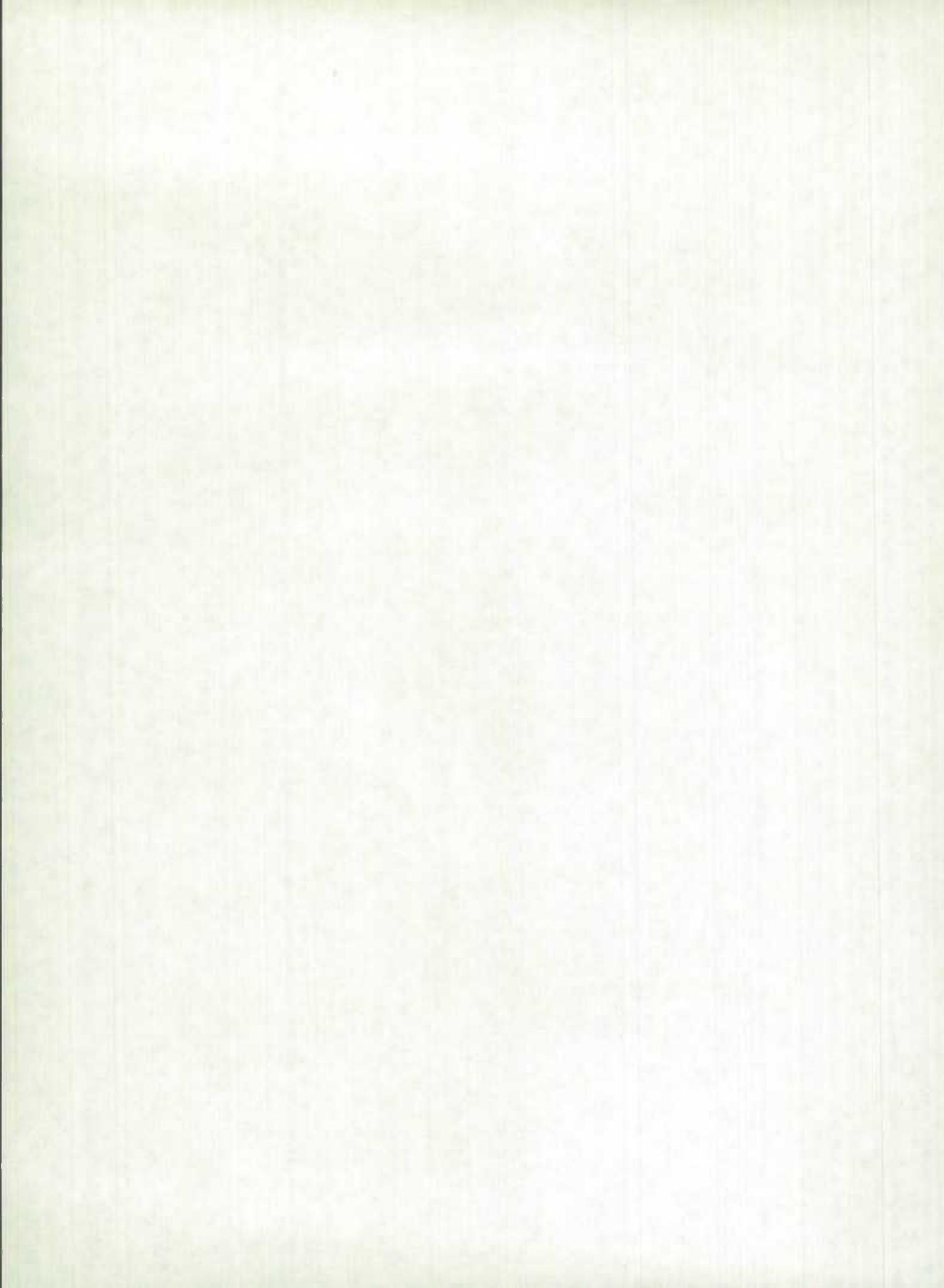


Figure 2 Flow Chart of the FORTRAN Program in Step 4 EOF stands for "End of File".



In Step 1, the user is asked to declare variables to be used in the program besides the program specified variables (see lines 4400-5400 in the listing). Use of a variable name that is already declared or used without declaration in the program must be avoided. Then the following variables should be assigned with appropriate values:

- a. NMNTH (number of months)
- b. SURVEY (vector of survey months) — it is initialized as 12 zeros. When a new value is assigned, a 4-digit integer number (e.g. 8506) should be used.
- c. C (vector of coefficients of linear combination) — it is initialized as zeros.
- d. NC (number of characteristics) — this number cannot exceed 200.

The read statement (lines 7600 & 7700) should then be altered according to the data organization of the input file. Note that the 10-digit recoded PSU number should be read as two 5-digit numbers since FORTRAN can not handle full range of 10-digit integer numbers as an integer variable. The next thing to do is to define characteristics in line 17700- and onward. Right after the characteristic definition, there is a line where user defined weight can be given. If this is not redefined, the final weight will be used. In Step 4, there are lines (starting from 29300) where characteristic names can be assigned through variables NAME1 and NAME2. These are double-precision real-valued vectors with 200 components and initialized as blank. Up to sixteen character names can be given by NAME1 and NAME2. If no name is given here, 16 blank spaces will be printed.

Even though the number is usually small, there may be single replicate strata. The variance formula (7) & (8) would not work for the single replicate strata because $n_h - 1 = 0$. The GP system handles these cases by simply putting $n_h = 2$ and using the same formula (8). The effect of this treatment is to add squares of $L^{(2)}$ -values of such strata to the total variance. That is, it gives conservative estimate of variance. However, this effect is not significant as long as the number of such strata is small. The system prints out contribution to the total variance by the single replicate strata so that the user can check the extent of over-estimation due to this treatment.

ACKNOWLEDGEMENT

The author would like to thank J. Gambino and J. Armstrong of Statistics Canada for their helpful comments and Judy Clarke for her excellent typing.

REFERENCES

- Arora, H.R., and Brackstone G.J. (1977a). An investigation of the properties of raking ratio estimators: I with simple random sampling. *Survey Methodology*, 3, 62-83.
- Arora, H.R., and Brackstone, G.J. (1977b). An investigation of the properties of raking ratio estimators: II with cluster sampling. *Survey Methodology*, 3, 232-252.
- Choudhry, G.H., and Lee, H. (1987). Variance estimation for the Canadian Labour Force Survey. *Survey Methodology* 13, 147-162.
- Keyfitz, N. (1957). Estimates of sampling variance where two units are selected from each stratum. *Journal of the American Statistical Association*, 52, 503-510.
- Kish, L. (1965). *Survey Sampling*. New York: John Wiley.
- Platek, R., and Singh, M.P. (1976). Methodology of the Canadian Labour Force Survey. Catalogue No. 71526, Statistics Canada.
- Rust, K. (1985). Variance estimation for complex estimators in sample surveys. *Journal of Official Statistics*, 1, 381-397.
- Wolter, K.M. (1985). *Introduction to Variance Estimation*. New York: Springer-Verlag.
- Woodruff, R.S. (1971). A simple method for approximating the variance of a complicated estimate. *Journal of the American Statistical Association*, 66, 411-414.

APPENDIX

A.1. Definition of the Recorded PSU Number

The recoded PSU number is the replicate identification number which is unique throughout the whole nation. It consists of 10 digits, 4 more than in the old design, retaining all geographic information. In general, the 10-digit number will consist of the following fields:

Field 1 : Position 1 : Province (0-9)

Field 2 : Positions 2-3 : CMA/non-CMA code. For areas not belonging to a CMA, this field will be "00". For CMA's it will contain the 2-digit CMA (or CA) code, of which there are 35 possibilities.

Field 3 : Positions 4-5 : 2nd and 3rd digits of ER.

Field 4 : Position 6 : 3rd digit of PSU : i.e. 0 = NSRU
9 = Special Areas
1-8 = SRU

Field 5 : Position 7 : Old/New sample code: 0 = old sample
1 = new sample (redesign)

Field 6 : Position 8-9 : Stratum Number within Type of Area.

In NSR areas, this will be the design stratum number, i.e. 00, 20, 40, 60, 80 or 99 for special NSR strata. In PEI this field will contain the PSU number, i.e. the last 2 digits of the NSR PSU. In SRU areas position 8 will be super stratum; position 9 the stratum. For apartments, the strata will be numbered sequentially within the SRU as 10, 20, 30, The "0" in position 9 will distinguish the apartment strata from the regular SRU. For the special areas this field will essentially contain the group number, acceptable values being 01, 02, 03 and 04. The entire Quebec remote urban stratum will have 04 in this field. In Alberta, for remote areas only, ER 870 will be stratum 03, ER 880 will be stratum 04 to distinguish the two frames. There are no other occurrences of stratum 04.

Field 7 : Position 10 : Replicate Number within Stratum.

For SRU, apartments, PEI NSR, and special NSR strata ending in "99", this field contains the rotation number. In special areas, the replicate number will be the rotation number, except in the Quebec remote urban stratum. There the two clusters selected in the first selected town will have replicate number 1, the two clusters in the second selected town will have replicate number 2. In the rest of the NSR areas (excluding PEI and the special strata ending in "99"), the replicate number will be a number for each PSU, numbered sequentially within the stratum.

The following chart provides a summary and overview of the definition of the recoded PSU number.

Field	Description	Position	SRU	APT	NSR	Special
1	Province	1	0-9	0-9	0-9	0-9
2	CMA/Non-CMA Code	2-3	CMA Code; "00" if non-CMA	CMA Code	00	CMA Code; "00" if non-CMA
3	2nd & 3rd digits of ER	4-5	00-99	00-99	00-99	00-99
4	3rd digit of PSU	6	1-8	1-8	0	9
5	Old/new sample code	7	0 = old 1 = new	0 = old 1 = new	0 = old 1 = new	0 = old 1 = new
6	Stratum	8-9	8 - super stratum 9 - stratum 9 - "0"	8 - sequential number of stratum within SRU number 99 as per design	PEI : 01-12 Rest: 00, 20, 40, 60, 80 or 99 as per design	01 - military 02 - institutions 03 - remote 04 - Quebec remote urban stratum & Alberta ER 880 remote
7	Replicate number within stratum	10	rotation	rotation	PEI & stratum 99 : rotation Rest: sequential number of PSU within stratum	rotation (Quebec remote urban coded to one rotation number per town selected)

A.2. Record Layout of the VARCOVAR File

Position	Size	Title and Description
1-280	280	as per TABS file
281-282	2	Age/Sex Recode : 1-13 for male 14-26 for female
283-292	10	Recoded PSU number
293	1	Area Type : 1 - NSR 2 - SRU 3 - Special area
294-295	2	Person Number within Household (1-99)
296-297	2	*Number of Selection per Rotation for Apartment Records Only (blank for other records)
298-299	2	*Number of Selection per Stratum for Apartment Records only (blank for other records)
300	1	Blank

*These numbers are used for Kish adjustment

A.3. Definitions of Variables Used in the GP Computer System

In the following, definitions of variables used in the GP system are given in the order declared in the program specification statements. If a variable is a vector or matrix, dimensionality is indicated by numbers in parentheses.

Double precision real variables in Step 1:

EX(200)	- final weighted estimates of LF characteristics
FXSHI(200)	- final weighted characteristic estimates at replicate level
FPSHIA(26)	- final weighted population estimates for 26 age.sex groups at replicate level
FXS(200)	- final weighted characteristic estimates at subprovincial level
SPS	- subweighted population estimate at subprovincial level
FPSA(26)	- final weighted population estimates for 26 age.sex groups at subprovincial level
FXA(200,26)	- final weighted characteristic estimates for 26 age.sex groups at provincial level
FPA(26)	- final weighted population estimates for 26 age.sex groups at provincial level
SW	- subweight
FW	- final weight
U	- 10 digit recoded PSU number in real number of the record which is being processed
OLDU	- this is the same as U until a record with a different recoded PSU number is read. This is used to check replicate break.
C(12)	- coefficients in forming a linear combination of 12 monthly estimates
UW	- user defined weight

Integer variables in Step 1:

SURVEY(12)	- survey year and month (e.g. 8605)
ID1	- the first 5-digit of the recoded PSU number (which is the subprovincial area identification) of the record currently processed
OLDID1	- this is the same as ID1 until a record with a new value for ID1 is read, and is used to check subprovincial area break
ID2	- the second 5-digit of the recoded PSU number of the record currently processed
OLDID2	- this is the same as ID2 until ID2 changes
A	- age/sex code as in VARCOVAR file
Y	- survey year (2-digit)
M	- survey month (2-digit)
P	- province code (e.g. 0 -Newfoundland)
NC	- number of characteristics
NMNTNH	- number of months
MCNT	- month count
YM	- survey year and month (e.g. 8605)
OLDYM	- same as YM until it changes and is used to check survey break
RECNO	- record number in direct-access files

Logical variables in Step 1:

FLAG	- flag indicating whether all records are read in or not
RLAG	- flag indicating whether the record read in is the first record or not.
CHAR(200)	- indicator variable which indicates whether or not the record has given characteristics

Temporary variables in Step 1:

L, K7, K8	- variables which define direct-access files
IC, IA	- index variables in DO-loops

Double precision real variables in Step 4:

X(200), FXSHI(200), SPSHI, FPSHIA(26)- same as in Step 1

FXS(12,200), SPS(12), FPSA(12,26), FXA(12,200,26), FPA(12,26) - same as in Step 1 except month can be identified by the first argument of the vectors or matrices here.

DVAL(200)	- $D^{(2)}$ -values
NAME1(200)	- the first eight characters of characteristic names
NAME2(200)	- the second eight characters of characteristic names
SDVAL(4,6,200)	- special area $D^{(2)}$ -values by type of special area, replicate and characteristic
VAR(200)	- variance estimates
SRCV(200)	- single replicate contributions to the total variance
DSUM(200)	- sums of $D^{(2)}$ -values of replicate
SSUM(200)	- sums of squares of $D^{(2)}$ -values of replicate
TV	- variance contribution of replicate (used as temporary variable)
U,OLDU	- same as in Step 1
ZX,ZP,RFA	- intermediate values to calculate $D^{(2)}$ -value.
C(12)	- same as in Step 1

Integer variables in Step 4:

ID1, ID2, A, P, NC, MCNT, OLDDID1 - same as in Step 1

OLDP	- same as P until a new value is read and used to check province break
SA	- special area code
OLDSA	- same as SA until a new value is read and used to keep SA value before change of the value occurs
STR	- stratification number (the first 9-digit of the 10-digit recoded PSU number)
OLDSTR	- same as STR until a new value is read and used to check stratum break

NH - number of replicates in a stratum
NSRS(10) - numbers of simple replicate strata within province
SNH(4) - number of replicates in a special area stratum
NMNTH,RECNO,SURVEY(12) - same as in Step 1

Logical variables in Step 4:

FLAG,RFLAG - same as in Step 1
SRFLAG - indicator variable which indicates if the stratum currently processed has only one replicate or not
PSCODE - indicator variable which indicates whether the province currently processed has special area or not
SCNTFG(4) - special area stratum counting flags which is used to get proper count of replicates in each type of special area stratum

Temporary variables in Step 4:

L,K7,K8 - same as Step 1
IM,IC,IA,I,J - index variables in DO-loops
IS - last two digits of STR which gives the type of special area

A.4. Listing of the GP Computer System

The whole listing of the GP system is given below. It gives also an example from which we can compute variances for Employed and Unemployed for ten provinces from the June 1986 survey.

START
COL

```

1 //SLFSVAR JOB (3111,TA48,,10),HLEE,MSGCLASS=A,CLASS=E,TIME=(2,59),      00000100
1 // NOTIFY=HLEE,MSGLEVEL=(2,0)                                              00000200
1 //*****                                                               00000300
1 /**
1 /**
1 //** THIS PROGRAM IS TO OBTAIN ESTIMATES AND THEIR VARIANCE      *** 00000400
1 //** ESTIMATES OF LINEAR COMBINATIONS OF LFS MONTHLY TOTALS OF      *** 00000500
1 //** USER SPECIFIED CHARACTERISTICS. FOR VARIANCE ESTIMATION      *** 00000600
1 //** METHODOLOGY AND DETAILED DESCRIPTION OF THIS PROGRAM,        *** 00000700
1 //** REFER TO THE TECHNICAL REPORT BY HYUNSHIK LEE, ENTITLED       *** 00000800
1 //** "VARIANCE ESTIMATION METHODOLOGY AND GENERAL PURPOSE        *** 00000900
1 //** VARIANCE ESTIMATION SYSTEM FOR THE LABOUR FORCE SURVEY."     *** 00001000
1 /**
1 /**
1 //** IF THIS STEP IS GOING TO BE USED TO SEPARATE DATA FROM A      *** 00001100
1 //** MASTER FILE, GIVE THE MASTER FILE AS INPUT FILE IN THE        *** 00001200
1 //** FOLLOWING DD STATEMENT NAMED 'TAPEIN' AND CHANGE THE CONTROL    *** 00001300
1 //** STATEMENTS ACCORDINGLY (CAUTION: THE INPUT FILE SHOULD BE      *** 00001400
1 //** SORTED BY SURVEY MONTH AND RECODED PSU NUMBER). OTHERWISE,      *** 00001500
1 //** INSERT '/*' IN FRONT OF LINES 2300-3100 TO NULLIFY THIS STEP. *** 00001600
1 /**
1 //*****                                                               00002100
1 //STEP0 EXEC FSPLIT                                                 00002200
1 //TAPEIN DD DSN=RLFS.HLEE.MAR85.FEB86.DATA,DISP=OLD                00002300
1 //OUT1 DD DSN=&&SPLIT,UNIT=TAPEMK,DISP=(,PASS),DCB=*.TAPEIN          00002400
1 //SYSIN DD *                                                       00002500
1 INPUT=TAPEIN,CONTROL=(1,4,A)                                         00002600
1 OUTPUT=OUT1                                                       00002700
1 SELECT=OUT1,CF1=(= 685)                                             00002800
1 OPTION=PRINT                                                       00002900
1 /*
1 //STEP1 EXEC FORTGCLG,TIME.GO=10                                     00003000
1 //FORT.SYSIN DD *                                                   00003100
1 C
1 C   PROGRAM SPECIFICATION STATEMENTS                               00003200
1 C
1 C   REAL*8 X(200)/200*0.D0/,FXSHI(200)/200*0.D0/,SPSHI/0.00/,      00003300
1 C   # FPSHIA(26)/26*0.D0/,FXS(200)/200*0.D0/,SPS/0.D0/,           00003400
1 C   # FPSA(26)/26*0.D0/,FXA(200,26)/5200*0.D0/,FPA(26)/26*0.D0/, 00003500
1 C   # SW,FH,U,OLDU,C(12)/12*0.D0/,UM                                00003600
1 C   INTEGER SURVEY(12)/12*0/,ID1,OLDID1,ID2,OLDID2,A,Y,M,P,NC,      00003700
1 C   # NMNTH/1/,MCNT/1/,YH,OLDDYM,RECNO                            00003800
1 C   LOGICAL#1 FLAG//.FALSE./,RFLAG//.FALSE./,CHAR(200)/200*.FALSE./ 00003900
1 C
1 C   USER SPECIFICATION STATEMENTS -- THE INPUT FILE SHOULD CONTAIN 00004000
1 C   VARIABLES FH(FINAL WEIGHT),SW(SUBWEIGHT),ID1(THE FIRST 5-DIGIT OF 00004100
1 C   10-DIGIT RECODED PSU NUMBER),ID2(THE SECOND 5-DIGIT OF 10-DIGIT 00004200
1 C   RECODED PSU NUMBER),A(AGE/SEX CODE),Y(SURVEY YEAR),M(SURVEY MONTH), 00004300
1 C   ALL OTHER VARIABLES IN THE INPUT FILE TO BE READ SHOULD BE      00004400
1 C   DECLARED HERE. CAUTION: NEVER USE THE SAME VARIABLE NAMES USED IN 00004500
1 C   THE PROGRAM SPECIFICATION STATEMENTS AND L, K7, K8, IC, IA      00004600
1 C   WHICH ARE ALREADY USED IN THE PROGRAM.                           00004700
1 C
1 C   INTEGER ROT,LFS,IND                                           00004800
1 C
1 C   INTEGER ROT,LFS,IND                                           00004900
1 C
1 C   INTEGER ROT,LFS,IND                                           00005000
1 C
1 C   INTEGER ROT,LFS,IND                                           00005100
1 C
1 C   INTEGER ROT,LFS,IND                                           00005200
1 C
1 C   INTEGER ROT,LFS,IND                                           00005300
1 C
1 C   INTEGER ROT,LFS,IND                                           00005400

```


DATASET: RLFS.HLEE.GENERAL.LFSVAR.CNTL

DATE: 87/07/17
TIME: 13:25
PAGE: 2

START

```

COL -----+-----1-----+-----2-----+-----3-----+-----4-----+-----5-----+-----6-----+-----7-----+-----8

1 C
1 C DIRECT ACCESS DATA SETS
1 C
7     DEFINE FILE 7(480,10400,L,K7)
7     DEFINE FILE 8(120,208,L,K8)
1 C
1 C ASSIGN NMNTH (NO. OF MONTHS) VALUE AND SPECIFY SURVEY YEAR & MONTH.
1 C
7     NMNTH=1
7     SURVEY(1)=8506

1 C
1 C SPECIFY COEFFICIENTS OF LINEAR COMBINATION. THEY ARE INITIALIZED
1 C AS ZERO.
1 C
7     C(1)=1.0
1 C
1 C ASSIGN NC (NO. OF CHARACTERISTICS) VALUE.
1 C
7     NC=20
1 C
2 10  READ(4,12,END=20)M,Y,ROT,LFS,IND,SM,FH,A,ID1,ID2
2 12  FORMAT(I2,I2,2I1,I2,2F9.4,I2,I5,I5)
7     P=ID1/10000
7     U=ID1*100000.D0+ID2
7     YM=Y*100+M
1 C
1 C CHECK IF THE RECORD READ IN IS THE FIRST RECORD.
1 C
7     IF (RFLAG) GOTO 40
7     IF (YM.NE.SURVEY(1)) GOTO 950
7     OLDDYM=YM
7     OLDP=P
7     OLDU=U
7     OLDid1=ID1
7     OLDid2=ID2
7     RFLAG=.TRUE.
7     GOTO 100
1 C
1 C END-OF-FILE
1 C
2 20  FLAG=.TRUE.
7     GOTO 50
1 C
1 C CHECK FOR SURVEY BREAK.
1 C
2 40  IF (YM.EQ.OLDDYM) GOTO 60
7     IF (YM.EQ.SURVEY(MCNT+1)) GOTO 50
7     IF (MCNT+1.LE.NMNTH) GOTO 950
7     GOTO 20
2 50  WRITE(5)OLDid1,OLDid2,MCNT,FXSHI,SPSHI,FPSHIA
7     WRITE(6)OLDid1,MCNT,FXS,SPS,FPSA
7     RECNO=OLDP*12+MCNT
7     WRITE(7'RECNO) ((FXA1 IC,IA),IC=1,50),IA=1,26)

```


DATASET: RLFS.HLEE.GENERAL.LFSVAR.CNTL

DATE: 87/07/17

TIME: 13:25

PAGE : 3

START

```

COL -----+-----1-----+-----2-----+-----3-----+-----4-----+-----5-----+-----6-----+-----7-----+-----8-----+
7      WRITE(7'RECNO+120) ((FXA(IC,IA),IC=51,100),IA=1,26)          00010900
7      WRITE(7'RECNO+240) ((FXA(IC,IA),IC=101,150),IA=1,26)          00011000
7      WRITE(7'RECNO+360) ((FXA(IC,IA),IC=151,200),IA=1,26)          00011100
7      WRITE(8'RECNO) FPA                                         00011200
7      IF (FLAG) GOTO 200                                         00011300
7      DO 52 IC=1,NC                                           00011400
9        FXSHI(IC)=0.D0                                         00011500
9        FXS(IC)=0.D0                                         00011600
9        DO 52 IA=1,26                                         00011700
11       FXA(IC,IA)=0.D0                                         00011800
2      52  CONTINUE                                         00011900
7      DO 54 IA=1,26                                         00012000
9        FPSHIA(IA)=0.D0                                         00012100
9        FPSA(IA)=0.D0                                         00012200
2      54  FPA(IA)=0.D0                                         00012300
7      SPS=0.D0                                         00012400
7      SPSHI=0.D0                                         00012500
7      OLDMY=YM                                         00012600
7      OLDP=P                                         00012700
7      OLDU=U                                         00012800
7      OL DID1=ID1                                         00012900
7      OL DID2=ID2                                         00013000
7      MCNT=MCNT+1                                         00013100
7      GOTO 100                                         00013200
1      C
1      C  CHECK FOR REPLICATE BREAK.
1      C
2      60  IF (OLDU.GT.U) GOTO 930                         00013400
7      IF (OLDU.EQ.U) GOTO 100                         00013500
7      WRITE(5)OLDID1,OLDID2,MCNT,FXSHI,SPSHI,FPSHIA
7      OLDU=U                                         00013600
7      OL DID2=ID2                                         00013700
7      DO 70 IC=1,NC                                         00013800
2      70  FXSHI(IC)=0.D0                                         00013900
7      DO 72 IA=1,26                                         00014000
2      72  FPSHIA(IA)=0.D0                                         00014100
7      SPSHI=0.D0                                         00014200
00014300
2      70  FXS(IC)=0.D0                                         00014400
7      DO 72 IA=1,26                                         00014500
2      72  FPSA(IA)=0.D0                                         00014600
1      C
1      C  CHECK FOR ER/CMA BREAK.
1      C
7      IF (OLDID1.EQ.ID1) GOTO 100                         00014700
7      WRITE(6)OLDID1,MCNT,FXS,SPS,FPSA
7      OLDID1=ID1                                         00014800
7      SPS=0.D0                                         00014900
7      DO 80 IC=1,NC                                         00015000
2      80  FXS(IC)=0.D0                                         00015100
7      DO 82 IA=1,26                                         00015200
2      82  FPSA(IA)=0.D0                                         00015300
00015400
2      80  FXS(IC)=0.D0                                         00015500
7      DO 82 IA=1,26                                         00015600
2      82  FPSA(IA)=0.D0                                         00015700
1      C
1      C  CHECK FOR PROVINCE BREAK.
1      C
7      IF (OLDP.EQ.P) GOTO 100                         00015800
7      RECNO=OLDP*12+MCNT
7      WRITE(7'RECNO) ((FXA(IC,IA),IC=1,50),IA=1,26)          00015900
00016000
7      RECNO=OLDP*12+MCNT
7      WRITE(7'RECNO) ((FXA(IC,IA),IC=1,50),IA=1,26)          00016100
00016200

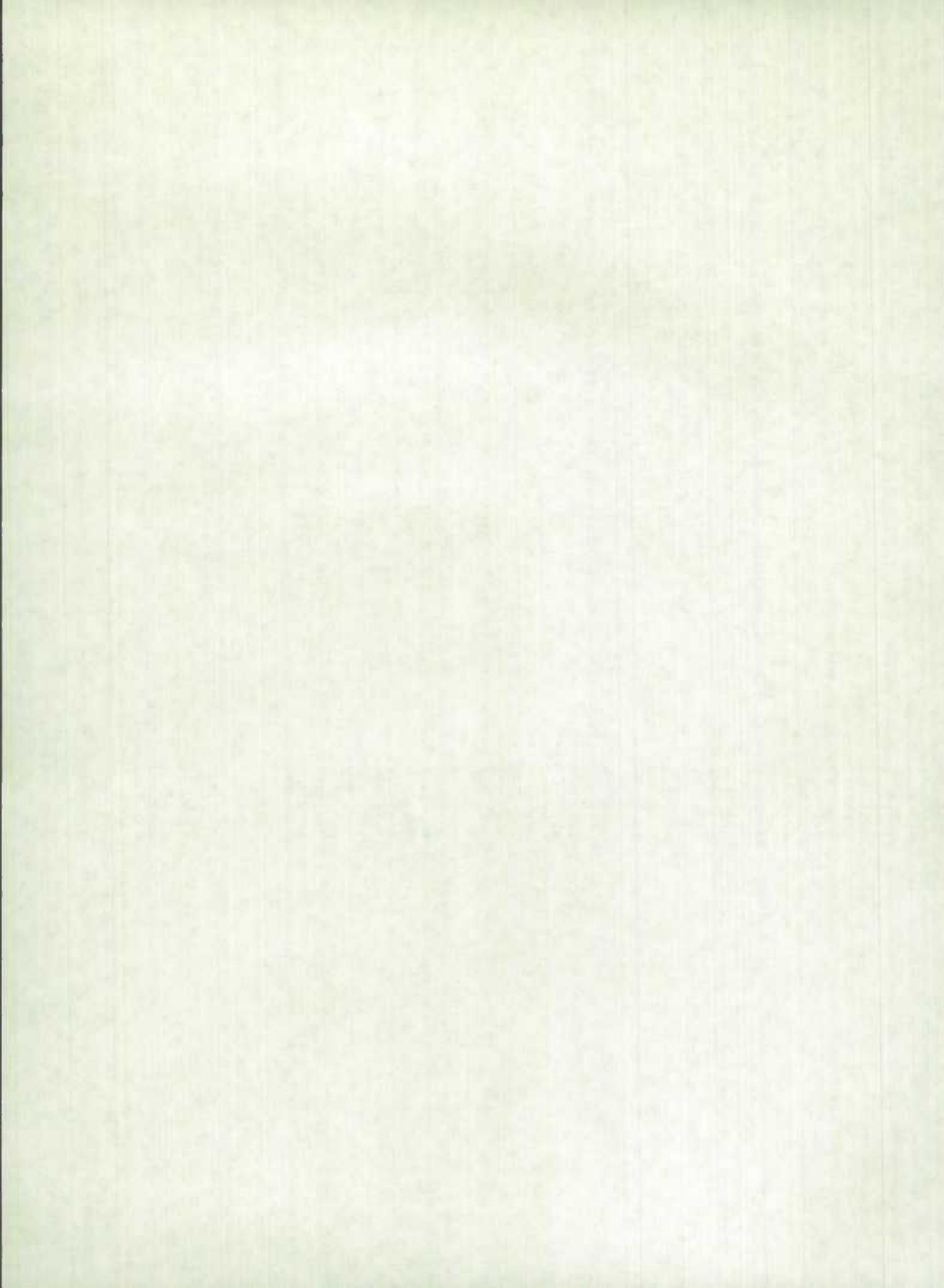
```


START

COL	1	2	3	4	5	6	7	8
7	WRITE(7'RECNO+120)	((FXA(IC,IA),IC=51,100),IA=1,26)			00016300			
7	WRITE(7'RECNO+240)	((FXA(IC,IA),IC=101,150),IA=1,26)			00016400			
7	WRITE(7'RECNO+360)	((FXA(IC,IA),IC=151,200),IA=1,26)			00016500			
7	WRITE(8'RECNO) FPA				00016600			
7	OLDP=P				00016700			
7	DO 90 IA=1,26				00016800			
9	FPA(I)=0.D0				00016900			
9	DO 90 IC=1,NC				00017000			
11	FXA(IC,IA)=0.D0				00017100			
2	90 CONTINUE				00017200			
1	C				00017300			
1	C CHARACTERISTIC DEFINITIONS (UP TO 200)				00017400			
1	C				00017500			
2	100 CONTINUE				00017600			
7	CHAR(1)=P.EQ.0.AND.LFS.EQ.1				00017700			
7	CHAR(2)=P.EQ.1.AND.LFS.EQ.1				00017800			
7	CHAR(3)=P.EQ.2.AND.LFS.EQ.1				00017900			
7	CHAR(4)=P.EQ.3.AND.LFS.EQ.1				00018000			
7	CHAR(5)=P.EQ.4.AND.LFS.EQ.1				00018100			
7	CHAR(6)=P.EQ.5.AND.LFS.EQ.1				00018200			
7	CHAR(7)=P.EQ.6.AND.LFS.EQ.1				00018300			
7	CHAR(8)=P.EQ.7.AND.LFS.EQ.1				00018400			
7	CHAR(9)=P.EQ.8.AND.LFS.EQ.1				00018500			
7	CHAR(10)=P.EQ.9.AND.LFS.EQ.1				00018600			
7	CHAR(11)=P.EQ.0.AND.LFS.EQ.2				00018700			
7	CHAR(12)=P.EQ.1.AND.LFS.EQ.2				00018800			
7	CHAR(13)=P.EQ.2.AND.LFS.EQ.2				00018900			
7	CHAR(14)=P.EQ.3.AND.LFS.EQ.2				00019000			
7	CHAR(15)=P.EQ.4.AND.LFS.EQ.2				00019100			
7	CHAR(16)=P.EQ.5.AND.LFS.EQ.2				00019200			
7	CHAR(17)=P.EQ.6.AND.LFS.EQ.2				00019300			
7	CHAR(18)=P.EQ.7.AND.LFS.EQ.2				00019400			
7	CHAR(19)=P.EQ.8.AND.LFS.EQ.2				00019500			
7	CHAR(20)=P.EQ.9.AND.LFS.EQ.2				00019600			
1	C				00019700			
1	C GIVE USER-DEFINED HEIGHT WHICH IS USED FOR QUANTITATIVE				00019800			
1	C VARIABLES (OTHERWISE, IT IS SET EQUAL TO THE FINAL HEIGHT).				00019900			
1	C				00020000			
7	UH=FH				00020100			
7	UH=UH*C(MCNT)				00020200			
1	C				00020300			
1	C CALCULATE VARIOUS SUMS.				00020400			
1	C				00020500			
7	SPSHI=SPSHI+SM				00020600			
7	SPS=SPS+SM				00020700			
7	FPSHI(A)=FPSHI(A)+FH				00020800			
7	FPSA(A)=FPSA(A)+FH				00020900			
7	FPA(A)=FPA(A)+FH				00021000			
7	DO 110 IC=1,NC				00021100			
7	IF (.NOT.CHAR(IC)) GOTO 110				00021200			
9	X(IC)=X(IC)+UH				00021300			
9	FXSHI(IC)=FXSHI(IC)+UH				00021400			
9	FXS(IC)=FXS(IC)+UH				00021500			
9	FXA(IC,A)=FXA(IC,A)+UH				00021600			

START

COL	1	2	3	4	5	6	7	8
2	110	CONTINUE				00021700		
7		GOTO 10				00021800		
1	C					00021900		
1	C	END-OF-FILE				00022000		
1	C					00022100		
2	200	WRITE(9)MCNT,SURVEY,C,NC,X				00022200		
7		STOP				00022300		
1	C					00022400		
1	C	ERROR MESSAGES				00022500		
1	C					00022600		
2	930	PRINT 932,OLDYM,OLDU,U				00022700		
2	932	FORMAT('1'/1X,'*** ERROR MESSAGE ***'//1X,				00022800		
6		* 'THE INPUT FILE IS NOT SORTED BY SURVEY DATE',				00022900		
6		* ' AND RECODED PSU NUMBER.'//5X,'OLDYM=',I4,2X,				00023000		
6		* 'OLDU=',F12.0,2X,'U=',F12.0)				00023100		
7		STOP 10				00023200		
2	950	PRINT 952				00023300		
2	952	FORMAT('1'/1X,'*** ERROR MESSAGE ***'//1X,				00023400		
6		* 'SURVEY YEAR AND MONTH FOR WHICH ESTIMATES'/				00023500		
6		* ' 1X,' ARE REQUIRED IS NOT IDENTICAL TO THE'/				00023600		
6		* ' 1X,' SURVEY DATE PROVIDED'')				00023700		
7		STOP 10				00023800		
7		END				00023900		
1	*****					00024000		
1	**					*** 00024100		
1	/* IF STEP0 (FSPLIT) IS NOT USED, REMOVE '/*' FROM LINE 24800					*** 00024200		
1	/* AND GIVE INPUT FILE IN THAT LINE (CAUTION: THE INPUT FILE					*** 00024300		
1	/* MUST BE SORTED BY SURVEY MONTH AND RECODED PSU NUMBER).					*** 00024400		
1	/* THEN, INSERT '/*' IN FRONT OF LINE 24700.					*** 00024500		
1	/*					*** 00024600		
1	*****					00024700		
1	//GO.FT04F001 DD DSN=&SPLIT,DISP=(OLD,DELETE)					00024800		
1	///*//GO.FT04F001 DD DSN=RLFS.HLEE.MAR85.FEB86.DATA,DISP=OLD					00024900		
1	//GO.FT05F001 DD DSN=&TEMP1,UNIT=TAPEWK,DISP=(,PASS),					00025000		
1	// DCB=(RECFM=VBS,LRECL=1832,BLKSIZE=16000)					00025100		
1	//GO.FT06F001 DD DSN=&TEMP2,UNIT=WORK,DISP=(,PASS),					00025200		
1	// SPACE=(16233,(300,30)),DCB=(RECFM=VBS,LRECL=1828,BLKSIZE=6233)					00025300		
1	//GO.FT07F001 DD DSN=&TEMP3,UNIT=WORK,DISP=(,PASS),					00025400		
1	// SPACE=(10400,480)					00025500		
1	//GO.FT08F001 DD DSN=&TEMP4,UNIT=WORK,DISP=(,PASS),					00025600		
1	// SPACE=(1208,120)					00025700		
1	//GO.FT09F001 DD DSN=&TEMP5,UNIT=WORK,DISP=(,PASS),					00025800		
1	// SPACE=(6233,1),DCB=(RECFM=VBS,BLKSIZE=6233)					00025900		
1	/*					00026000		
1	//STEP2 EXEC SORT					00026100		
1	//SORTIN DD DSN=&TEMP1,DISP=(OLD,DELETE)					00026200		
1	//SORTOUT DD DSN=&TEMP1,UNIT=TAPEWK,DISP=(,PASS),DCB=*.SORTIN					00026300		
1	/*//SORTOUT DD DSN=&TEMP1,UNIT=WORK,DISP=(,PASS),DCB=*.SORTIN,					00026400		
1	/* SPACE=(6233,(100,10))					00026500		
1	//SYSIN DD *					00026600		
2	SORT FIELDS=(5,4,FI,A,9,4,FI,A,13,4,FI,A)					00026700		
1	/*					00026800		
1	//STEP3 EXEC SORT					00026900		
1	//SORTIN DD DSN=&TEMP2,DISP=(OLD,DELETE)					00027000		



START

COL	1	2	3	4	5	6	7	8
1	//SORTOUT DD DSN=&TEMP2,UNIT=WORK,DISP=(,PASS),DCB=*.	SORTIN,			00027100			
1	// SPACE=(6233,(600,601)				00027200			
1	//SYSIN DD *				00027300			
2	SORT FIELDS=(5,4,FI,A,9,4,FI,A)				00027400			
1	/*				00027500			
1	//STEP4 EXEC FORTGCLG,TIME.GO=10				00027600			
1	//FORT.SYSIN DD *				00027700			
7	REAL*8 X(200),FXSHI(200),SPSHI,FPSHIA(26),				00027800			
6	* FXS(12,200),SPS(12),FPSA(12,26),				00027900			
6	* FXA(12,200,26),FPA(12,26),DVAL(200)/200*0.D0/,				00028000			
6	* NAME1(200)/200*' /,				00028100			
6	* NAME2(200)/200*' /,SDVAL(4,6,200)/4800*0.D0/,				00028200			
6	* VAR(200)/200*0.D0/,SRCV(200)/200*0.D0/,				00028300			
6	* DSUM(200)/200*0.D0/,TV,U,OLDU,				00028400			
6	* SSUM(200)/200*0.D0/,ZX,ZP,RFA,C(12)				00028500			
7	INTEGER ID1,ID2,A,P,NC,MCNT,OLDID1,OLDP,SA,OLDSA,STR,OLDSTR,				00028600			
6	* NH//,NSRS(10)/10*0//,SNH(4)/4*0//,NMNTH,REONO,SURVEY(12)				00028700			
7	LOGICAL*1 FLAG//.FALSE./,RFLAG//.FALSE./,SRFLAG//.FALSE./,PSCODE//				00028800			
6	* .FALSE./,SCNTFG(4)/4*.FALSE./				00028900			
1	C				00029000			
1	C DEFINE CHARACTERISTIC NAMES.				00029100			
1	C				00029200			
7	DATA NAME1(1)/*NFLD */				00029300			
7	DATA NAME1(2)/*PEI */				00029400			
7	DATA NAME1(3)/*NS */				00029500			
7	DATA NAME1(4)/*NB */				00029600			
7	DATA NAME1(5)/*PQ */				00029700			
7	DATA NAME1(6)/*ONT */				00029800			
7	DATA NAME1(7)/*MAN */				00029900			
7	DATA NAME1(8)/*SAS */				00030000			
7	DATA NAME1(9)/*ALT */				00030100			
7	DATA NAME1(10)/*BC */				00030200			
7	DATA NAME2(1)/*EMP */				00030300			
7	DATA NAME2(2)/*EMP */				00030400			
7	DATA NAME2(3)/*EMP */				00030500			
7	DATA NAME2(4)/*EMP */				00030600			
7	DATA NAME2(5)/*EMP */				00030700			
7	DATA NAME2(6)/*EMP */				00030800			
7	DATA NAME2(7)/*EMP */				00030900			
7	DATA NAME2(8)/*EMP */				00031000			
7	DATA NAME2(9)/*EMP */				00031100			
7	DATA NAME2(10)/*EMP */				00031200			
7	DATA NAME1(11)/*NFLD */				00031300			
7	DATA NAME1(12)/*PEI */				00031400			
7	DATA NAME1(13)/*NS */				00031500			
7	DATA NAME1(14)/*NB */				00031600			
7	DATA NAME1(15)/*PQ */				00031700			
7	DATA NAME1(16)/*ONT */				00031800			
7	DATA NAME1(17)/*MAN */				00031900			
7	DATA NAME1(18)/*SAS */				00032000			
7	DATA NAME1(19)/*ALT */				00032100			
7	DATA NAME1(20)/*BC */				00032200			
7	DATA NAME2(11)/*UNEMP */				00032300			
7	DATA NAME2(12)/*UNEHP */				00032400			

START

COL	1	2	3	4	5	6	7	8
7	DATA NAME2(13)/*UNEMP	'/				00032500		
7	DATA NAME2(14)/*UNEMP	'/				00032600		
7	DATA NAME2(15)/*UNEMP	'/				00032700		
7	DATA NAME2(16)/*UNEMP	'/				00032800		
7	DATA NAME2(17)/*UNEMP	'/				00032900		
7	DATA NAME2(18)/*UNEMP	'/				00033000		
7	DATA NAME2(19)/*UNEMP	'/				00033100		
7	DATA NAME2(20)/*UNEMP	'/				00033200		
1	C					00033300		
1	C DIRECT ACCESS DATA SETS					00033400		
1	C					00033500		
7	DEFINE FILE 7(480,10400,L,K7)					00033600		
7	DEFINE FILE 8(120,208,L,K8)					00033700		
7	READ(9)NMNTH,SURVEY,C,NC,X					00033800		
7	DO 2 IM=1,NMNH					00033900		
9	READ(6)OLDID1,MNTH,(FXS(IM,IC),IC=1,200),SPS(IM),					00034000		
6	* (FPSA(IM,IA),IA=1,26)					00034100		
9	IF (IM.NE.MNTH) GOTO 900					00034200		
9	OLDP=OLDID1/10000					00034300		
9	RECNO=OLDP*12+IM					00034400		
9	READ(7'RECNO)(FXA(IM,IC,IA),IC=1,50),IA=1,26)					00034500		
9	READ(7'RECNO+120)(FXA(IM,IC,IA),IC=51,100),IA=1,26)					00034600		
9	READ(7'RECNO+240)(FXA(IM,IC,IA),IC=101,150),IA=1,26)					00034700		
9	READ(7'RECNO+360)(FXA(IM,IC,IA),IC=151,200),IA=1,26)					00034800		
9	READ(8'RECNO)(FPA(IM,IA),IA=1,26)					00034900		
2	2 CONTINUE					00035000		
2	10 READ(5,END=12)ID1,ID2,MCNT,FXSHI,SPSHI,FPSHIA					00035100		
7	STR=ID1*10000+ID2/10					00035200		
7	U=ID1*100000.D0+ID2					00035300		
7	SA=ID2/10000					00035400		
1	C					00035500		
1	C CHECK IF THE RECORD JUST READ IN IS THE FIRST RECORD.					00035600		
1	C					00035700		
7	IF (RFLAG) GOTO 20					00035800		
7	NH=1					00035900		
7	OLDU=U					00036000		
7	OLDSA=SA					00036100		
7	OLDSTR=STR					00036200		
7	RFLAG=.TRUE.					00036300		
7	GOTO 20					00036400		
1	C					00036500		
1	C END-OF-FILE					00036600		
1	C					00036700		
2	12 FLAG=.TRUE.					00036800		
7	GOTD 22					00036900		
1	C					00037000		
1	C CHECK FOR REPLICATE BREAK					00037100		
1	C					00037200		
2	20 IF (OLDU.EQ.U) GOTO 72					00037300		
2	22 IF (OLDSA.EQ.9) GOTO 26					00037400		
7	DO 24 IC=1,NC					00037500		
9	DSUM(IC)=DSUM(IC)+DVAL(IC)					00037600		
2	24 SSUM(IC)=SSUM(IC)+DVAL(IC)**2					00037700		
7	GOTO 30					00037800		

START

COL	1	2	3	4	5	6	7	8
2	26	IS=OLDSTR-(OLDSTR/100)*100			00037900			
7		PSCODE=.TRUE.			00038000			
7		DO 28 IC=1,NC			00038100			
2	28	SDVAL(IS,NH,IC)=SDVAL(IS,NH,IC)+DVAL(IC)			00038200			
2	30	IF (FLAG) GOTO 34			00038300			
7		DO 32 IC=1,NC			00038400			
2	32	DVAL(IC)=0.D0			00038500			
7		OLDU=U			00038600			
1	C				00038700			
1	C	CHECK FOR STRATUM BREAK			00038800			
1	C				00038900			
7		IF (OLDSTR.EQ.STR) GOTO 70			00039000			
1	C				00039100			
1	C	COMPUTE AND ADD STRATUM CONTRIBUTION TO THE TOTAL VARIANCE.			00039200			
1	C				00039300			
2	34	IF (OLDSA.NE.9) GOTO 36			00039400			
7		IS=OLDSTR-(OLDSTR/100)*100			00039500			
7		SCNTFG(IIS)=.TRUE.			00039600			
7		GOTO 44			00039700			
1	C				00039800			
1	C	THE PROGRAM CHECKS IF THE STRATUM HAS ONLY ONE REPLICATE.			00039900			
1	C	IF SO, IT SETS NH (NO. OF REPLICATES) VALUE EQUAL TO TWO.			00040000			
1	C	THIS ACTION CAUSES OVERESTIMATION OF VARIANCE.			00040100			
1	C	HENCE, A CAUTION SHOULD BE GIVEN WHEN THERE ARE MANY SUCH CASES.			00040200			
1	C	(THE NUMBER OF SINGLE REPPLICATE STRATUM WILL BE PRINTED AT THE			00040300			
1	C	END OF THE PROGRAM RUN).			00040400			
1	C				00040500			
2	36	IF (NH.NE.1) GOTO 40			00040600			
7		NH=2			00040700			
7		NSRS(OLDP+1)=NSRS(OLDP+1)+1			00040800			
7		SRFLAG=.TRUE.			00040900			
2	40	DO 42 IC=1,NC			00041000			
9		TV=(NH*SSUM(IC)-DSUM(IC)**2)/(NH-1)			00041100			
9		VAR(IC)=VAR(IC)+TV			00041200			
9		SSUM(IC)=0.D0			00041300			
9		DSUM(IC)=0.D0			00041400			
9		IF (.NOT.SRFLAG) GOTO 42			00041500			
9		SRCV(IC)=SRCV(IC)+TV			00041600			
2	42	CONTINUE			00041700			
7		SRFLAG=.FALSE.			00041800			
2	44	IF (FLAG) GOTO 46			00041900			
7		NH=0			00042000			
7		OLDSTR=STR			00042100			
7		OLDSA=SA			00042200			
1	C				00042300			
1	C	CHECK FOR ER/CMA BREAK			00042400			
1	C				00042500			
7		IF (OLDDID1.EQ.ID1) GOTO 70			00042600			
7		DO 45 IM=1,MNTH			00042700			
9		READ(6)ID1,MNTH,(FXS(IM,J),J=1,200),SPS(IM),			00042800			
6	*	(FPSA(IM,J),J=1,26)			00042900			
9		IF (IM.NE.MNTH) GOTO 900			00043000			
2	45	CONTINUE			00043100			
7		P=ID1/10000			00043200			

START

COL	1	2	3	4	5	6	7	8
7	OLDID1=ID1				00043300			
1	C				00043400			
1	C CHECK FOR PROVINCE BREAK.				00043500			
1	C				00043600			
7	IF (OLDP.EQ.P) GOTO 70				00043700			
1	C				00043800			
1	C COMPUTE AND ADD SPECIAL AREA CONTRIBUTION TO THE TOTAL VARIANCE.				00043900			
1	C				00044000			
2	46 IF (.NOT.PSCODE) GOTO 60				00044100			
7	DO 50 IS=1,4				00044200			
9	IF (SNH(IS).EQ.0) GOTO 50				00044300			
9	NH=SNH(IS)				00044400			
9	DO 52 IH=1,NH				00044500			
9	DO 52 IC=1,NC				00044600			
11	DSUM(IC)=DSUM(IC)+SDVAL(IS,IH,IC)				00044700			
11	SSUM(IC)=SSUM(IC)+SDVAL(IS,IH,IC)**2				00044800			
2	52 SDVAL(IS,IH,IC)=0.00				00044900			
9	IF (NH.NE.1) GOTO 54				00045000			
9	NH=2				00045100			
9	NSRS(OLDP+1)=NSRS(OLDP+1)+1				00045200			
9	SRFLAG=.TRUE.				00045300			
2	54 DO 56 IC=1,NC				00045400			
11	TV=(NH*SSUM(IC)-DSUM(IC)**2)/(NH-1)				00045500			
11	VAR(IC)=VAR(IC)+TV				00045600			
11	DSUM(IC)=0.00				00045700			
11	SSUM(IC)=0.00				00045800			
11	IF (.NOT.SRFLAG) GOTO 56				00045900			
11	SRCV(IC)=SRCV(IC)+TV				00046000			
2	56 CONTINUE				00046100			
9	SRFLAG=.FALSE.				00046200			
2	50 CONTINUE				00046300			
2	60 IF (IFLAG) GOTO 100				00046400			
7	PSCODE=.FALSE.				00046500			
7	NH=0				00046600			
7	OLDP=P				00046700			
7	DO 62 I=1,4				00046800			
9	SNH(I)=0				00046900			
2	62 SCNTFG(I)=.FALSE.				00047000			
7	DO 64 IM=1,NMNTN				00047100			
9	RECNO=OLDP*12+IM				00047200			
9	READ(7'RECNO)(FXA(IM,IC,IA),IC=1,50),IA=1,26)				00047300			
9	READ(7'RECNO+120)(FXA(IM,IC,IA),IC=51,100),IA=1,26)				00047400			
9	READ(7'RECNO+240)(FXA(IM,IC,IA),IC=101,150),IA=1,26)				00047500			
9	READ(7'RECNO+360)(FXA(IM,IC,IA),IC=151,200),IA=1,26)				00047600			
9	READ(8'RECNO)(FPA(IM,IA),IA=1,26)				00047700			
2	64 CONTINUE				00047800			
2	70 NH=NH+1				00047900			
7	IF (OLDSA.NE.9) GOTO 72				00048000			
7	IS=OLDSTR-(OLDSTR/100)*100				00048100			
7	IF (SCNTFG(IS)) GOTO 72				00048200			
7	SNH(IS)=SNH(IS)+1				00048300			
1	C				00048400			
1	C COMPUTE D-VALUE.				00048500			
1	C				00048600			

DATASET: RLFS.HLEE.GENERAL.LFSVAR.CNTI

DATE: 87/07/17
TIME: 13:25
PAGE: 10

START

```

COL ---+---1---+---2---+---3---+---4---+---5---+---6---+---7---+---8---

2   72   DO 76 IC=1,NC
9     ZX=FXSHI(IC)-SPSHI*FXS(MCNT,IC)/SPS(MCNT)
9     ZP=0.D0
9     DO 74 IA=1,26
11    IF (FPA(MCNT,IA).EQ.0) GOTO 74
11    RFA=FXA(MCNT,IC,IA)/FPA(MCNT,IA)
11    ZP=ZP+RFA*(FPSHIA(IA)-SPSHI*FPSA(MCNT,IA)/SPS(MCNT))
2   74   CONTINUE
9     DVAL(IC)=DVAL(IC)+ZX-ZP
2   76   CONTINUE
7     GOTO 10

1 C
1 C   PRINT THE RESULTS.
1 C

2   100  PRINT 102,(SURVEY(I),I=1,NMNT)
2   102  FORMAT('1'//1X,'SURVEY MONTHS:',12I6)
7     PRINT 104,(C(I),I=1,NMNT)
2   104  FORMAT('0','COEFFICIENTS:',1X,12F6.2)
7     PRINT 110,(NSRS(I),I=1,10)
2   110  FORMAT('1'//1X,'THE NUMBER OF SINGLE REPLICATE',
6     * ' STRATA FOR EACH PROVINCE IS GIVEN BELOW.'//1X,
6     * 'CAUTION: VARIANCE COULD BE SERIOUSLY OVERESTIMATED',
6     * ' IF THESE NUMBERS ARE LARGE'//T5,'NFLD',T10,
6     * 'PEI',T15,'NS',T20,'NB',T25,'PQ',T30,'ONT',T35,'MAN',
6     * T40,'SAS',T45,'ALT',T50,'BC'//T5,10(I2,3X))
7     PRINT 120
2   120  FORMAT(////1X,'CHARACTERISTICS',5X,'ESTIMATE',6X,
6     * 'VARIANCE',4X,'CV',3X,'SINGLE REP. ST. CONTRIB.'//
6     * 18X,'(IN THOUSAND)',2X,'(IN MILLION)',2X,'(%)',7X,
6     * '(IN MILLION)'//)
7     DO 130 IC=1,NC
9       X(IC)=X(IC)/1000
9       VAR(IC)=VAR(IC)/1000000
9       SRCV(IC)=SRCV(IC)/1000000
9       CV=100*DSQRT(VAR(IC))/X(IC)
9       PRINT 132, NAME1(IC),NAME2(IC),X(IC),VAR(IC),CV,SRCV(IC)
2   132  FORMAT(1X,2A8,3X,F9.2,4X,F9.2,3X,F6.2,10X,F9.4)
2   130  CONTINUE
7     STOP

1 C
1 C   ERROR MESSAGE
1 C

2   900  PRINT 910,IM
2   910  FORMAT('1'//1X,'*** ERROR MESSAGE ***'//1X,
6     * 'MONTH ',I2,' DATA IS MISSING FROM FILE 6 OR 7.')
7     STOP 10
7     END

1 //GO.FT05F001 DD DSN=&&TEMP1,DISP=(OLD,DELETE)
1 //GO.FT06F001 DD DSN=&&TEMP2,DISP=(OLD,DELETE)
1 //GO.FT07F001 DD DSN=&&TEMP3,DISP=(OLD,DELETE)
1 //GO.FT08F001 DD DSN=&&TEMP4,DISP=(OLD,DELETE)
1 //GO.FT09F001 DD DSN=&&TEMP5,DISP=(OLD,DELETE)
1 //

```

STATISTICS CANADA LIBRARY
BIBLIOTHÈQUE STATISTIQUE CANADA



1010179807

OOS

