SURVEY MAINTENANCE - PHILOSOPHY AND PRACTICE[1]

F. Mayda and P. Timmons[2]

An aspect of surveys not always given adequate consideration is maintenance.

The scope and importance of survey maintenance are discussed and a case is made for a more scientific methodological approach. Practical applications to various stages of surveys are illustrated by examples from the Canadian Labour Force Survey.

1. INTRODUCTION

Survey maintenance is an indispensable part of any continuing survey; however, its components are usually treated as separate activities rather than as an overall program. This ad hoc approach can result in gaps in the program, inadequate documentation or dissemination of results, inefficiency, and lack of funding due to inadequate understanding of the problems addressed by maintenance. The recognition of survey maintenance as a distinct methodological domain can encourage the application of a more scientific approach. One such example is the cost-benefit approach to controls in surveys of Platek and Singh [1]. This paper will deal with a philosophy of sample maintenance and will also present some illustrations of various aspects of its application in a large scale continuous survey.

For the purposes of this paper, survey maintenance can be considered

---

1 This paper was presented at the Canadian Conference in Applied Statistics (1981) at Concordia University, Montreal, April 29, 1981.

2 F. Mayda and P. Timmons, Census and Household Survey Methods Division, Statistics Canada.

as the sum total of the activities and programs, both regular and occasional:

a) which ensure that the survey design is respected in all of the operations of the survey,

b) which measure the quality of operations and of survey data,

c) which modify or adapt the survey process to meet changing requirements.

The maintenance aspect of the design and conduct of large scale continuing surveys seldom receives sufficient recognition. There can be many reasons for this. Theoreticians, anxious to break new ground and develop new, and in some sense better, mathematical approaches, find the concept of maintenance unglamorous and mundane. Managers, concerned with competing priorities, budgets and production, often fail to see the relationship of maintenance to operational productivity and question the need for these expenditures. The very word "maintenance" has the connotation of "just getting by" or "avoiding deterioration" and seems to imply "no change" for many people.

In spite of this lack of recognition, maintenance is a most necessary part of continuing surveys for many good reasons. The requirements of the survey may change, the conditions in the population being surveyed or the sample frame itself may change, there may be changes in policy or budget, new techniques or equipment may become available. Adapting to these changes and ensuring that quality and efficiency are not compromised is a regular part of the maintenance of a survey. Even if such obvious changes do not occur, such things as the turnover of all levels of staff, the passage of time since principles and procedures were first learned and the gradual separation of the designers and developers from the operations staff can lead to a dilution of experience and the possible degeneration of quality.

The purpose of this paper is neither to try to glamourize survey maintenance nor to develop some all-encompassing theoretical approach, model or package which can be applied to any survey. Rather, the purpose of this paper is to spotlight the role of survey maintenance, to bring to the fore and emphasize how that role applies to large scale continuing surveys and, by demonstrating its function, to encourage a more scientific and theoretical study to be brought to bear on the subject. In order to do this, it must be realized that survey maintenance is more a philosophy than a procedure. Individual programs must be tailored to the needs of the survey.

Most of the discussions on maintenance and the specific examples, which follow, although drawn from the Canadian Labour Force Survey, are relevant for many large scale continuous surveys.

2. THE MAINTENANCE PHILOSOPHY

Whenever large scale continuing surveys are planned and developed, significant effort and resources are devoted to implementing the best features that the available money can buy. This ranges from the original sample frame through data collection procedures to final estimation and data dissemination. Once the survey has become operational, especially in the case of surveys used to gather official government statistics, there is a need continually to ensure and to demonstrate the quality of the data and the efficiency of the survey methods and procedures. This ability to demonstrate the validity of the survey is required to allow for quality certification of data, to withstand criticism, to assist the organization in performing program quality audits and to encourage the development of quality improvement programs.

Perhaps the most important feature of this philosophy is to maintain programs which continue in an organized fashion to question operations, procedures and survey materials in order to verify their adequacy.

The anomaly of a good maintenance program is that the more effective it is in maintaining high quality in the survey process, the less recognition it may receive as a necessary program. A simple hypothetical example could be cited. Suppose in a personal interview survey the interviewers and supervisory staff all know that there is a continuous and prescribed program of reinterview. This mere presence of the reinterview program may result in a better standard of data collection. The more effective this program is, however, the less dramatic will be the results of the reinterview. When faced with the requirement to reduce costs it is very tempting for managers to assume that the interview process is properly conducted, as evidenced by the good reinterview reports, and therefore to cut back on the reinterview program.

The philosophy of a unified survey maintenance program approach implies a broad scope. Survey maintenance touches on every facet of a survey from the initial planning to the final dissemination of the survey data.

Many steps or stages can be identified in the process of a survey according to the degree of detail one wishes. For convenience, we will broadly divide the survey process into the following five stages:

a) Survey Planning and Design
b) Sample Selection and Control
c) Data Collection
d) Data Capture and Processing
e) Estimation and Dissemination

All of the above can be recognized as common to any large scale survey. When the survey is continuous the planning and design stages are frequently replaced by periodic improvements and occasional redesigns and revisions.

3. ASPECTS OF MAINTENANCE

Various maintenance programs, to be discussed in detail later, are operative at each of these five stages of a survey. These programs can be classified as Measuring, Controlling and Adapting.

The distinction as to which aspect of maintenance a particular maintenance program falls under is not essential. This is so because often a quality measure, for example, can be used both for diagnostic purposes and as feed-back to operations. What is important is the recognition of the necessity of these aspects in a maintenance program.

Measuring

Maintenance programs, classified as Measuring, provide certain indicators of performance at various stages of the survey. The measures may be used as a guide to operational control or by data analysts to improve their insight into the reliability of the data and its suitability for particular purposes.

Measurement programs can be identified according to their use:

- regulatory: those which serve to measure the conduct of specific aspects of the survey operations.
- diagnostic: those which measure how well the survey functions in relation to the survey output.
- metadata: measurements of aspects of the survey data used by analysts and managers to evaluate the data itself.

It is understood, of course, that the same measurement may serve more than one of these purposes.

Controlling

Maintenance programs used in controlling provide measures of survey performance for comparison against standards to identify aspects requiring correction. This implies a feed-back mechanism which will adjust operations to ensure that these standards are met.

Adapting

These programs are essentially means of coping with change, whether due to changes of objectives or conditions or to the availability of new methods or equipment.

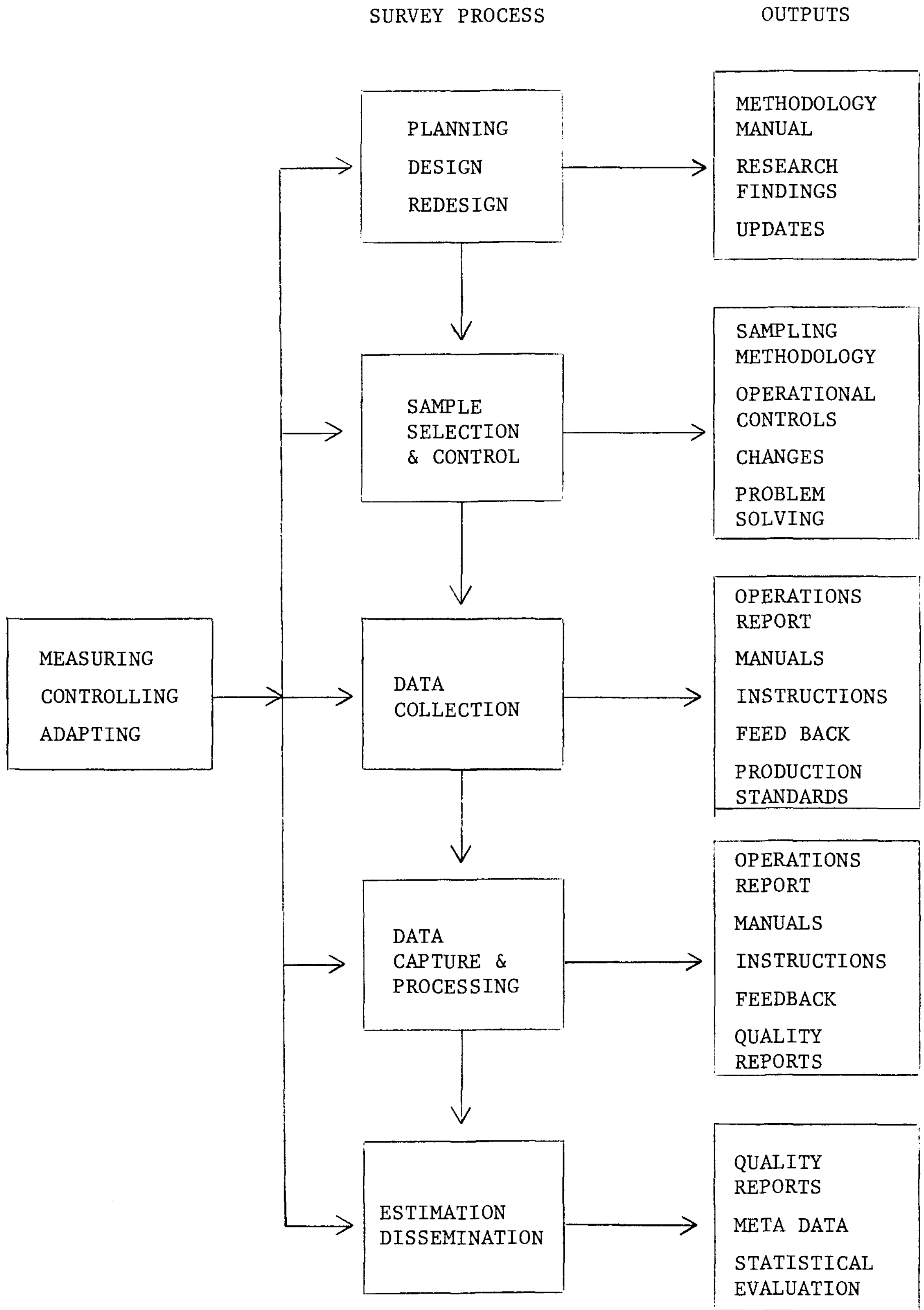
Examples of maintenance activity in the Canadian L.F.S. for the purpose of adapting to changed conditions or requirements:

- Parallel run. The running of two surveys in parallel, the old design and the new, to explain any differences and to link up the old and new time series.
- Sample size increase. Differential increase in sample size to improve provincial estimates.
- Stabilization program. To maintain a stable sample size while allowing for natural population growth.
- Sample Update. Partial redesign of new sample units to account for unequal growth.
- Sub-sampling in growth clusters. Reducing the interviewing burden on Field operations while allowing for growth.

Data from Maintenance Programs

Data from maintenance programs, whether in the form of quality measures, feed-back for remedial action or methods review and evaluation can be considered as outputs from the survey process. The following schematic diagram, although by no means complete, serves to illustrate how the maintenance program can influence all the stages of the survey process.

SURVEY MAINTENANCE



4. MAINTENANCE PROGRAMS AND THE SURVEY PROCESS

In order to control and minimize the impact of errors in large scale surveys, numerous quality control and evaluation programs are traditionally used. Most of these are familiar to survey methodologists and variations of them appear in almost any ongoing survey. The more obvious programs include tabulation and evaluation of non-response, undercoverage, cost of enumeration, observation, reinterview, variance/covariance, error rates and so on.

What we would like to do is demonstrate how survey maintenance impacts on every stage of the survey process. Examples are drawn from the Canadian Labour Force Survey.

a. Planning and Design

Obviously if a totally new survey is being planned and designed, there can be no maintenance program which affects the exercise directly. However, maintenance programs play two roles in the planning and design of a continuous survey. The first is in the fact that survey designers will draw on their own previous experience and that of others in the area of survey maintenance in order to evaluate possible features of the survey design. For example, the type of frame chosen will depend not only on what is available but also on what experience the designer has been exposed to in regard to frame maintenance.

The second way in which maintenance programs influence survey design is in periodic re-designs or programs of survey updating. For example, the LFS is normally redesigned every 10 years, shortly after the Decennial Census. The re-design after the 1971 Census was particularly extensive, incorporating many changes which were suggested based on the experience of maintenance on the older survey. Details of the changes made are elaborated in [2]. It is expected that more changes, based on maintenance experience with the current survey, will be introduced in the 1981 Redesign.

No on-going survey regardless of how adequate its initial design was, can remain without change for an extended period of time withot some deterioration. Populations of study change, concepts and objectives are modified, parameters which govern the sample selection become out of date and new procedures and technologies are developed. To prevent deterioration in the level of reliability of survey output, the survey maintenance function must evaluate these new factors constantly and implement required changes.

A specific example of this concerns the up-to-dateness of the LFS sample frame in large cities (SRU areas). After the 1976 Census it became possible to identify population growth in SRU areas from 1971 to 1976. The effect of this growth, which was not uniform even within individual SRU areas, was that size measures used in unequal probability of selection of sampling units became out of date, resulting in increased sampling variances. The changes made to the survey in the 1971 redesign allowed the development of methods to update the design in the SRU areas [3]. By using the information provided from the population comparisons, a special program was introduced to redesign specified sub-units within the SRU [4]. The impact of the re-specification of size measures can be seen in table 1. The effect of the program is to avoid increases in sampling variability due to highly clustered growth. This is particularly significant for estimates at the Census Metropolitan Area level.

TABLE 1

Increase in number of Random Groups due to SRU Update

December 1977 to March 1981

Province	No. of sub-units up-dated	Resulting no. of sub-units	Original no. of groups	New no. of groups
NFLD	4	4	42	60
PEI	5	5	90	112
NS	15	12	102	128
NB	12	13	144	214
QUE	26	32	162	252
ONT	46	53	300	402
MN	9	8	96	154
SASK	6	8	108	194
ALTA	23	39	276	518
BC	25	27	174	250
CANADA	170	204	1494	2284

Note: a sub-unit is a contiguous area stratum within a Self-Representing area comprising a number of Random Groups. A Random Group is a random collection of clusters (usually city blocks). The total number of sub-units in the initial design was 734.

b. <u>Sample Selection and Control</u>

Numerous activities are involved in the selection and control of a sample. In the case of a continuing survey, these relate to the maintenance of the sample frame and the selection and rotation of sample units at various stages.

In the LFS the second last stage of selection is a small well-defined area called a cluster. All of the dwellings located in the cluster are identified and listed in the field. The list is stored on a computerized data base in Ottawa and the final sample consists of a systematic sample of dwellings drawn by computer from the clusters. Comparisons of the expected number of dwellings based on the count when designing the area, to those actually listed frequently show significant differences. Most often the differences are due to construction or demolition of dwellings since the time that the cluster was first defined. In a number of cases, however, differences were due to incorrect listing or boundary errors. Such errors result in under or over sampling. To minimize the possibility of errors in listing clusters, a special program called "Cluster Yield Monitoring" was established.

Each month, Regional Offices are asked to identify reasons for significant differences between the design count of dwellings and the actual number listed for all newly introduced clusters. The timing of the program is such that field or design errors can frequently be corrected before interviewing in selected dwellings has begun. The following table illustrates some results of the program.

TABLE 2

Cluster Yield Monitoring Program
Number of Exception Clusters Checked (October 1979-December 1980)

Type of discrepancy	No. of Clusters
Valid differences	968
No correction necessary or no correction possible	244
Correctable errors	62
Not determined	24
Total Exceptions	1298

This evaluation is based on a total of 11804 clusters entering the active sample during the period.

Another example of maintenance in the Sample Selection and Control stage is stabilization of the sample size. Because of the self-weighting feature of the LFS design and the fact that the Canadian population continues to grow, the sample size would normally continue to grow at the same rate. As a means of holding down survey costs an automated procedure known as Sample Size Stabilization has been developed to keep the sample size from growing [5]. Each month for a specific rotation group and type of area within a province, the number of dwellings selected is compared to a predetermined base figure. Should the number selected be less than or equal to the base, nothing further is done. However, should the number selected exceed the base then the excess of dwellings is systematically dropped from the set of selections. A compensating weight is calculated and applied to all the non-dropped dwellings.

Fluctuations in sample size, due to sampling variability among clusters and unequal growth rates, introduce slight changes in the actual number of dwellings selected each month. However, due to the stabilization program, the net sample size remains fairly stable. The following table illustrates the net sample reductions per month. It can be seen that although the net decrease due to stabilization varies somewhat from month to month, there is an increasing reduction through time compensating for the natural growth in the sample.

TABLE 3

SAMPLE SIZE STABILIZATION: OCTOBER 1979 TO MARCH 1981

SURVEY DATE	NUMBER OF DWELLINGS DROPPED
1079	521
1179	544
1279	519
0180	610
0280	544
0380	598
0480	643
0580	667
0680	548
0780	740
0880	677
0980	693
1080	745
1180	868
1280	755
0181	861
0281	847
0381	897

At the current rate this amounts to a direct saving in interviewing costs of around $4,500 per month. There are also additional savings due to reduced processing and hiring and training of additional interviewers.

c. <u>Data Collection</u>

This phase of the survey process encompasses all collection activities and the materials used. In the case of the LFS, data is collected by personal and telephone interviews by a large staff of highly trained interviewers. The forms used by interviewers are preprinted for specific households and often the second and subsequent interview show certain data reproduced from the month before. There are many opportunities in this stage of the survey process for the effective use of maintenance programs to maintain quality and to improve methods and procedures. For example, tabulation and examination of edit changes can lead to improvements in training, changes in questionnaire design or changes in edit rules depending on the results of such analyses. This operation, in the LFS, is called the Field Edit Module and is maintained on a monthly basis.

Other significant modifications can derive from re-interview, observation and cost monitoring programs. It is essential to make results from such programs visible so that their importance can be recognized in order to ensure their continued support.

A phenomenon in large scale probability surveys is the problem of under-coverage. In the LFS the extent to which the survey underrepresents the population is called slippage. Slippage is the accumulated result of many things such as errors in clusters or cluster lists, missed dwellings, missed persons within dwellings, errors in coding and inaccurate population estimates to which the survey estimates are compared.

The continual monitoring of slippage is part of the maintenance program of the LFS. A significant change in the slippage rate triggers remedial action, for example: special list checks or special interviewer instructions.

Another problem in Data Collection is non-response. Continuous monitoring of response rates has shown a consistent trend toward higher non-response due to higher no one at home and temporary absent category during the summer months. In an effort to improve response a procedure known as "Post Survey Week Follow Up" has been developed [6]. In essence this is a special procedure of contacting, mostly by telephone, as many non-responding dwellings as possible one or two days after the normal survey period. Due to considerations such as timeliness, recall length and cost, the procedure is only used within specific restrictions and essentially during the summer months. On occasion, the procedure is also permitted where there are special situations where non-response is expected to be exceptionally high. An example of the sort of improvement that is obtained is shown in the following, Table 4.

TABLE 4

POST SURVEY WEEK FOLLOW-UP OTTAWA R. O.

JULY 1978

Type of Non-Response	Number			Reduction in Non-response Rate (%)
	At end of Survey Week	Followed Up	Successful Follow-Up	
T	160	117	43	1.61
N	46	36	11	0.47
K	4	2	2	0.04
TOTAL TOTAL	210	155	55	2.12

T = The household was temporarily absent for the entire week.

N = The occupants could not be contacted after several attempts.

K = Circumstances within the household, e.g. sickness, language problems.

d. Data Capture and Processing

The next step of the survey process consists of data capture and processing. In any large scale survey, the survey data are transformed into machine readable form; the data are edited and coded and imputations are made. In the LFS, in order to ensure the accuracy of the data capture, a quality control program using complete and sample verification is maintained. The program is designed to ensure that data entry errors do not exceed 3%. Continued monitoring of the program results in very high levels of data capture accuracy and ensures the efficiency of data entry operators.

Even when high levels of data capture accuracy are maintained, the data are still subject to errors which may have been introduced during the field collection. These would represent enumerator or respondent errors and they are detected in the edit process. In the LFS there is a program to identify all data fields where changes have been made during editing. The program is called the Field Edit Module and, as has been mentioned earlier, is used as a feed-back to interviewers, questionnaire designers and editors.

The FEM does not include all errors that might have been made but only those where the data entered (or omitted) causes an edit failure. Nevertheless the FEM results have been shown to be a very good measure of the relative number of errors made. The error rates generated by the FEM are a sensitive indicator of the quality of interviewers' work and also a measure of the awareness of field staff of survey requirements. Since the implementation of the program, there has been steady improvement in the error rates to a currently stable level. Even this stable level shows some improvement, however, each time efforts are made to emphasize accuracy in completing the questionnaire. For example, each time special training sessions for field staff focus on improving accuracy, there is a corresponding improvement in error rates for several occasions immediately after the sessions.

In addition to the FEM analysis of edit failures, the editing section monitors the number of error-containing records each month. Results of this monitoring are regularly discussed with field staff to keep them aware of the need to minimize these errors. In cases where error rates show abnormal increase, there is feed-back to individual Regional Offices with identification of the specific types of errors and suggestions for eliminating them in future.

Table 5 shows how the maintenance of this program has contributed to reducing the error rates over time for the household record (F03) and individual questionnaire (F05).

TABLE 5

EDIT ERROR RATES BY FORM TYPE

1977 - 1980

F03	JAN.	FEB.	MAR.	APR.	MAY	JUNE	JULY	AUG.	SEP.	OCT.	NOV.	DEC
1977	5.8	4.7	4.8	5.2	5.1	5.3	5.3	5.4	5.4	5.1	4.9	4.6
1978	4.8	4.2	4.0	3.4	3.6	3.3	3.1	3.0	3.1	2.9	2.8	2.9
1979	2.3	2.3	2.3	2.6	2.3	2.6	2.3	2.5	2.3	1.9	2.1	2.2
1980	1.8	1.9	1.8	1.9	1.9	1.9	2.1	1.8	N/A	1.9	1.8	1.9

F05	JAN.	FEB.	MAR.	APR.	MAY	JUNE	JULY	AUG.	SEP.	OCT.	NOV.	DEC.
1977	20.3	18.9	17.7	17.7	18.0	15.6	16.7	15.6	15.4	16.4	15.2	13.6
1978	14.4	13.8	13.5	14.5	14.3	11.9	12.6	11.6	11.3	12.7	11.1	9.8
1979	10.3	9.7	9.8	7.7	9.9	9.7	9.6	8.9	8.4	7.5	8.5	8.4
1980	7.7	7.4	7.1	8.6	8.0	8.0	8.1	7.1	N/A	8.2	7.6	6.4

e. Estimation and Dissemination

In this final phase of the survey process, the estimates are published and distributed to the various users. In a sense the data leave the hands of the survey methodologists and enter the domain of the data analysts and policy makers.

Survey data, especially those from large scale continuing surveys, are usually collected for two reasons. The first is to document and chart the phenomenon of interest. Thus, in the LFS, the survey data serve to provide a comprehensive summary statement about the labour force activity of the Canadian population. The second reason is for policy makers to combine data from various sources to evaluate existing social policy, to predict trends and to devise new policy aimed at improving the social situation.

It is not immediately obvious how maintenance programs can affect this phase of the survey process. Estimation procedures are usually fixed in that they depend on the probability design of the survey. Changes are only made if the probabilities of selection change. We have an example in the LFS. As mentioned earlier, in an effort to put a limit to natural sample size growth, a procedure of "stabilization" was implemented. The effect was the requirement to add a special weight to compensate for the sample reduction. This weight was incorporated into the estimation process.

Other changes too are being incorporated which must be considered part of the regular maintenance function. A program is currently under way to expand from two-digit to three-digit occupational codes to respond to requests from users for more detail. Such a change will not be implemented without careful assessment of the impact on editing and coding operations, data processing, tabulation and printing and finally estimation of data reliability.

By carefully monitoring the data processing operations, it has been possible to improve head office processing schedules to such a degree that the press release date for LFS data has been advanced. It has moved forward four days from the Tuesday at the start of the third week after survey week to the Friday at the end of the second week.

Perhaps the most neglected aspect of survey maintenance and quality evaluation in general is its impact on the uses to which data is put after publication. Too frequently implicit assumptions are made to the effect that the published data is "true" without considering outside factors which should temper any analysis. Most data analysts and users recognize the existence and significance of sampling errors, but may be less appreciative of the uncertainties in the data caused by non-sampling errors and non-respose (missing data).

Maintenance programs are effective in reducing not-sampling errors and non-response (missing data). They also provide valuable information (metadata) which should be considered by the analyst in using the data. Difficulties in data estimation and evaluation caused by non-sampling errors and non-response are most difficult to deal with. They must then be controlled by programs of prevention and this in addition to maintaining operations is the purpose of maintenance programs.

5. CONCLUSION

This has been a rather short and not very detailed overview of some of the maintenance programs of the LFS. Particular attention has been given to some of the less well known programs in an effort to demonstrate how fundamental they can be and how they form a part of overall survey maintenance.

We hope the foregoing makes the case for the pervasiveness and importance of survey maintenance and its contribution to better and more useable statistics. We recognize that there are, however, substantial costs involved and methods need to be developed to produce dynamic indicators similar to cost-variance studies used in sample design. A start in this direction has been made by Platek and Singh [1].

The relative value and cost of the various procedures should be used to control the scope, incidence and intensity of the components of the survey maintenance program.

RESUME

La coordination est un aspect des enquêtes auquel on n'accorde pas toujours suffisamment d'attention.

Les auteurs analysent l'envergure et l'importance de la coordination des enquêtes et préconisent une approche méthodologique plus scientifique. Des applications pratiques à diverses étapes des enquêtes sont illustrées par des exemples tirés de l'Enquête sur la population active du Canada.

REFERENCES

[1] Platek, R. and Singh, M.P., (1980), "Cost Benefit Analysis of Controls in Surveys", Symposium on Survey Sampling, Ottawa.

[2] Statistics Canada (1977), "Methodology of the Canadian Labour Force Survey", Catalogue No. 71-526.

[3] Platek, R. and Singh, M.P., (1978), "A Strategy for Updating Continuous Surveys", Metrika, Vol. 25, pp 1-7.

[4] Drew, D., Choudhry, H. and Gray, G., (1978), "Some Methods for Updating Survey Frames and Their Effects on Estimation", Proceedings of the American Statistical Association, pp 62-71.

[5] Drew, D. (1977), "LFS Sample Size Stabilization", Technical Memorandum, Census and Household Survey Methods Division, Statistics Canada.

[6] Gower, A. (1979) "Non-Response in the Candian Labour Force Survey", Survey Methodology, Statistics Canada, Vol. 5, No. 1.