

# A STRATEGY FOR THE PROVISION OF FRAME DATA AND USE OF TAX DATA FOR ECONOMIC SURVEYS

Michael Colledge and Robert Lussier, Statistics Canada

A new strategy has been developed at Statistics Canada, firstly, to provide centrally organized, more coherent frames for all major economic surveys and, secondly, to complement data collection by consistent and effective use of income tax data. This paper describes the fundamental elements of the strategy and outlines the plan for implementation.

The new 'Central Frame Data Base' will comprise two components. The first component will provide coverage of all large, complex, or otherwise significant economic operations. It will involve several types of records, corresponding in particular to legal entities, statistical units and survey reporting units, linked into enterprise structures. The second component will provide coverage of the remainder of the economic universe. It will comprise two sets of records based upon two independent sources. Each set will involve essentially only a single type of record, i.e. legal, statistical and reporting units will generally be presumed to coincide.

Income tax data will supplement annual survey data collection for small units and will help validate the coverage of the enterprise structures.

## INTRODUCTION

The objective of this paper is to describe the general strategy and plan for a major Statistics Canada redevelopment, namely the Business Survey Redesign Project. The project includes the complete redesign of systems for the provision of frame data and for the use of income tax data in economic surveys. It will affect virtually all the major business and financial surveys in the Agency. The project is scheduled to last three years and to involve about 100 people on a full time basis throughout this period. Completion will be by March 1988.

A paper presented to the 1983 ASA Annual Meetings(1) described the problems encountered in the creation and maintenance of frame data at Statistics Canada. These problems are briefly reviewed in Section 2. Section 3 contains a description of the objectives of the project and a summary of the strategy developed to meet these objectives. Sections 4 to 7 present its key elements, namely strategies for provision of frame data, for acquisition and use of income tax data to supplement or replace survey data, for design of annual surveys of production, and finally for design of sub-annual surveys of production. Organizational arrangements to support the strategy are described in Section 8 while implementation notes are presented in Section 9.

## PROBLEMS IN THE CREATION AND MAINTENANCE OF FRAME DATA

### Design Considerations

In the design of economic surveys, there are a number of significant factors which have to be taken into account, as illustrated in the following paragraphs.

- (a) The target population is often difficult to define precisely, for example whether a business should be identified as a legal entity or a site of economic production?
- (b) The items for which data are to be collected are numerous and diverse - production, finance, taxation, capital stock and expenditure, employment, payroll, etc. Also, there are subtle differences of meaning between economic, accounting and statistical concepts. Examples are the definition of "wholesale" trade and the distinction between "sales", "shipments", "revenue" and "income".
- (c) The target population and data requirements cannot be fully specified without practical consideration of the availability and ease of extraction of the data from respondents' accounting records. The "profiling" of a large enterprise, that is the delineation of the set of statistical units within the enterprise for which the data are to be collected and the arrangements by which such data may actually be acquired, is a complex operation.
- (d) The target population (however defined) tends to be rather large. There are of the order of 0.5 to 2.0 million units in Canada, depending upon the particular unit being considered. Furthermore the units tend to be very heterogeneous in size. For example the largest 12% of corporations account for 82% of corporate gross revenue.
- (e) Classification of data by industry though sometimes difficult to establish and maintain is essential for most purposes. Regular revisions to the classification scheme are necessary as a result of evolution of the economic world such as the introduction of new industries and the quasi-disappearance of others. Conversion to a new classification scheme is a major undertaking. Not all Statistics Canada surveys have adopted the latest Standard Industrial Classification (SIC)(2) which was developed in 1980.
- (f) The requirements for classification of data by geographic location vary, according to users, from all-Canada totals to small area breakdowns. As an illustration, the Canadian System of National Accounts requires national and provincial figures whereas regional development programs require major urban area or even county data. This results in designs which are compromises.
- (g) Response burden is another important feature. Businesses get requests for data from many agencies on various subjects. Response rates of 40% or less to an initial survey mailout are not unusual. Respondent costs as well as data collection costs are factors which must be taken into account in the design of a survey and suggest that maximal use should be made of data collected for administrative purposes, in particular income tax data.

The conclusion to be drawn from factors (a) -

(g) are that well defined systems for delineation and classification of survey frames and for blending administrative and survey response data are essential prerequisites for a successful economic statistics program. In the opinion of the authors, the complexity of economic survey design lies more in the clear definition of concepts and standards, in the creation and maintenance of frames, and in the use of administrative data, than in the sampling and estimation schemes.

#### Current Systems for Providing Frame and Income Tax Data at Statistics Canada

In general, lists of statistical units for economic surveys do not exist a priori; they have to be built. Administrative data to which Statistics Canada has access are a vital component of the construction process. In particular, substantial use can be made of personal income tax (T1), corporate income tax (T2) and employer payroll deduction (PD) files available from Revenue Canada. Unfortunately no single administrative file has all the required characteristics of coverage, currency and quality and the various files are not linked. Tax files do not provide coverage of government departments, hospitals, schools, etc., and are not available until well after the period to which they refer. PD files contain more current information but do not cover non-employers. As a consequence the system presently providing frame data is split into several parts. The principal component is the Business Register Master File (BRMF) based essentially on employer PD data. There are separate systems containing income tax (T1 and T2) and intercorporate ownership data collected under the Corporations and Labour Unions Returns Act (CALURA). In combination the BRMF, T1 and T2 tax and also the intercorporate ownership data support directly or indirectly the frames for most major economic surveys. Various survey operations also maintain separate, survey specific, systems to assist in frame maintenance.

Some of the problems with this arrangement are, firstly, that coordination between the BRMF and the survey operations is not sufficiently regulated to ensure complete exchange of frame information and absence of duplicated effort. Secondly, coverage is not fully controlled by the BRMF. Thirdly, there is the impression that a portion of the resources currently spent on the maintenance of smaller units should be devoted to the larger ones. Finally, the quality of BRMF classification data leads one to question the maintenance procedures. A recent study(3) has confirmed, for example, that 21% of the BRMF units are misclassified by industry. In short the BR has not yet succeeded in fulfilling its mandate(5) of providing complete up-to-date frames for all economic surveys.

The procedures for sampling, acquisition, processing and use of income tax data to supplement or replace survey response data have developed on an experimental basis over the past decade in the Tax Record Access (TRA) program but they have never been fully rationalized nor evaluated. Annual surveys of production in different industrial sectors use such data in completely different ways, and there is no all-

purpose income tax data base.

"In summary, the following are the key weaknesses of the current business statistics program, all fundamentally deriving from the lack of funding for programs with less than annual frequency:

- (1) lack of a complete and unduplicated list of businesses;
- (2) inadequate updating of the "profiles" of the larger businesses;
- (3) the inadequacy of procedures for ensuring that reporting units in the business universe are comprehensively and consistently classified by industry, on the basis of the most up-to-date classification system, and that new additions to the universe and changes in the status of existing units are promptly handled;
- (4) the inability of present systems to exploit tax records to their full potential as a substitute for direct data collection from smaller businesses, with the consequence that such records cannot in fact provide an adequate proxy for the survey data they are meant to replace; ....."(4).

### **BUSINESS SURVEY REDESIGN PROJECT**

#### Project Initiation

In early 1984, a Statistics Canada task force was created to address the general concerns regarding quality of data from the economic statistics program and the particular concerns about the Business Register. The task force proposed a new strategy for provision and use of frame and income tax data(7). Federal government Treasury Board approval and funding for the "Business Survey Redesign Project" was obtained in June 1984 and a project team was established two months later.

#### Objectives

The initial piece of work for the project team was to convert the task force report into general objectives and a comprehensive strategy acceptable to the wide range of survey and service operations which would be affected by their implementation. In the first project team report, the objectives were defined as follows.

General objectives for the development of an effective infrastructure for the conduct of surveys:

Identification of all economic entities to ensure total coverage of economic activity.

Development of a repository which will be used as the source of statistical units and will contain linked reporting units for all annual and sub-annual economic surveys.

Provision of frames that ensure no duplication or undercoverage within and between economic surveys.

Provision and ensured use of facilities to integrate and compare financial, production and labour statistics at the lowest level of statistical unit conceptually possible". (8,p.13).

Specific objectives of the Project Team:

Improvement of the accuracy and timeliness of industrial, geographic, and size coding of all units on the frames.

Improvement of the accuracy and timeliness of the statistical and reporting structures of all

significant and all multi-unit businesses.  
Extension of the coverage of the frame to units currently excluded.

Reduction of response burden and operational costs by more effective use of administrative data.

Improvement of procedures for the sampling, acquisition, SIC coding, transcription and weighting of tax data.

Subsequent to the Project Team report, the objectives were extended to include the complete redevelopment of the wholesale/retail trade survey program. In redesigning this program, procedures and systems will be generalized as far as possible so that they can be applied to the redevelopment of other surveys.

### Strategy

The strategy developed by the Project Team to meet its objectives is described in some detail in the following sections. In general terms, it requires:

"Enforced use of standardized concepts, definitions and procedures related to the maintenance of survey frames, data collection and acquisition of income tax data.

Clear definition of the role and responsibilities of each participant in the maintenance of the central frame and individual survey frames or mailing lists.

Standardization of the frame requirements of individual surveys in the context of a comprehensive economic statistics program.

Implementation of standard structures for complex businesses and their linkage to non-standard and associated reporting units.

Implementation of an audit function to ensure the integrity of the frame maintenance system". (8,p.14).

The strategy will be implemented within a framework of significant overall reduction in operating costs compared with current costs.

## **STRATEGY FOR FRAME DATA**

### Concepts and Standards

The starting point for a strategy for frame data is a clear definition of concepts and standards.

There will be five different structures on the central frame to represent the economic world:

1. administrative structure containing edited administrative data;
2. legal structure containing the legal image of the economic universe;
3. operating structure describing how the legal structure operates on a daily basis. This structure must be further processed to derive standard statistical records;
4. statistical structure which is the structure from which actual samples will be selected;
5. reporting structure identifying the reporting arrangements for each survey. It is developed by negotiation between survey managers and respondents, and is related to the statistical units appropriate to the survey.

In the statistical structure, there will be a hierarchy of four standard statistical units which, largest first, will be enterprise, company, establishment and location. All surveys of

employment, labour income and economic production, with the exception of monthly retail trade will be based upon the establishment. Monthly retail trade will be location based. Surveys of capital stocks and expenditures will use the establishment or the company. Surveys of financial and taxation statistics will be company based. The intercorporate ownership survey will define and utilize the enterprise.

There will be Bureau-wide policies and standards for industrial and geographic coding. The standards currently comprise:

- (a) 1980 Standard Industrial Classification(2) applied to the establishment as a statistical unit;
- (b) 1980 Company Classification applied to the company; and
- (c) 1981 Standard Geographical Classification applied to any unit.

Additional standards will be developed in the course of the Project.

### Creation and Maintenance of a Central Frame Data Base

Classified lists of statistical units and associated frame data items will be created and maintained to cover the whole universe of economic production. The lists will cover every unit in the universe, either individually, or by statistical sample, i.e. by a weighted subset of units which aggregate to the total. Reporting arrangements associated with the statistical units will be established for each survey. These data will be stored in a Central Frame Data Base (CFDB).

The CFDB will be maintained using administrative sources, information obtained from direct contacts made as part of routine survey data collection, and information obtained through direct frame enquiries. Standards, guidelines, procedures, software, reports and studies will be provided for the acquisition, processing, and quality assurance of frame data.

A Central Frame Data Base (CFDB) will be developed for the storage of all frame data items.

The CFDB will provide multiple frames for the following operational reasons. The principal administrative sources of frame data to the CFDB are the T1 and T2 income tax and the employer PD files from Revenue Canada. However, these files do not carry common identifiers. Their complete integration into a single frame would be expensive and technically difficult due to the large numbers and volatility of small businesses and the costs of tracking them through changes of ownership, amalgamations, mergers, etc.

The CFDB will be structured into an "integrated portion" and a "non integrated portion". The integrated portion of the frame will be a unique and unduplicated list of all large, complex, or otherwise significant statistical units with their associated administrative, legal, operating and reporting structure records. It will be derived through complete linkage and unduplication of the input sources. Maintenance will be comprehensive. All units will be followed through time.

The non-integrated portion of the CFDB will comprise multiple lists of all of the other units. These units, although the majority in terms of numbers, will account for only a small

fraction of the economic activity represented by the integrated portion and thus will not warrant the cost of detailed linkage and unduplication. The non-integrated portion will be divided into two and possibly three overlapping groups of statistical units. The first group will be derived from T1 and T2 source records and will be termed "income tax based statistical units". The second group will be derived from PD source records and will be termed "PD based statistical units". Consideration is presently being given to the introduction of an area frame, the units of which would then form a third overlapping group.

The division of the CFDB into an integrated portion and a non-integrated portion is a cornerstone of the strategy. It is illustrated simply in Appendix 1.

#### STRATEGY FOR INCOME TAX DATA

T1 (personal) and T2 (corporate) income tax data are available in essentially two forms:

- (a) Revenue Canada machine readable files, containing a subset of the reported data items, at no extraction cost, but with no explicit means of controlling quality; and
- (b) Tax return (or facsimile), including associated schedules and financial statements, from which all the data actually reported can be extracted; the quality of such data are determined by the quality of the extraction process (which can be controlled) and the quality of reporting (which cannot be controlled).

Income tax data will be used for two basic purposes, firstly as a building block in the construction of frames for annual surveys, secondly as a supplement or replacement to annual survey response data, as outlined in the following paragraphs.

All T2 tax filers, and all T1 filers reporting business, professional, farming, fishing, commission, or rental income, will be considered in scope for the economic survey program.

Records in Revenue Canada machine readable files above a specified size threshold will be associated with the integrated portion of the frame and will thus be linked to all corresponding units on other major administrative and statistical lists. Additional records from these files may also be linked to the integrated portion by virtue of their association with other units in this portion of the frame.

Tax data items used to supplement or replace survey data are not all available in machine readable form on Revenue Canada files thus data capture in Statistics Canada from the original returns is necessary. Because of the large number of returns, a sample must be selected. Sampling of the returns and extraction of selected items to meet survey requirements will be a central service function.

Each sampled return will be assigned a weight determined by the sampling procedure, adjusted to ensure that the sampling weights sum to the universe totals for a given reference year. The weighted returns will be used for estimation of counts and other universe totals and will form the "weighted sample file". This file in turn will provide a basis for estimation for units

not covered by the integrated portion of the frame.

Sampled T1 tax returns not in scope for the integrated portion of the frame, which refer to a partnership, will be assigned a "partnership weight" on the "weighted sample file". No attempt will be made to establish linkages to other partners, as the weight will compensate for the fact that the total activity of the business is reported on each partner's tax return, and is represented more than once in the universe.

Distinct statistical units will be created on the "weighted sample file" for each separate activity which is identified by a tax filer and for which separate statements are filed.

#### STRATEGY FOR ANNUAL SURVEYS OF PRODUCTION

Annual surveys of production form a particular class of surveys concerned with the collection and publication of "principal" and other production statistics for establishments(2). The basic objective of these surveys will be to provide structural information. It is the intention of the strategy to standardize procedures for these surveys so that collectively they may be regarded as constituting a single "annual survey of economic production".

These surveys will draw their frames from the CFDB and will provide feedback to the CFDB from survey results. Only in cases where alternative and better frame sources have been explicitly recognized and authorized by senior management will these be used.

A standard methodology for annual surveys has been proposed(9). It covers the frame, the collection strategy and the use of tax data to supplement or replace survey data as summarized in the following paragraphs.

For data collection purposes, the target population of establishments will be divided, essentially on the basis of size and operational significance, into three categories: "large", "small" and "out-of-scope". The large units will be maintained in the integrated portion of the CFDB. The complement of the frame will be provided by tax returns not associated with these large units.

Large establishments will be sampled annually. The sample will be mailed a full scale questionnaire requesting the complete range of financial, production and commodity statistics appropriate for the particular industry. Very large units will be sampled with certainty.

Small establishments will be those which are not included in the large category but which, measured in terms of a prespecified lower size threshold are large enough to be in scope. The threshold will be expressed in terms of one or more data items available in machine readable form from Revenue Canada for all tax returns, such as gross business income. A sample of these small establishments will be selected annually, the tax returns corresponding to the sample will be located and financial data will be obtained from them. A subsample of these units may be mailed a reduced questionnaire requesting a limited range of labour, production or commodity statistics. In both cases sampling

will be on a probability basis with sampling rates increasing with size.

Any economic production reported upon tax returns falling below the specified lower threshold will be defined as out-of-scope for survey. The corresponding tax returns will not be sub-sampled for mailout purposes. However, the data on such returns will provide the basis for estimating the effect of omitting such activity from explicit survey coverage.

Appendix 2 summarizes the frame for annual surveys.

#### STRATEGY FOR SUB-ANNUAL SURVEYS OF PRODUCTION

The basic objective of sub-annual surveys of economic production is to produce monthly estimates of changes in economic activities for various industrial and geographical aggregations thus, in principle, stepping annual benchmarks forward in time. In cases where no comparable annual data are available, sub-annual surveys will be designed to produce estimates of levels.

The frame of any given sub-annual survey will be drawn from the statistical units in the integrated portion, the PD based statistical units in the non-integrated portion and possibly an area frame. The PD based statistical units will be constantly updated based on information received from Revenue Canada. An area frame will be put in place if the feasibility study which is being conducted indicates it would be cost-beneficial.

In drawing a sample from the frame, very large units will be selected with certainty. It is the intention to develop generalized sample selection and mailout systems for all sub-annual surveys of economic production as part of the CFDB.

Survey questionnaires will be mailed to the sample units. Estimates will be produced monthly and revised annually to agree with annual survey estimates which will be considered benchmarks.

Appendix 2 summarizes the frame for sub-annual surveys.

#### ORGANIZATIONAL ARRANGEMENTS

A Central Service Function (CSF) will be created. It will be responsible for all services which can be more efficiently and effectively provided centrally than by individual survey operations and will replace and extend the existing Business Register and Tax Record Access operations. It will perform two distinct but interrelated functions:

1. The CSF(1) will manage the CFDB and will support: a full range of services including provision of a classified list of statistical and reporting units; the maintenance of these data using administrative and other sources; the provision of standards, guidelines and procedures; the sampling of survey universes and generation of mailing lists for surveys.
2. The CSF(2) will coordinate sampling of income tax data; and the acquisition, processing and use of such data to replace survey data.

#### Time Frame

The life span of the project will be three years. In very general terms, the activities will be scheduled as follows:

Phase I (until April 1986) will include specification of the complete CFDB system, implementation of income tax data subsystems for weighting and error estimation, and redesign of the wholesale/retail program;

Phase II (until March 1987) will include implementation of the CFDB and of all remaining income tax data subsystems, and development of systems for the wholesale/retail program;

Phase III (until March 1988) will see implementation of the Central Service Function, completion of wholesale/retail program development, and integration of all other economic surveys to the new systems and procedures.

#### Prototyping Approach

The "life cycle" approach to systems development, widely used over the last two decades, is along the lines: general statement of requirements, detailed statement, detailed design, systems proposal, pseudo coding, coding, testing, production and evaluation. In this project, by contrast, "prototyping" will be adopted for the major systems developments in particular for the CFDB. It is an approach in which "an early version of the system exhibits the essential features of the later operational system"(10). The prototype may evolve into the actual production system or may be used only for experimentation. Four CFDB prototypes, each displaying additional functions to the previous one, will be created before the final production system, see Appendix 3.

The benefits of prototyping are that it provides a tangible means for comprehending and evaluating the proposed system before it is built; it provides scope for experimentation and for hardening fuzzy requirements; and it helps maintain enthusiasm during a long development by providing real evidence of progress. The major drawback to prototyping is that, milestones and central procedures are less well established than with the conventional life cycle. It is hoped that defining an outline of what is to be achieved by each of the prototypes with a time scale, as in Appendix 3, will provide a basis for effective control.

#### Methodological Issues and Problems

Although the project strategy has been specified there are a number of specific issues to be addressed and problems of standardization to be solved (9,p.22-29) as indicated in the following paragraphs.

#### Concepts

The target population for annual surveys of economic production has yet to be precisely defined in terms of the categories of the self-employed income reported on an income tax return which are to be considered "in scope", and the minimum size threshold in each category. Methods for monitoring economic production below the thresholds have to be devised so that the extent of the corresponding undercoverage can be

estimated and periodic adjustments made to the thresholds when need be.

The identification of the four standard statistical units and their relationships to one another, to the administrative units and to the business operating and reporting units have not been finalized. The definitions of and procedures for specifying "standard updates" to these data items when stored in the CFDB have also to be fully developed. A particular example is the need to have a precise definition of establishment "continuity" over time to enable year-to-year micro-level comparison of such data for validation and imputation purposes. It must be stated under exactly what set of changes in size, location, activity, output, ownership, etc. an establishment is deemed to have "ceased to exist" as opposed to "continued".

#### Frame for Annual Surveys

The general framework for annual surveys of economic production (9) leaves the choice of boundary between full scale and reduced questionnaires still to be decided, balancing cost and quality considerations. The proposed procedures for income tax and reduced questionnaire sampling, and for merging these data with full scale questionnaire data have to be validated.

#### Framework for Sub-annual Surveys

The possibility of supplementing or even replacing employer PD data for frame maintenance with area frame data is being investigated. The area frame would comprise division of the country into a number of geographical areas which would be sampled on a rotational basis. The businesses within each sampled area would be enumerated and would provide coverage of non-employers and of new businesses not included in the PD based frame. Such a procedure would be an extension of that presently used for the retail trade survey. However, it would entail extensive changes to sub-annual survey procedures and systems already in place.

#### CFDB Boundaries

A central feature of the strategy is the division of the CFDB into integrated portion and non-integrated portion, the former being the only part common to both annual and sub-annual surveys. Criteria for specifying the integrated/non-integrated boundary and the rules for transfer across the boundary have yet to be established. The higher maintenance costs associated with a large integrated portion must be balanced against the improved frame data quality.

A related issue is the maintenance of industrial classification data in the non-integrated portion. One option is to maintain SIC codes for all units in the universe defined by T1 and T2 tax and employer PD data thus enabling pre- or post-stratification by industry. The alternative is to apply SIC codes only to a sample and use domain estimation for SIC breakdown. The balance is between reduced sampling variance but higher probability of bias due to poor quality, out-of-date classification in the first case versus increased sampling variance but reduced misclassification in the second case.

#### Income Tax Data Sampling, Acquisition and Storage

The proposed strategy for annual income tax data sampling (9 p. 10-13) involves definition of a general purpose "cross sectional" tax sample, covering all industries, in combination with various survey specific samples to swell sample sizes in certain industries which would otherwise be poorly represented. A multiple frame technique based on raking ratios (11) is proposed for estimation. This procedure which has been applied for a number of years in the annual construction survey has to be validated for application in other industries.

Procedures for acquisition of information from tax returns additional to that captured by Revenue Canada as part of the regular assessing process will be reviewed. It is possible that there are some efficiencies to be gained by rationalization and integration of Revenue Canada and Statistics Canada operations.

A data base system for processing and storage of tax data will be developed. It will replace and extend a number of separate processing modules which have been put in place over the past decade. The possibility of survey data feedback to the tax data base and use of the latter in combination with the CFDB for production of small area and small business data will be investigated.

#### Generalized Procedures and Systems

Every opportunity will be taken, especially during the redesign and implementation of the wholesale/retail program, to develop software and systems which can be generalized and utilized for future survey redesign. This will include systems for data capture, edit and imputation, confidentiality, etc.

#### Transition

A task as complex or even more complex than that of designing new systems for frame and income tax data processing is that of phasing them into the production environment. It is the present intention to avoid as far as possible "parallel runs" for production purposes and to aim for change over at a given point in time. It is not yet clear whether such an approach can guarantee the operational and the historical continuity required, in varying measures, by all surveys.

#### **CONCLUDING REMARKS**

A previous paper to the ASA Annual Meeting (1) described the problems of frame and income tax data acquisition and usage at Statistics Canada. This paper has outlined the strategy and some proposals for solution of these problems adopted by the Business Survey Redesign Project. By the Annual Meeting in 1988 the success or otherwise of the project in addressing the problems will be evident.

For more information on the project, the reader may contact the authors at:

Business Survey Methods Division,  
5th Floor, Jean Talon Building,  
Tunney's Pasture,  
Ottawa, Ontario, K1A 0T6, Canada.  
Telephone: 613-990-9430 (Michael Colledge)  
613-990-9433 (Robert Lussier)

**ACKNOWLEDGEMENTS**

The authors would like to acknowledge the contribution made by members of the Business Survey Redesign Project Team to contents of this paper.

**REFERENCES**

Copies of references may be obtained by application to the Business Survey Redesign Project Secretariat, Statistics Canada, Jean Talon Building 13-B, Ottawa, Ont. K1A 0T7, Can.

1. Colledge, M.J. (1983). "Frame Maintenance for Economic Surveys at Statistics Canada", Proceedings of the Section on Survey Research Methods, American Statistical Association.
2. Statistics Canada (1980). "Standard Industrial Classification 1980", Catalogue 12-501.
3. Estevao, V. (1985). "A Report on the Quality of Data in the BRMF", Statistics Canada.
4. Statistics Canada (1984). "The Cyclical Requirements for Maintaining Statistics Canada's Program of Business Statistics".
5. Statistics Canada (1973). "Policy for the Implementation and Use of the Business Register and for the Conduct of Business Profiling", Statistics Canada.
6. Colledge, M.J., Estevao, V., Lussier, R. (1985). "Quality Assurance of Business Register Data, presented to the Statistical Society of Canada Annual Meeting, June 1985.
7. Ansmits, A., Colledge, M., Monty, A., Sande, I., Winkworth, A. (1984). "Strategy for the Provision of Frame Data and Use of Income Tax Data in Economic Statistics", Statistics Canada.
8. Infrastructure Project team (1984). "Business Survey Redesign Project: Phase I - Infrastructure Development Objectives, Policies and Strategy", Statistics Canada.
9. Colledge, M.J. (1985). "Annual Surveys of Economic Production: Methodology Framework", Statistics Canada.
10. Sprague, R.H. and Carlson, E.D (1982). "Building Effective Decision Support Systems". Prentice-Hall, Englewood Cliffs, N.J.
11. Bankier, M.B. (1983). "Estimators Based on Several Stratified Samples with Applications to Multiple Frame Surveys", Proceedings of the Section on Survey Research Methods, American Statistical Association.

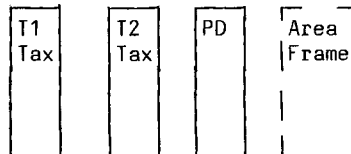
**Appendix 1: Central Frame Data Base (CFDB)**  
**Integrated and Non-Integrated Portions**

Integrated  
Portion



(Includes all large or otherwise significant statistical and reporting units, linked and followed through changes of ownership etc.)

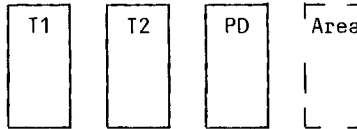
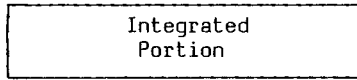
Non-  
Integrated  
Portion



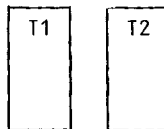
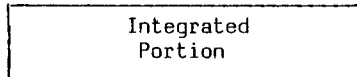
(Includes the remainder of the universe; not linked, not followed through changes; with duplication between T1/T2 tax and PD files)

Appendix 2: Frames for Annual and Sub-Annual Surveys

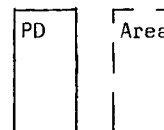
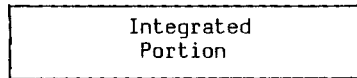
CFDB



Frame for  
Annual Surveys



Frame for  
Sub-annual Surveys





## Appendix 3: Description of the CFDB Prototypes

### Prototype No. 1

- Content of the CFDB
- Records, Structures or Relationships & Data Elements
- Functions: Load, Store and Display Data
- Sample of Records
- Delivery Date: June 1985

### Prototype No. 2

- Upgraded Version of Prototype No. 1
- Standard Updating Procedures
- Functions: Proto 1; Add, Update and Delete Records & Data: Identification of Segments (Integrated, Non-integrated); Structure Audit
- Sample of Records
- Delivery Date: October 1985

### Prototype No. 3

- Upgraded Version of Prototypes No. 1 & No. 2
- Processing of Administrative Sources of Data (T2, T1, & PD)
- Processing of Profiling Updates (Including Survey Feedback)
- Functions: Proto 1; Proto 2; Survey Interfaces
- Sample of Records
- Delivery Date: February 1986

### Prototype No. 4

- Upgraded Version of Prototypes No. 1, No. 2 & No. 3
- Final Prototype
- Functions: All Functions Including Sampling and Population Estimates
- All Records
- Delivery Date: July 1985